Lecture notes on Quantum Statistical Inference

October 15, 2024

Contents

Chapter 1

Quantum Hypothesis Testing Part I: State discrimination

1.1 Introduction

In our every day lives we are constantly confronted with the necessity of having to take a decision among alternative possibilities. Our decision, naturally, is based on the information available and the processing we do of it. Often, the information at our disposal and our intuition lead us to take the correct option. However, more often than not we wish we had more information and we beg for more data that perhaps are impossible to obtain. In situations where it is possible to systematically accumulate data (e.g. with experiments) a formalisation and and a rigorous theoretical analysis can increase the possibilities of a correct decision. In this scenario, we would like to answer questions such as, given some available data, what is the best decision that can taken as well as well as to quantify the probabilities of error or success. Equivalently, given some required precision we may want to know how many data need to be gathered to achieve it. The theory behind these problems is named statistical inference. It started to be systematically developed in the middle of last century and had since profound impact in areas that range from medicine, to sociology or economy. It also plays a crucial role in the certification of any scientific discovery. At that time quantum mechanics was already a mature theory able to provide a plethora of applications in virtually all scientific fields and technology branches. In the last decades, with the advent of the theory of Quantum Information, new and revolutionary applications have appeared. The field of statistical inference has been deeply revised and has incorporated the distinctive features of quantum mechanics. Statistical quantum inference has provided again a profound impact in the developments of quantum communication protocols (as e.g. quantum cryptography), metrology or sensing applications. On the theoretical side also many new results have been discovered recently.

In these lectures we will provide a comprehensive introduction to this topic. The first part is devoted to discrimination of quantum states. We first review the classical results of hypothesis testing and introduce some of the specific language. We then move to the the quantum setting and present the minimum error problem, Helstrom measurements, Holevo conditions and unambiguous discrimination. We then give a brief introduction to semidefinite program (SDP) techniques, which are specially suited for quantum state discrimination tasks, but actually had their origins in optimisation problems outside the quantum realm. The Holevo conditions are re-derived in this formulation. We also show that for pure states the matrix of overlaps (Gram matrix) encapsulates all the discrimination properties of the set of the states. In the second chapter asymptotic results of the error probabilities and celebrated error exponents in the classical and quantum realm are obtained. Then the discrimination of quantum channels and applications is discussed. The last two chapters deal with the estimation problem. In the third chapter we present the fundamentals of the classical the frequentist and Bayesian approaches. We then analyze the Quantum case. The last chapter is mainly devoted to state tomography, i.e., the task of estimating a quantum state (or some features of it) from quantum systems. Single shot and sequential protocols are discussed as well as maximum likelihood and confidence regions. Some alternative schemes for quantum state tomography are also presented. Technical details not covered in the lectures, but necessary to make this notes self-contained are presented in appendices of each section.

1.2 Classical hypothesis testing

The most elementary decision problem consists of choosing between two hypothesis H_0 and H_1 from data gathered from measurements or tests. For instance, one may have to decide if a patient is healthy (H_0) or has a given illness (H_1) from the results of a clinical test. One may also want to assess if a drug has an effect (H_1)or no effect (H_0). Usually H_0 is called the null hypothesis, specially in clinical trials, where it corresponds to the hypothesis that the standard treatment of some disease is more effective, while H_1 is termed alternate hypothesis, generally associated to an experimental treatment one wants to test. In general, the two hypothesis may not be considered in a symmetric way. For instance, in a doping test if an athlete is not doped and the test gives a positive answer, this situation may ruin his/her career. However, if a doped athlete tested negative, it is considered a failure of the test, but has much less serious consequences. The first case is a false positive, that is, the test decides the alternate hypothesis when actually the null hypothesis is the true one. The second case is named false negative. These errors are dully known in the statistical inference literature as Type I and Type II errors, respectively. The corresponding error probabilities can be written as

$$\begin{aligned} \text{Type I} \quad \alpha &= p(\hat{H}_1|H_0) \\ \text{Type II} \quad \beta &= p(\hat{H}_0|H_1), \end{aligned} \tag{1.2.1}$$

where the hat indicates that $\hat{H}_i$ is the guess inferred from the test.

Very much related to Type I error is the so-called p-value. Roughly speaking is the probability of obtaining the observed outcomes under the null hypothesis. The smaller this probability is the larger is our belief that that the null hypothesis is not true. This concept is not very easy to grasp and has often been misused in statistical inference, up to the point that more than 800 scientists have recently written a manifesto complaining about its wrong use in many publications, and even suggesting removing it from the scientific literature [1]. Let us nevertheless give a glimpse of this notion with an example without entering into too many details, as it is beyond the scope of these lectures.

Let a coin be tossed a number of times. The null hypothesis is that the coin is fair. From the outcomes one can construct a *statistic* that encapsulates (i.e. summarises) a relevant property of the observed data. In our example the statistic can simply be the number T of heads (but one can consider other statistics, e.g. the number of two consecutive heads, etc). The p-value is the probability $\Pr[T > t|H_0]$, i.e.the probability of observing at least t heads. In many cases it is more meaningful to consider the two sided p-value defined as $2\min\{\Pr[T \geq t|H_0], \Pr[T \leq t|H_0]\}$. If the distribution of the statistic T is symmetric around $T = 0$, the two-sided p-value can be written as $\Pr[|T| \geq |t||H_0]$. If we want to test if a coin is fair or not, indeed it is more meaningful to consider the statistic that counts the absolute value of the difference between the number of heads and tails observed. A large deviation from a 50% heads of the observed tosses will entail a very small p-value. Conventionally $p = 0.05$ is considered the value to claim a disproval of the null hypothesis.

The problem with p-values is that often is confused with the probability that the null hypothesis is true, or the probability that the alternate is wrong. The p-value should be viewed as a measure of compatibility between the observed data and the model described by the null hypothesis. A rejection of the null hypothesis based on p-value below 0.05 only indicates that the observed data are fairly incompatible with the model, it does not prove that the alternate hypothesis is true, as the possibilities of alternates hypotheses are immense in general.

Exercise 1 *A coin is tossed 100 times. We obtain 60 heads. Check that the symmetric p-value is* 0.057*, (take into account both tails). Thus, the fair coin hypothesis would not be rejected. Check that if we toss one more time and the outcome is heads, then the fair coin hypothesis would be rejected under the conventional rules. Just for fun. Check that if one tosses one more time and obtains tails, the fairness hypothesis is not rejected again.*

Let's go back to Type I and Type II errors. It would be desirable to devise procedures that minimise both type of errors. However this is not possible in general as the reduction of one type of error involves the increase of the other. Hence, a common approach is to consider a minimisation of a linear combination of both type of errors:

$$P_e = \eta_0 p(\hat{H}_1|H_0) + \eta_1 p(\hat{H}_0|H_1), \tag{1.2.2}$$

or, equivalently maximise the success probability

$$P_s = \eta_0 p(\hat{H}_0|H_0) + \eta_1 p(\hat{H}_1|H_1), \tag{1.2.3}$$

where η_i can be viewed as the prior probabilities of occurrence of each hypothesis. The quantity P_e (P_s) is the average error (success) probability and the goal is to minimise (maximise) P_e (P_s). This procedure is called minimum error approach.

In the classical case the optimization of Eq. (1.2.3) and its interpretation is rather straightforward. Let's consider again the coin case. The test is simply a toss and the outcome is either heads or tails. If one obtains heads, the best guess is the coin that has the highest likelihood given this outcome (see below for an elementary proof), i.e., it is determined by the condition $\max\{p(H_0|\text{heads}), p(H_1|\text{heads})$ and likewise for tails. The overall success probability is then

$$P_s = p(\text{heads}) \max\{p(H_0|\text{heads}), p(H_1|\text{heads})\} + p(\text{tails}) \max\{p(H_0|\text{tails}), p(H_1|\text{tails})\}, \tag{1.2.4}$$

where $p(\text{heads})$ and $p(\text{tails})$ are the total probabilities of obtaining heads and tails, respectively. We can rewrite (1.2.4) in terms of joint probabilities as

$$P_s = \max\{p(H_0, \text{heads}), p(H_1, \text{heads})\} + \max\{p(H_0, \text{tails}), p(H_1, \text{tails})\}, \tag{1.2.5}$$

and by using Bayes theorem

$$p(r)p(h|r) = p(h)p(r|h) = p(r, h), \tag{1.2.6}$$

easily get a more useful expression in terms of the prior probabilities of each hypothesis $p(H_i) = \eta_i$ and the conditional outcome probabilities $p(r|H_i)$:

$$P_s = p(H_0) \max\{p(\text{heads}|H_0,), p(\text{heads}|H_1)\} + p(H_1) \max\{p(H_0, \text{tails}), p(H_1, \text{tails})\}. \tag{1.2.7}$$

Denoting the conditional outucome probabilities by $p(\text{heads}|H_0) = p_0, p(\text{tails}|H_0) = 1-p_0$ and $p(\text{heads}|H_1) = p_1, p(\text{tails}|H_1) = 1 - p_1$. we can finally write

$$P_s = \max\{\eta_0 p_0, \eta_1 p_1\} + \max\{\eta_0 \bar{p}_0, \eta_1 \bar{p}_1\} \tag{1.2.8}$$

where $\bar{p}_i = 1 - p_i$.

The seemingly natural statement that for a given outcome the optimal choice is the one corresponding to the maximum likelihood, can be rigorously proven for any two probability distributions with d outcomes (e.g. we can think of two roulettes) [2]. We denote the outcomes by $r = 1, 2, \ldots, d$ and define a decision function $\Theta(r) \to \{0, 1\}$, where $\{0, 1\}$ denote our decision in favour of H_0 or H_1, respectively. The error probability simply reads

$$P_e[\Theta] = \eta_0 \Pr[\Theta = 1|H_0] + \eta_1 \Pr[\Theta = 0|H_1] \tag{1.2.9}$$

i.e. the weighted sum that we decided in favour of H_1 (H_0) when the true hypothesis was H_0 (H_1). The natural decision function is the one that given an outcome r chooses the hypothesis with higher posterior probability as given by the Bayes rule

$$\Pr[H = i|r] = \frac{\eta_i \Pr[r|H_i]}{\eta_0 P[r|H_0] + \eta_1 \Pr[r|H_1]} = \frac{\eta_i \Pr[r|H_i]}{p(r)} \tag{1.2.10}$$

i.e. the Bayes decision rule is

$$\Theta(r) = \begin{cases} H = 0 & \text{if} \quad \eta_0 \Pr[r|H_0] > \eta_1 \Pr[r|H_1] \\ H = 1 & \text{if} \quad \eta_1 \Pr[r|H_1] > \eta_0 \Pr[r|H_0] \\ \text{random} & \text{if} \quad \eta_0 \Pr[r|H_0] = \eta_1 \Pr[r|H_1] \end{cases} \tag{1.2.11}$$

Then the error probablitiy (1.2.9) writes

$$P_e[\Theta] = \eta_0 \sum_{r=1}^{d} \Theta(r) \Pr[r|H_0] + \eta_1 \sum_{r=1}^{d} [1 - \Theta(r)] \Pr[r|H_1] \tag{1.2.12}$$

We now prove that (1.2.11) is the best decision rule. For any other decision rule Θ', we have

$$P_e[\Theta'] - P_e[\Theta] = \eta_0 \sum_{r=1}^{n} \Big(\Theta'(r) - \Theta(r)\Big)\Big(\eta_0 \Pr[r|H_0] - \eta_1 \Pr[r|H_1]\Big) \tag{1.2.13}$$

If the decision functions are different, there are outcomes r such that $\Theta'(r) \neq \Theta(r)$. E.g., if $\Theta'(r) = 0$ and $\Theta(r) = 1$ and $\eta_0 \Pr[r|H_0] - \eta_1 \Pr[r|H_1] < 0$, the term is positive. The same analysis applies if $\Theta'(r) = 1$ and $\Theta(r) = 1$ with $eta_0 \Pr[r|H_0] - \eta_1 \Pr[r|H_1] > 0$. Hence

$$P_e(\Theta') > P_e(\Theta), \tag{1.2.14}$$

as we wanted to demonstrate.

Exercise 2 *Demonstrate that for equal priors the success probability of discriminating tow coins with biases p_0 and p_1 after one toss can be written as*

(a) $P_s = \dfrac{1}{2}\left[1 + \max\{p_0, p_1\} - \min\{p_0, p_1\}\right]$

(b) $P_s = \dfrac{1 + |p_0 - p_1|}{2}$.

(c) Give the generalization of (b) for arbitrary priors η_0 and η_1.

Exercise 3 *Consider arbitrary priors η_0 and η_1, is it possible that for some values of the priors the guess is independent of the outcome obtained? If this is possible, state the conditions and give the success probability in this case.*

The straightforward generalisation of (1.2.8) for n distributions (i.e. hypotheses) of d outcomes with probabilities $\Pr[\text{Outcome} = r|H_i] = p_i(r)$, $i = 0, 1, \ldots, n$ and $r = 0, 1, \ldots, d-1$ is

$$P_s = \sum_{r=0}^{d-1} \max\{\eta_0 p_0(r), \eta_1 p_1(r), \ldots, \eta_n p_n(r)\} \tag{1.2.15}$$

Note that a repeated test can be viewed as a single shot experiment with larger number of outcomes (more on this in the asymptotic section).

Exercise 4 *Consider the following two pairs of hypotheses (i.e. four coins) distributed with probability distributions $H_0 \sim \{0.96, 0.04\}$, $H_1 \sim \{0.04, 0.96\}$, and $J_0 \sim \{0.9, 0.1\}$, $J_1 \sim \{0, 1\}$, where within each pair the prior probabilities are equal.*

(a) Check that the success probabilities oof getting the right answer in the the first and the second pair fulfil $P_s(H_0, H_1) > P_s(J_0, J_1)$. Hence you may be tempted to say that the first pair of coins is more distinguishable than the second.

(b) Consider however tossing the coins twice. Obtain the corresponding probability distributions and check that then $P_s(H_0^2, H_1^2) < P_s(J_0^2, J_1^2)$. So, paradoxically tossing twice the first pair becomes less distinguishable than the second! Discuss this result.

1.3 Discrimination of quantum states

In the quantum realm hypotheses correspond to quantum states and the hypothesis testing task becomes a discrimination problem between quantum states. The crucial difference with respect to the classical case is that in quantum mechanics to obtain information about the states, one needs to perform a measurement, mathematically described by a Positive Operator Value Measure (POVM) that can be chosen to optimise the success probability of discrimination. Here the playground is much richer. First, in the single shot scenario the optimal measurement is already a non trivial task (recall that in the

classical case there was no freedom one could only observe). Second, if more than one identically prepared system [1] are available for measurement one can consider measuring all the systems with the most general measurement, i.e. a single shot approach, or one may consider individual measurements on each system which can be identical or can be modified at each step to include the information gathered from previous outcomes in a sort of adaptive procedure. This procedure again can be Markovian, in the sense that only the information of previous step is included in the next, non-Markovian, etc. Third, one my also consider two-way measurements, i.e. one system is weakly measured, with the information (usually small) acquired from the outcome one goes back to previous systems, and then forward again. In spite of the complexity described in some cases it is possible to find analytical results or useful bounds for the success probabilities. Finally, in some cases with appropriate quantum measurements it is possible to identify the state with absolute certainty, i.e. without error. Again, one can consider global vs local measurements, adaptivity, etc.

We first tackle the binary case and present the Helstrom measurements. We then focus on the pure state case and consider the situation when more than one system is available for measurement. The asymptotics of this case is discussed in Chapter II of these lectures. Next we analyze the discrimination problem for more than two states and obtain the optimality conditions, a.k.a, Holevo conditions. We then make an interlude and present the basic notions of semidefinite programming (SDP)[3, 4], a technique employed in many optimisation problems that proves to be extremely useful in discrimination problems. We rederive some of the previous results using SDP and address other cases with this technique.

1.4 Minimum error: binary case

Let hypothesis H_0 and H_1 correspond to a quantum system that either has been prepared in state ρ_0 or ρ_1 and the prior probabilities of occurrence are, as in the classical case, assumed to be η_0 and η_1, respectively. We have to measure the system to gain information about the identity of the state. The measurement can in principle have many outcomes, but at the end they will be lumped in two groups: those indicating state ρ_0 and those indicating state ρ_1. Therefore, effectively the POVM has only two elements E_0 and E_1, which satisfy $E_i \geq 0$, $i = 0, 1$, and $E_0 + E_1 = \mathbb{1}$. The Born rule gives the conditional probabilities $p(E_i|\rho_j)$ associated to a measurement. When no confusion arises, we simply denote them as $p(i|\rho_j)$, $i, j = 0, 1$. They read $p(i|\rho_j) = \mathrm{tr}(E_i\rho_j)$ and with all what we have already learnt [see Eq. (1.2.4)] we have that the success probability reads

$$P_s(E) = \eta_0 \,\mathrm{tr}(E_0\rho_0) + \eta_1 \,\mathrm{tr}(E_1\rho_1), \tag{1.4.16}$$

where E reminds the dependence on the measurement. The maximum success probability is obtained when the measurement is optimised over all possible measurements. Let's first solve the symmetric case $\eta_0 = \eta_1$,

$$\begin{aligned} P_s = &\max_E \ \frac{1}{2}\left[\mathrm{tr}(E_0\rho_0) + \ \mathrm{tr}(E_1\rho_1)\right] \\ &\max_{0\leq E_0\leq \mathbb{1}} \frac{1}{2}\left\{1 + \mathrm{tr}[E_0(\rho_0 - \rho_1)]\right\} \\ &= \frac{1}{2}\left(1 + \frac{1}{2}||\rho_0 - \rho_1||_1\right) \end{aligned} \tag{1.4.17}$$

In going from the first to the second line we have simply used the POVM conditions for E_o and E_1. In the third line we have used that the optimal operator E_0 has to be the the projector onto the positive spectrum of $\Gamma = (\rho_0 - \rho_1)/2$, for any projection on the negative spectrum would only decrease the value fo the trace. In the diagonal basis of Γ, the operator E_0 is diagonal with ones in the slots corresponding to the positive eigenvalues of Γ and zeroes in the rest. I.e. it "reads" the positive eigenvalues of Γ . Since $\mathrm{tr}\,\Gamma = 0$ the sum of the negative eigenvalues must be the same in absolute value. The sum of the absolute values of an hermitian operator A is named trace norm and denoted as $||A||_1$. Often the double bar and the subindex 1 is omitted. The general form of the trace norm valid also for non-hermitian operators is

$$||A||_1 = \mathrm{tr}\,\sqrt{A^\dagger A} \tag{1.4.18}$$

[1] Some relaxation of this i.i.id. condition will be considered later in these lectures

Exercise 5 *If the states to be discriminated are pure*

(a) Argue that without loss of generality they any two pure states can be written as $|\psi_0\rangle = \cos\theta/2\,|0\rangle + \sin\theta/2\,|1\rangle$ *and* $|\psi_1\rangle = \cos\theta/2\,|0\rangle - \sin\theta/2\,|1\rangle$, $0 \leq \theta \leq \pi/2$. *Compute the overlap* $c = \langle\psi_0|\psi_1\rangle$.

(b) Compute the success probability for equal priors and write it terms of the overlap

Exercise 6 *For two qubit mixed states* $\rho_i = \dfrac{\mathbb{1} + \mathbf{r}_i \cdot \sigma}{2}$, $i = 0, 1$, *that can occur with equal priors,*

(a) Compute the success probability in terms of the Bloch vectors $\mathbf{r}_i$.

(b) Check that you recover the results of the previous exercise for pure states. Compute the success probability if one of the states is pure and the other is completely depolarised. This result gives you an idea of how well one can discriminate with respect to a completely unknown state.

Exercise 7 *A completely unknown qubit state can be represented as the statistical mixture*

$$\int \frac{d\Omega}{4\pi} |\psi(\Omega)\rangle \langle\psi(\Omega)|,$$

with $|\psi(\Omega)\rangle = \cos\theta/2\,|0\rangle + e^{i\phi}\sin\theta/2\,|1\rangle$ *and* $d\Omega = \sin\theta d\theta d\phi$

(a) Check that indeed this mixture gives the state $\mathbb{1}/2$

(b) Compute the density matrix corresponding to two unknown but identical qubits

(c) Compute the density matrix corresponding to two unknown qubits that point in opposite directions

(d) What is the success probability of distinguishing if we have been given identical or opposite qubits with equal priors?

(e) Can you obtain this probability doing almost no computations? Hint: Integrals have to be proportional to projectors (identities) in subspaces (a corollary of Schur lemma)

Let's finally solve the case for general priors η_0 and η_1 Using the same reasoning as in Eq. (1.4.17)

$$\begin{aligned} P_s = &\max_E \ \mathrm{tr}(E_0\eta_0\rho_0) + \ \mathrm{tr}(E_1\eta_1\rho_1) \\ &\max_{0\leq E_0\leq \mathbb{1}} \eta_1 + \mathrm{tr}[E_0(\eta_0\rho_0 - \eta_1\rho_1) \end{aligned} \tag{1.4.19}$$

which tells us that the optima E_0 is the projector onto the positive spectrum of $\Gamma = \eta_0\rho_0 - \eta_1\rho_1$. Note that we also have

$$\begin{aligned} P_s = &\max_E \ \mathrm{tr}(E_0\eta_0\rho_0) + \ \mathrm{tr}(E_1\eta_1\rho_1) \\ &\max_{0\leq E_1\leq \mathbb{1}} \eta_0 - \mathrm{tr}[E_1(\eta_0\rho_0 - \eta_1\rho_1) \end{aligned} \tag{1.4.20}$$

which, as expected, ells us that the optima E_1 is the projector onto the negative spectrum of Γ. Adding both expressions and taking into account that $\eta_0 + \eta_1 = 1$ we get

$$P_s = \frac{1}{2} + \frac{1}{2}||\eta_0\rho_0 - \eta_1\rho_1||_1 \tag{1.4.21}$$

The optimal measurement for minimum error discrimination of two quantum states is often referred to Helstrom measurement.

Exercise 8 *Justify that a Helstrom measurement is a von Neumann measurement.*

Exercise 9 *What is the value of P_s when ρ_0 and ρ_1 have no common support, i.e are orthogonal $\rho_0 \cdot \rho_1 = 0$. What are E_0 and E_1 in this case?*

Exercise 10 *Find an example of priors and state such that Γ has no negative spectrum. What is E_0 and E_1 in this case? Give an interpretation of this somewhat paradoxical result. Demonstrate that for pure states there no priors that leas to this result. If the sates are mixed, do always exist priors that yield this result?*

Exercise 11 *When ρ_0 and ρ_1 are pure qubit states*

(a) Check that the success probability is

$$P_s = \frac{1+\sqrt{1-4\eta_0\eta_1 c^2}}{2}, \quad c^2 = |\langle\psi_0|\psi_1\rangle|^2 \tag{1.4.22}$$

(b) Argue that this expression is valid for states in any dimension

(c) Compute the success probability when the hypotheses correspond to a set of N identical copies of pure states: $|\Psi_0\rangle = |\psi_0\rangle^{\otimes N}$ and $|\Psi_1\rangle = |\psi_1\rangle^{\otimes N}$

(d) Compute the success probability is the set is not of identical copies, i.e., $|\Psi_0\rangle = |\psi_1\rangle|\psi_2\rangle\cdots|\psi_N\rangle$ and $|\Psi_1\rangle = |\phi_1\rangle|\phi_2\rangle\cdots|\phi_N\rangle$

Exercise 12 *Bayesian updating. From previous exercise compute the success probability for two copies of pure states with equal priors*

(a) Check that a succession of the following two local measurements gives the same success probability. Measure the first copy with the optimal measurment. Measure the second measurement with updated priors, if the first outcome was 0 then $\eta_0' = P_s$ if was 1 then $\eta_1' = P_s$. Does this procedure give the optimal success probability for any n?. Give an interpretation

(b) From the results of Exercise 10, argue why this procedure will not work for mixed states.

1.5 Minimum error for multi-hypotheses

Going beyond two states, although at first sight looks as a simple generalisation, it is a much harder problem and no general closed expressions exist, except in symmetric cases. In this general scenario we have a source that can prepare states ρ_i with prior probabilities η_i, $i = 1, 2, \ldots, n$ and, as before, we want to find the procedure that gives the maximum success probability of correct identification of the state that has been prepared. Naturally, as in previous sections, If the systems are composed, one can consider single (global) shot as well as multiple shot measurements with all their variants.

The success probability in this case reads

$$P_s = \max_E \sum_i \eta_i \operatorname{tr} E_i\rho_i, \quad \text{s.t. } E_i \geq 0 \ \text{ and } \ \sum_i E_i = \mathbb{1} \tag{1.5.23}$$

(sorry for the change of notation of the POVM operators, I clean it in next version of these notes). Although no analytical results giving the optimal measurement exit, there exist necessary and sufficient conditions that allow to prove if some ansatz is optimal or not. These are known as Holevo conditions.

1.5.1 Holevo conditons

Theorem 1 *The optimal discrimination of states given by the ensemble $\mathcal{E} = \{\eta_i, \rho_i\}$ or equivalently $\mathcal{E} = \{\tilde{\rho}_i\}$ with $\tilde{\rho}_i = \eta_i\rho_i$ is provided by a quantum measurement $\{E_i\}_{i=1}^n$ if and only if [5]*

$$\boxed{Y = \sum_{i=1}^{n} E_i \tilde{\rho}_i \qquad \textit{satisfies} \qquad Y - \tilde{\rho}_i \geq 0, \quad i = 1, \cdots, n} \tag{1.5.24}$$

Notice that probability of success reads

$$P_s = \operatorname{tr} Y \tag{1.5.25}$$

and that Y is hermitian, i.e., $Y = Y^\dagger = Y$, or $\sum_{i=1}^n E_i \tilde{\rho}_i = \sum_{i=1}^n \tilde{\rho}_i E_i$, because Y is a semidefinite positive operator.

This theorem is derived in appendix

Corollary. The following two properties follow from Eq.(1.5.24) :

$$\boxed{Y \geq 0 \qquad \text{and} \quad E_i(\tilde{\rho}_i - \tilde{\rho}_j)E_j = 0 \quad \forall i, j} \tag{1.5.26}$$

Exercise 13 *Easier derivation of the sufficiency of the Holevo's conditions. Consider another POVM* $\{E'_i\}$ *and* $Y' = \sum_j E'_j \tilde{\rho}_j$

(a) Argue that optimality implies that $\operatorname{tr} Y - \operatorname{tr} Y' \geq 0$ *and use the closure relation* $\sum_j E'_j = \mathbb{1}$ *to obtain the condition*

$$\sum_j \operatorname{tr}\left[E'_j \left(Y - \tilde{\rho}_j\right)\right] \geq 0 \tag{1.5.27}$$

(b) Prove that for $A \geq 0$ *and* $B \geq 0$ *it follows that* $\operatorname{tr}[AB] \geq 0$. *(Does it follow that* $AB \geq 0$*?).*

(c) From the trivial identity $\mathbb{1}Y = Y$ *finally obtain that*

$$Y - \tilde{\rho}_j \geq 0, \quad j = 1, 2, \ldots, n \tag{1.5.28}$$

is a sufficient condition.

Exercise 14 *Re-derivation of the corollary Eq.* (1.5.26) .

(a) Compute $\operatorname{tr}[\sum_j (Y - \tilde{\rho}_j)E_j]$

(b) Prove that if $A, B \geq 0$, $\operatorname{tr} AB = 0 \Leftrightarrow AB = 0$

(c) Derive the conditions $(Y - \tilde{\rho}_j)E_j = 0$ *and* $E_j(Y - \tilde{\rho}_j) = 0$ *for all* $j = 1, 2, \cdots, n$ *to obtain*

$$E_i(\tilde{\rho}_i - \tilde{\rho}_j)E_j = 0 \quad \forall i, j$$

1.5.2 Symmetric states and square root measurement

There exist only few cases of multi-hypothesis discrimination that can be solved exactly. These comprise sets of N pure states that have equal prior probability $\eta_i = 1/N$ (i.e., there is no preference of occurrence towards any particular state) and satisfy

$$|\psi_{k+1}\rangle = U |\psi_k\rangle, \quad k = 0, 1, \cdots, N-1 \quad \text{with} \quad U^N = \mathbb{1}, \tag{1.5.29}$$

where U is a unitary operation.

Exercise 15 *Let* $\Omega = \sum_k |\psi_k\rangle\langle\psi_k|$, *prove that* $U\Omega U^\dagger = \Omega$ *and* $[U, \Omega] = 0$.

We consider next the square root measurement defined by the elements

$$E_k = \Omega^{-1/2} \left|\psi_k\right\rangle\!\left\langle\psi_k\right| \Omega^{-1/2} \tag{1.5.30}$$

Note that indeed $E_k \geq 0$ and $\sum_k E_k = \mathbb{1}$. (This measurement is also known as pretty good measurement and for general states and priors is defined as $\Omega = \sum_k \eta_k \rho_k$ and $E_k = \eta_k \Omega^{-1/2} \rho_k \Omega^{-1/2}$). To prove its optimality we simply check that the Holevo conditions are satisfied, i.e. has to prove that

$$A_j = N(Y - \tilde{\rho}_j) = \sum_k E_k \left|\psi_k\right\rangle\!\left\langle\psi_k\right| - \left|\psi_j\right\rangle\!\left\langle\psi_j\right| \geq 0 \tag{1.5.31}$$

that is, for any state $|\phi\rangle$ one has $\langle\phi| A_j |\phi\rangle \geq 0$. First note that using the results of Exercise 15 one can easily prove that $\langle\psi_k| \Omega^{-1/2} |\psi_k\rangle = \langle\psi_0| \Omega^{-1/2} |\psi_0\rangle$, thus

$$\sum_k E_k \left|\psi_k\right\rangle\!\left\langle\psi_k\right| = \langle\psi_0| \Omega^{-1/2} |\psi_0\rangle \, \Omega^{1/2} = \langle\psi_j| \Omega^{-1/2} |\psi_j\rangle \, \Omega^{1/2}$$

$$\langle\phi| A_j |\phi\rangle = \langle\psi_j| \Omega^{-1/2} |\psi_j\rangle \, \langle\phi| \Omega^{1/2} |\phi\rangle - |\langle\psi_j|\phi\rangle|^2 \tag{1.5.32}$$

Using Schwarz inequality for the states $\Omega^{-1/4} |\psi_j\rangle$ and $\Omega^{-1/4} |\phi\rangle$, one gets

$$\langle\phi| A_j |\phi\rangle \geq \langle\psi_j| \Omega^{-1/4} \Omega^{1/4} |\phi\rangle \, |^2 - |\langle\psi_j|\phi\rangle|^2 = 0 \tag{1.5.33}$$

Note that the conditional success probability is the same for all states and the overall success probability then reads

$$P_s = |\, \langle\psi_0| \Omega^{-1/2} |\psi_0\rangle \, |^2 = |\, \langle\psi_k| \Omega^{-1/2} |\psi_k\rangle \, |^2 \quad (\text{for any } k = 0, 1, \ldots, N-1) \tag{1.5.34}$$

Exercise 16 *Compute the success probability of discrimination fort the set* $\{|\psi_k\rangle\}_{k=0}^{N-1}$ *with equal priors* $\eta_k = 1/N$ *and*

$$\frac{|0\rangle + e^{i\frac{2\pi}{N}k} |1\rangle}{\sqrt{2}}$$

1.5.3 Pure states and the Gram matrix formulation

If the discrimination set is comprised by pure states $\mathcal{E} = \{\eta_i, |\psi_i\rangle\}_{i=1}^N$, then all the discrimination properties are encapsulated in the Gram matrix G whose elements are

$$G_{ij} = \sqrt{\eta_i}\sqrt{\eta_j} \langle\psi_i|\psi_j\rangle = \langle\tilde{\psi}_i|\tilde{\psi}_j\rangle \tag{1.5.35}$$

In symmetric cases of equal priors we may omit the constant factor of the priors and consider the gram matrix to be just the matrix of overlaps $G_{ij} = \langle\psi_i|\psi_j\rangle$. (note that G has dimensions $N \times N$ (not the dimensions d of the Hilbert space of the states).

We first note that G can be written as

$$G = X^\dagger X, \tag{1.5.36}$$

where X is the matrix with the components states $|\psi_i\rangle$ as columns, i.e. $X_{ki} = \langle w_k|\psi_i\rangle$ where $\{|k\rangle\}_{k=1}N$ is a set of states such that $\sum_k |w_k\rangle\!\langle w_k| = \mathbb{1}$ and the matrix $X^\dagger$ simply has $\langle\psi_i|$ as rows. More explicitly

$$X^\dagger = \begin{pmatrix} \left\langle\tilde{\psi}_1\right| & \cdot & \cdots & \cdot \\ \left\langle\tilde{\psi}_2\right| & \cdot & \cdots & \cdot \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \quad X = \begin{pmatrix} \left|\tilde{\psi}_1\right\rangle & \left|\tilde{\psi}_2\right\rangle & \cdots & \left|\tilde{\psi}_N\right\rangle \\ \cdot & \cdot & \cdot & \cdot \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}, \tag{1.5.37}$$

Exercise 17 *Prove that* G *is a non-negative matrix. Prove that any non-negative matrix is actually a Gram matrix for some set of unnormalized states*

We note that by definition $E_k = |w_k\rangle\langle w_k|$ defines a POVM. It easy to convince oneself that the optimal POVM for the discrimination of a set of pure states has rank-one elements, hence the choice of the vectors $|w_k\rangle$ defines a POVM that gives a success probability

$$P_s = \sum_k |X_{kk}|^2 \tag{1.5.38}$$

Thus, the optimal success probability is equivalent to finding the optimal Gram matrix decomposition $G = X^\dagger X$ in the sense that $\sum_k |X_{kk}|^2$ is maximal. Each choice of vectors $|w_k\rangle$ defines a measurement which seen to be equivalent to a particular choice X of the matrix decomposition of G. When the states are linearly independent the POVM becomes a von Neumann measurement in the space spanned by the set of vectors, because the only way to satisfy the completeness relation with d vectors is when they are orthogonal (see next section).

Off all the choices of X , there is a special one: when $X = X^\dagger$, that we denote it with S, i.e, $G = S^2$. Note that hermiticity condition implies that

$$\langle\tilde{\psi}_i|w_j\rangle = \langle w_i|\tilde{\psi}_j\rangle \tag{1.5.39}$$

(Remark: for linearly independent states it translates into $|\psi_k\rangle\langle w_k| = |w_k\rangle\langle\psi_k|$)

This choice is the square root of the non-negative matrix G and naturally coincides with the square root measurement of symmetric sources discussed above. For symmetric sources one just computes the square root of the Gram matrix and add the squares of its diagonal terms. Actually all the terms must be equal by symmetry. The success probability i this case simply read

$$P_s = |X_{kk}|^2 = |\langle\psi_1|w_k\rangle|^2 \quad \text{for any } k \tag{1.5.40}$$

Exercise 18 *Compute the success probability of exercise 16 for $N = 3$ using the Gram matrix formulation*

Exercise 19 *Consider the following set of states*

$$\{|\Psi_k\rangle = |0\rangle_1 |0\rangle_2 \cdots |0\rangle_{k-1} |\phi\rangle_k |0\rangle_{k+1} \cdots |0\rangle_N\}_{k=1}^N$$

with equal probabilities $\eta_i = 1/N$ and $\langle 0|\phi\rangle = c$.

(a) Compute the Gram matrix G

(b) Find the square-root S

(c) Find the success probability of correctly identifying one of the states

(d) Find a local measurement that at leading order gives the same success probability as the optimal measurement

1.6 Unambiguous discrimination

In this task one is asked to identify a state without error, naturally at the expense of having the possibility of a non-conclusive error: a "I don't know" outcome. The optimisation problem amounts to minimise the rate of inconclusive outcomes or maximise the rate success probability of a zero-error guess. Let us analyse first the binary case with equal priors, and then consider general priors and multi-hypothesis cases.

In the binary case the POVM has three elements E_0, E_1, E_Q, where the last corresponds to an inconclusive answer. Note that here we are in front of a genuine POVM that is not a von Neumann measurement. The no-error conditions read

$$p(E_0|\rho_1) = p(E_1|\rho_0) = 0 \tag{1.6.41}$$

Note that if ρ_i is a density matrix of full rank these conditions cannot be satisfied. For qubits it amounts that the states must be pure, which is the case we discuss next (remark: two pure states can always be considered to belong to a Hilbert space of dimension 2, $\mathbb{C}^2$, with a conveniently chosen basis).

Any two states can be written as

$$|\psi_{0,1}\rangle = \cos\theta/2\,|0\rangle \pm \sin\theta/2\,|1\rangle\,. \tag{1.6.42}$$

Conditions (1.8.70) are satisfied iff $E_0 \propto \left|\psi_1^\perp\right\rangle\!\left\langle\psi_1^\perp\right|$ and $E_1 \propto \left|\psi_0^\perp\right\rangle\!\left\langle\psi_0^\perp\right|$ For equal priors the proportional parameter must be equal

$$E_0 = \mu\left|\psi_1^\perp\right\rangle\!\left\langle\psi_1^\perp\right|, E_1 = \mu\left|\psi_0^\perp\right\rangle\!\left\langle\psi_0^\perp\right| \tag{1.6.43}$$

and actually the optimisation simply amount to find the maximum value of $\mu \geq 0$ such that

$$E_Q = \mathbb{1} - \mu(\left|\psi_1^\perp\right\rangle\!\left\langle\psi_1^\perp\right| + \left|\psi_0^\perp\right\rangle\!\left\langle\psi_0^\perp\right|) \geq 0 \tag{1.6.44}$$

Using

$$\left|\psi_{0,1}^\perp\right\rangle = \sin\theta/2\,|0\rangle \mp \cos\theta/2\,|1\rangle \tag{1.6.45}$$

is straight forward to see that the maximum value is

$$\mu^* = \min\{\frac{1}{2\sin^2\theta/2}, \frac{1}{2\cos^2\theta/2}\} = \frac{1}{1+|\cos\theta|} = \frac{1}{1+|\langle\psi_0|\psi_1\rangle|} := \frac{1}{1+c} \tag{1.6.46}$$

The success probability is then

$$P_s = \frac{1}{2}\left(\langle\psi_0|\,E_0\,|\psi_0\rangle + \langle\psi_1|\,E_1\,|\psi_1\rangle\right) = \frac{1}{2}\frac{2\sin^2\theta}{1+|\cos\theta|} = 1 - c \tag{1.6.47}$$

and the inconclusive, or failure probability,

$$Q = 1 - P_s = c \tag{1.6.48}$$

In general, unambiguous discrimination is only possible for pure states, (actually. as seen below, only for linearly independent states). For mixed states unambiguous discrimination is only possible for very specific sets of states-Tthey cannot have identical support, e.g. not all of the them must be full rank. For one state to be able to be identified without error one needs that it has a non-vanishing overlap with the intersection of the kernels of the rest of the states.

The linear independence condition for pure states can be easily derived by noticing that measurement the vectors $|w_k\rangle$ must satisfy $\langle\psi_i|w_l\rangle = 0$ for all $k \neq l$. This set of states exists if the inverse of the matrix X (or $X^\dagger$) [see Eq. (1.5.37)] that has the components of the states as its columns exist. This occurs only for linearly independent sets.

We postpone the analysis of arbitrary priors and multi-hypothesis after presenting the basic notions of semidefinite programming , a technique that proves to be extremely useful in this context.

Exercise 20 *Game formulation of discrimination problems*

(a) Justify that the minimum error task can be formulated as a game where a player gets a reward of 1€ if the guess is correct and 0€ if the guess is wrong. The optimal strategy aims at maximising the average earnings. Justify that this average coincide with the maximum success probability,

(b) Now one includes the possibility of abstaining. In this new game a player gets 1€ for correct guessing, 0€ for wrong guessing and -1€ if he/she abstains. Do the maximum earnings differ from previous item?

(c) In a game with payments 1€ for correct guessing, -1€ wrong guessing and 0 € abstention, to which probability do the average earnings correspond to?

1.7 Semidefinite program (SDP)

Is an optimisation technique that can be viewed as an extension of the linear programming which includes non-linear constraints. It is an extremely useful technique used in many branches of science and technology. A standard reference of the theory and applications can be found in [3], however here we follow the formulation of [4], that is better suited for quantum information problems. Naturally, both formulations are completely equivalent. The virtue fo this technique is twofold. From the numerical side, SDP can be implemented very efficiently in numerical routines, up to the point that when a problem is cast as a SDP many authors consider this a solution of the problem. From the theoretical side, the primal and dual version of the problem often provide complementary points of view of a problem and connect seemingly unrelated magnitudes. Furthermore it can even provide the necessary insight to obtain analytical results.

An SDP is a triple (Φ, A, B), where $\Phi : \mathcal{X} \longrightarrow \mathcal{Y}$ is a hermiticity preserving linear map and A and B are hermitian matrices, i.e., $A \in \mathrm{Herm}(\mathcal{X})$ and $B \in \mathrm{Herm}(\mathcal{Y})$ for some Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$. Note that the dimension $\mathcal{X}$ can be very different to that of $\mathcal{Y}$. Hermiticity preserving linear map means that $\Phi[A] \in \mathrm{Herm}(\mathcal{Y})$, and $\Phi[A_1 + A_2] = \Phi[A_1] + \Phi[A_2]$.

Example: $\Phi(A) = A \oplus (-A)$, i,e. $\Phi(A) = \begin{pmatrix} A & 0 \\ 0 & -A \end{pmatrix}$

Exercise 21 *If A has elements a_{ij} $i, j = 1, \cdots d$, which of the following maps are valid and which are not:*

(a) $\Phi(A) = \mathrm{tr}\, A$ *(b)* $\Phi(A) = \mathrm{tr}\, A\, \mathbb{1}_{d-2}$ *(c)* $\Phi(A) = \det(A)\mathbb{1}_d$

(d) $\Phi(A) = a_{11}$ *(e)* $\Phi(A) = (a_{11} + a_{12})\mathbb{1}$ *(f)* $\Phi(A) = \left(\sum_{ij} a_{ij}\right) A$

(g) $\Phi(A) = \Re(a_{12})$ *(h)* $\Phi(A) = A \otimes A$ *(i)* $\Phi(A) = A \oplus \mathbb{1}$

The triple has associated two optimisation problems called primal and dual. The primal problem is written as

$$\boxed{\begin{array}{l} \max \mathrm{tr}\, AX \\ \text{s.t.} \quad \Phi[X] = B \\ \qquad X \geq 0 \end{array}} \tag{1.7.49}$$

The function $\mathrm{tr}\, AX$ is called the primal objective funcion, $\mathcal{A} = \{X | \Phi[X] = B, X \geq 0\}$ is called the feasibility set and

$$\alpha = \sup_{X \in \mathcal{A}} \mathrm{tr}\, AX \tag{1.7.50}$$

is the primal optimal value. If the feasible set is empty, i.e., there is no instances satisfying the constraints of the problem, we take $\alpha = -\infty$. Notice that any element of the feasibility set gives a lower bound of the optimisation problem.

To obtain the dual form, we write the Lagrange function

$$\mathcal{L} = \mathrm{tr}\, AX + \mathrm{tr}[Y(B - \Phi[X])] + \mathrm{tr}[ZX], \tag{1.7.51}$$

where $Y = Y^\dagger$ and $Z = Z^\dagger$ are the Lagrange multipliers. Note that if $Z \geq 0$, $\mathcal{L}$ upperbounds α for any X in the feasible set

$$\mathcal{L} = \mathrm{tr}\, AX + \mathrm{tr}\, ZX \geq \mathrm{tr}\, AX, \text{ if } X \in \mathcal{A} \text{ and } Z \geq 0 \tag{1.7.52}$$

Define the dual map $\Phi^\dagger[\bullet]$ from the condition $\mathrm{tr}\, Y\Phi[X] = \mathrm{tr}\, \Phi^\dagger[Y]X$ to rewrite Eq. (1.7.51) as

$$\mathcal{L} = \mathrm{tr}[BY] + \mathrm{tr}[(A - \Phi^\dagger[Y] + Z)X] \tag{1.7.53}$$

From the extremal condition $\partial d\mathcal{L}/\partial X = 0$ one obtains

$$A - \Phi^\dagger[Y] + Z = 0 \tag{1.7.54}$$

and we can write the optimization of the Lagrange function as

$$\begin{aligned} &\min \operatorname{tr} BY \\ &Z = \Phi^\dagger[Y] - A \\ &Z \geq 0 \end{aligned}$$

Note that Z is a slack variable, that can be eliminated to finally yield

$$\boxed{\begin{aligned} &\min \operatorname{tr} BY \\ &\text{s.t. } \Phi^\dagger[Y] \geq A \\ &\quad Y = Y^\dagger \end{aligned}} \tag{1.7.55}$$

As in the primal case one defines the dual objective function as $\operatorname{tr} BY$, the dual feasibility set as $\mathcal{B} = \{Y | \Phi[^\dagger[Y] \geq A, Y = Y^\dagger\}$, and the dual optimum as

$$\beta = \inf_{Y \in \mathcal{B}} \operatorname{tr} BY \tag{1.7.56}$$

If the dual feasibility set $\mathcal{B}$ is empty, one takes the convention that $\beta = +\infty$. Except for some atypical or specifically tailored cases, $\alpha = \beta$. This equality is called strong duality (more on this below)

Exercise 22 *Find the dual maps of* $\Phi[X] = \operatorname{tr} X$, $\Phi[X] = X \oplus \mathbb{1}$, $\Phi[X] = X \oplus (-X)$, *and* $\Phi[X] = x_{11}$

Example. Write the primal and dual SDP's that give the maximum eigenvalue of a Hermitian matrix M and identify all the elements of the programs. Solution:

Primal	**Dual**		
$\max \operatorname{tr} MX$	$\min t$	$M = A$, $\Phi[X] = \operatorname{tr} X$, $B = 1$	Note, that X
$\operatorname{tr}[X] = 1$	$t\mathbb{1} \geq M$	$Y = y$, $\Phi^\dagger[Y] = t\mathbb{1}$	
$X \geq 0$			

corresponds to a quantum state, for $X \geq 0$ and $\operatorname{tr}[X] = 1$. Thus, the primal problem can be seen as the variational formulation of the maximum eigenvalue. The dimensions of the unknown X are the same as the dimensions of M. The dimensions of dual space is 1, that's why the dual variable y is simply a real number and it's written in lower case to remind this property. The dual program is quite enlightening. The parameter y is directly the largest eigenvalue, as any lower value will not give $t\mathbb{1} - M \geq 0$. From the primal program any X in the feasible set provides a lowerbound. As a particular example consider $M = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$, with $a, b, c \in \mathbb{R}$. Using the ansätze $X_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $X_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, one obtains the obvious bound $\lambda_{\max} \geq \min\{a, b\}$. From the dual program one can go beyond the strict variational method and obtain upperbounds to the maximum eigenvalue.

Exercise 23 *For the* 2×2 *matrix* M *above prove the bound* $\lambda_{\max} \geq (a + b + 2|c|)/2$ *. Check that for* $a = b$ *gives the exact maximum eigenvalue. From the positivity of* $t\,\mathbb{1} - M$ *obtain the exact value of* $\lambda_{\max}$.

1.7.1 Variations

If instead of an equality constraint in the primal program one has an inequality, $\Phi[X] \leq B$,

$$\boxed{\begin{aligned} &\max \operatorname{tr} AX \\ &\quad \Phi[X] \leq B \\ &\quad X \geq 0\,, \end{aligned}} \tag{1.7.57}$$

one simply introduces a slack variable $Z \geq 0$, writes $\tilde{X} = X \bigoplus Z$, defines the map $\tilde{\Phi}[\tilde{X}] = \Phi[X] + Z$ and makes the extension $\tilde{A} = A \bigoplus \mathbf{0}$, where $\mathbf{0}$ is the null matrix. These definitions transform the problem

into the standard form for $\tilde{A}, \tilde{\Phi}[\bullet]$ and B [cf. Eq. (1.7.49)]:

$$\begin{aligned} &\max \operatorname{tr} AX = \operatorname{tr} \tilde{A}\tilde{X} \\ &\Phi[X] \leq B \Leftrightarrow \Phi[X] + Z = B \Leftrightarrow \tilde{\Phi}[\tilde{X}] = B \\ &\tilde{X} \geq 0 \Leftrightarrow X, Z \geq 0\,. \end{aligned} \tag{1.7.58}$$

Taking into account that $\tilde{\Phi}^\dagger[Y] = \Phi[Y] \oplus Y$, dual version is directly

$$\boxed{\begin{aligned} &\min \operatorname{tr} BY \\ &\quad \Phi^\dagger[Y] \geq A \\ &\quad Y \geq 0\,. \end{aligned}} \tag{1.7.59}$$

Observe the beautiful duality between Eqs. (1.7.57) and (1.7.59). We could have started with the inequality version in Eq. (1.7.49) but the equality version is better suited for some of the proofs. Note that any inequality $\Phi[X] \geq B$ can be reversed as $-\Phi[X] \leq -B$, and any equality $\Phi[X] = B$ can be substituted by conditions $\Phi[X] \geq B$ and $\Phi[X] \leq B$.

If one has several conditions, equalities or inequalities, $\Phi_1[X] = B_1, \Phi_2[X] \leq B_2, \cdots$, simply consider $\tilde{X} = X \oplus X \oplus \cdots$, $\tilde{B} = B_1 \oplus B_2 \oplus \cdots$, $\tilde{Y} = Y_1 \oplus Y_2 \oplus \cdots$, and $\Phi[\tilde{X}] = \Phi_1[X] \oplus \Phi_2[X] \oplus \cdots$. That is, constraints are incorporated as direct sums and for each extra conditions one has as many dual variables as number of constraints.

Exercise 24 *Write a primal and dual SDP that give the sum of the r largest eigenvalues of a hermitian matrix A. Give two SDP's able to compute a specific eigenvalue λ_k.*

Exercise 25 *Trace norm of an operator A.*

(a) Write a SDP that gives the trace norm From the dual derive that trace norm can be written as $\min \operatorname{tr} Y$ s.t. $Y \geq A$ and $Y \geq -A$. Give an interpretation of this dual feasibility condition.

(b) Justify that the program $\max_U \operatorname{tr} UA$ s.t. U is a unitary matrix also gives the trace norm. Why is it not an SDP?

1.8 Quantum applications of SDP

1.8.1 Holevo conditions as the dual problem

The primal problem of minimum error discrimination is

$$\begin{aligned} &\max \operatorname{tr} \tilde{\rho}\tilde{E} \\ &E_1 + E_2 + \cdots E_N = \mathbb{1} \\ &\tilde{E} \geq 0\,, \end{aligned} \tag{1.8.60}$$

where $\tilde{\rho} = \oplus_{k=1}^N \eta_k \rho_k$ and $\tilde{E} = \oplus_{k=1}^N E_k$. Here $A = \tilde{\rho}$, the unknown X is $\tilde{E}$ and the map $\Phi[X]$ sums the diagonal blocks and $B = \mathbb{1}$. We only need to identify the dual map :

$$\Phi^\dagger[Y] = \bigoplus_{i=1}^N Y = \tilde{Y} \tag{1.8.61}$$

to get the dual program

$$\begin{array}{lcll} \max \operatorname{tr} Y & & \min \operatorname{tr} Y & \\ \tilde{Y} \geq \tilde{\rho} & \Longrightarrow & Y - \eta_j \rho_j \geq 0, \quad j = 1, 2, \ldots, N & \text{Observe that the dual} \\ \tilde{E} \geq 0\,, & & Y = Y^\dagger & (1.8.62) \end{array}$$

constraints are exactly the Holevo conditions for $Y = \sum_i \eta_i E_i$. Optimality is proved as follows. Take any

POVM, from the primal program, then one is assured that the objective function gives a lower bound of the success probability. If $Y = \sum_i \eta_i E_i$ then $\mathrm{tr}\, Y$ give the same value as the primal. If Y satisfies the dual conditions. it gives an upper bound of the success probability,. If the same value is a lower and upper bound it is the exact result.

1.8.2 SDP for unambiguous discrimination.

Unambiguous discrimination is a much simpler problem than minimum error as the structure of the POVM is essentially fixed by the zero-error conditions. As argued in Section 1.6, it is essentially determined by the inverse of the matrix of states X in Eq. 1.5.37. The rows of X^{-1} define a set of (unnormalized) states $\langle w_k|$ with the desired zero-error property $\langle w_k|\psi_l\rangle = 0$ for all $k \neq l$. Hence the only parameters that need to be optimised are the proportionality constants γ_k of the POVM elements

$$E_k = \gamma_k \, |w_k\rangle\langle w_k| \tag{1.8.63}$$

Note that $\langle w_k|\psi_k\rangle = 1$, hence the success probability is simply $P_s = \sum_{k=1}^{N} \gamma_k$ and the constraints are

$$E_Q = \mathbb{1} - \sum_k E_k \geq 0 \quad \text{and} \quad \gamma_k \geq 0, \quad k = 1, 2, \ldots N \tag{1.8.64}$$

We observe that

$$\sum_k E_k = \sum_k |w_k\rangle \, \gamma_k \, \langle w_k| = {X^{-1}}^{\dagger} \Gamma_D X^{-1}, \quad \text{with} \quad \Gamma_D = \mathrm{diag}\{\gamma_1, \gamma_2, \ldots, \gamma_N\} \tag{1.8.65}$$

and multiplying left and right with $X^\dagger$ and X we obtain.

$$X^\dagger E_Q X = G - \Gamma_D, \quad G_{ij} = \langle \psi_i|\psi_j\rangle \tag{1.8.66}$$

Exercise 26 *Demonstrate that if $A \geq 0$ $X^\dagger A X \geq 0$. Demostrate that if X and A are full rank, $X^\dagger A X$ is also full rank.*

Finally, the SDP for the success probability of unambiguous discrimination reads

$$\begin{aligned} &\max \mathrm{tr}\, \hat{\eta}\Gamma \\ &\Gamma_D \leq G \\ &\Gamma \geq 0 \end{aligned} \tag{1.8.67}$$

where $\hat{\eta} = \mathrm{diag}\{\eta_1, \eta_2, \ldots, \eta_N\}$

For two states, it is convenient to define the failure probabilities $\bar{\gamma}_i = 1 - \gamma_i$ and the usual definition of the overlap $|\langle \psi_1|\psi_2\rangle| = c$. The optimisation transforms to $\min \eta_1\bar{\gamma}_1 + \eta_2\bar{\gamma}_2$ and the condition $G - \Gamma_D \geq 0$ gives the suggestive uncertainty relation

$$\bar{\gamma}_1\bar{\gamma}_2 \geq c^2 \tag{1.8.68}$$

We note that this condition does not depend on the priors, only on G. The solution of the SD follows directly:

$$\bar{\gamma}_1 = c\sqrt{\frac{\eta_2}{\eta_1}}, \qquad \bar{\gamma}_2 = c\sqrt{\frac{\eta_1}{\eta_2}} \tag{1.8.69}$$

if

$$c^2 \leq \frac{\eta_1}{\eta_2} \leq \frac{1}{c^2}, \tag{1.8.70}$$

Outside this range

$$\bar{\gamma}_1 = 1, \ \bar{\gamma}_2 = c^2 \ \text{ if } \ \frac{\eta_1}{\eta_2} \leq c^2, \tag{1.8.71}$$

or

$$\bar{\gamma}_2 = 1, \ \bar{\gamma}_1 = c^2 \ \text{ if } \ \frac{\eta_1}{\eta_2} \geq 1/c^2 \tag{1.8.72}$$

In these extremal cases the priors are so biased that the optimal measurement discards detecting the state with the lowest prior and the POVM changes from having three to two elements. For instance, in the case $\bar{\gamma}_1 = 1$ we only have elements E_1 and E_Q with $E_1 + E_Q = \mathbb{1}$. The symmetric case $\eta_1 = \eta_2 = 1/2$ falls inside the range (1.8.70) for any value of the overlap and yields the well-known minimum inconclusive probability $Q = c$ computed in previous sections.

We finally tackle some multi-hypothesis cases. If we have a symmetric source with $\eta_i = 1/N$ and constant state overap $\langle\psi_i|\psi_j\rangle = c$ for al i, j, then it is clear that all conditional success probabilities are equal and $\Gamma_D = \gamma\mathbb{1}$. In this case the SDP is

$$\begin{aligned} &\max \gamma && (1.8.73)\\ &\gamma\mathbb{1} \leq G \\ &\gamma \geq 0 && (1.8.74)\end{aligned}$$

Exercise 27 *Explain why the SDP of (1.8.73) computes the minimum eigenvalue of G. Find the success probability of uanambiguous discrimination of the states of the set defined in exercise 19.*

1.8.3 The dual SDP of unambiguous discrimination

The primal problem with the priors included in the Gram matrix, i.e.,

$$G_{ij} = \sqrt{\eta_i}\sqrt{\eta_j}\langle\psi_i|\psi_j\rangle$$

and

$$\Gamma_D = \mathrm{diag}\{\eta_1\gamma_1, \eta_2\gamma_2, \ldots, \eta_N\gamma_N\}$$

reads

$$\begin{aligned} &\max \mathrm{tr}\,\Gamma \\ &\Gamma_D \leq G \\ &\Gamma \geq 0, && (1.8.75)\end{aligned}$$

We can easily identify $A = \mathbb{1}, B = G, X = \Gamma$, and $\Phi[\Gamma] = \Gamma_D \leq G$, and $\Gamma_D = \mathrm{diag}\{\Gamma_{11}, \Gamma_{22}, \ldots, \Gamma_{nn}\}$. The dual map $\Phi^\dagger[Y]$ is the same as the direct one $\Phi^\dagger[Y] = Y_D$.

Hence the dual version reads

$$\begin{aligned} P_s = &\min\ \mathrm{tr}\,GZ \\ &Z_{kk} \geq 1\ , \quad k \geq 1, \ldots, n \\ &Z \geq 0\ . && (1.8.76)\end{aligned}$$

(recall that we construct G with the priors included).

Exercise 28 *Prove that Eq. (1.8.76) is the dual version of Eq. (1.8.75) .*

The nice thing about the dual version is that it transforms the optimisation into a variational problem. The idea is to take a convenient matrix Z satisfying the conditions $Z_{kk} \geq 1$ and $Z \geq 0$. For instance take any $Z = |u\rangle\langle u|$, with $|u\rangle = (u_1, u_2, ...)$ and and $|u_i| \geq 1$. Then compute the "induced" conditional probabilities $\gamma_k = u_k(G\,|u\rangle)_k$, these are chosen in such a way that $\mathrm{tr}\,GZ = \sum\gamma_k$. If this γ_k satisfy the primal condition $G - \Gamma \geq 0$ and $\gamma_k \geq 0$, then they are the optinal solution. yields a valid solution, i.e. $G - \gamma \geq 0$.

We illustrate this procedure for the two state $\{|\psi_1\rangle, |\psi_2\rangle\}$ with arbitrary priors $\eta_1, \eta_2\}$. As usual we take $c = \langle\psi_1|\psi_2\rangle$, which can be taken w.l.o.g. as positive and $0 \leq c \leq 1$. The Gram matrix is then

$$G = \begin{pmatrix} \eta_1 & c\sqrt{\eta_1\eta_2} \\ c\sqrt{\eta_1\eta_2} & \eta_2 \end{pmatrix} \qquad (1.8.77)$$

The ansatz is here $|u\rangle = (1, -1)$ and the efficiency matrix Γ_D reads

$$\Gamma_D = \begin{pmatrix} \gamma_1 = \eta_1 - c\sqrt{\eta_1\eta_2} & 0 \\ 0 & \gamma_2 = \eta_2 - c\sqrt{\eta_1\eta_2} \end{pmatrix} \tag{1.8.78}$$

Notice that if $\eta_1/\eta_2 \geq 1/c^2$ then γ_2 becomes negative (we can assume w.l.o.g. that $\eta_1 \geq \eta_2$). Beyond this point the ansatz becomes $|u'\rangle = (1, -c\sqrt{\eta_1/\eta_2})$ so that $\gamma_2' = 0$ and the new efficiencies are

$$\Gamma_D' = \begin{pmatrix} \gamma_1' = \eta_1(1 - c^2) & 0 \\ 0 & \gamma_2' = 0 \end{pmatrix} \tag{1.8.79}$$

Let's check that in both cases $G - \Gamma_D$ is a semidefinite positive matrix. In the first case when $\eta_1/\eta_2 \leq 1/c^2$ indeed

$$G - \Gamma_D = c\sqrt{\eta_1\eta_2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \geq 0, \tag{1.8.80}$$

while for $\eta_1/\eta_2 > 1/c^2$

$$G - \Gamma_D' = \begin{pmatrix} c^2\eta_1 & c\sqrt{\eta_1\eta_2} \\ c\sqrt{\eta_1\eta_2} & \eta_2 \end{pmatrix}, \tag{1.8.81}$$

Which has eigenvalues 0 and $\eta_1 c^2 + \eta_2 \geq 0$ (notice that the easiest way to find this eigenvalue is computing the trace).

It is interesting to note that the positivity condition does not yield constraints on the form of the ansatz. The relevant constraint is $\Gamma_D \geq 0$.

Exercise 29 *Prove that in the general case of arbitrary priors and states, the minimum eigenvalue of G is a lowerbound of the success probability.*

Appendix

1.A Holevo's derivation of Holevo conditions

The positivity condition is explicitly included by writing $E_i = A_i^\dagger A_i$ and the completeness condition is taken care of by a Lagrange multiplier Y. So, one has to maximize the function

$$L = \sum_i \mathrm{tr}[\tilde{\rho}_i A_i^\dagger A_i] - Y\left(\sum_i A_i^\dagger A_i - \mathbb{1}\right) \tag{1.A.82}$$

The extremality conditions over A_k and $A_k^\dagger$ yield

$$(\tilde{\rho}_k - Y)A_k^\dagger = 0, \quad A_k(\tilde{\rho}_k - Y) = 0 \tag{1.A.83}$$

Multiplying by A_k on the right of first equation and by $A_k^\dagger$ on the left of the second and summing over k one obtains

$$Y = \sum_k \tilde{\rho}_k E_k = \sum_k E_k \tilde{\rho}_k, \tag{1.A.84}$$

which implies the hermiticity of Y. Taking the second derivative in (1.A.82) and imposing the condition of maximum one gets

$$\tilde{\rho}_i - Y \leq 0 \to Y - \tilde{\rho}_i \geq 0, \quad i = 1, 2, \ldots, n, \tag{1.A.85}$$

which is Eq. (1.5.24)

Proof of the second property (1.5.26). Consider w.l.o.g. $i = 1, j = 2$. Because of positivity one can write $E_i = T_i^2$, $i = 1, 2$. Now consider the modified operators $E_i' = T_i'^\dagger T_i'$

$$T_1' = T_1 + \epsilon A^\dagger T_2; \quad T_2' = T_2 - \epsilon A T_1, \tag{1.A.86}$$

where A is an arbitrary operator, not needed to be hermitian. At first order in ϵ we have

$$E_1' = E_1 + \epsilon(T_2 A T_1 + T_1 A^\dagger T_2)\ , \tag{1.A.87}$$

$$E_2' = E_2 - \epsilon(T_2 A T_1 + T_1 A^\dagger T_2) \tag{1.A.88}$$

and $E_1' + E_2' = E_1 + E_2$. The difference of the success probabilities is

$$P_s' - P_s = \epsilon\,\mathrm{tr}\left[(\tilde{\rho}_1 - \tilde{\rho}_2)(T_2 A T_1 + T_1 A^\dagger T_2)\right] \tag{1.A.89}$$

$$2\epsilon \Re\,\mathrm{tr}\left[T_1(\tilde{\rho}_1 - \tilde{\rho}_2)T_2 A\right] \tag{1.A.90}$$

This difference must vanish for any arbitrary choice of A, otherwise one could construct a modified POVM with larger success probability (if the value in the r.h.s was negative, change the sign of A). Thus

$$T_1[\tilde{\rho}_1 - \tilde{\rho}_2]T_2 = 0 \tag{1.A.91}$$

$$\Rightarrow T_1^2[\tilde{\rho}_1 - \tilde{\rho}_2]T_2^2 = E_1(\tilde{\rho}_1 - \tilde{\rho}_2)E_2 = 0 \tag{1.A.92}$$

1.B Optimal POVM discriminating sets of linearly independent pure states is von Neumann

The derivation follows [6]. Let's denote the set of n pure states by $\{|\alpha_i\rangle\}_{i=1}^n$ and $\{\eta_i\}_{i=1}^n$ the corresponding priors. If they are linearly independent, there exists a set $\{|\phi_i\rangle\}$ such that $\langle\phi_i|\alpha_j\rangle = C_j\delta_{ij}$. An explicit construction from the Gram matrix formalism is presented in Section **??**, but one can trivially argue that the set of equations $\langle\phi_i|\alpha_j\rangle = 0 \;\; \forall i \neq j$ and $\langle\phi_i|\alpha_i\rangle = C_j$ has a non-trivial solution for $i = 1, \cdots, n$ because the states $\{|\alpha_i\rangle\}$ are linearly independent.

Now we use the optimality Holevo conditions to prove (i) $E_jE_i\,|\alpha_i\rangle = \delta_{ij}E_i\,|\alpha_i\rangle$ and (ii) the states $|\beta_k\rangle = E_k\,|\alpha_k\rangle$ are linearly independent. Then it will be trivial to prove that (iii) $E_iE_j = \delta_{ij}E_i$. This last condition implies that $\{E_j\}$ is a von Neumann measurement (notice that the orthogonality of operators requires that E_j are rank one).

Proof of (i):

From Holevo's condition (1.A.83):

$$\begin{aligned}\left(E_j\sum_i \eta_iE_i\,|\alpha_i\rangle\langle\alpha_i| - \eta_jE_j\,|\alpha_j\rangle\langle\alpha_j|\right)|\phi_k\rangle = 0\\ E_j\sum_i \eta_i\,|\beta_i\rangle\,C_i\delta_{ik} = \eta_kC_k\delta_{jk}\,|\beta_k\rangle\\ E_j\,|\beta_k\rangle = \delta_{jk}\,|\beta_k\rangle\end{aligned} \tag{1.B.93}$$

where in the last equation we have used that $C_k \neq 0$.

Proof of (ii):

If $\{|\beta_k\rangle\}$ were not independent there would exist a non-trivial state $|\gamma\rangle$ orthogonal to all states of the set, i.e, $\langle\gamma|\beta_k\rangle = 0 \;\; \forall k$. But this is impossible because using the Holevo condition (1.5.24) we get a contradiction

$$\begin{aligned}0 &\geq \langle\gamma|\left(\sum_i \eta_iE_i\,|\alpha_i\rangle\langle\alpha_i| - \eta_k\,|\alpha_k\rangle\langle\alpha_k|\right)|\gamma\rangle\\ &= \sum_i \eta_i\langle\gamma|\beta_i\rangle\langle\alpha_i|\gamma\rangle - \eta_k|\langle\alpha_k|\gamma\rangle|^2\\ &= -\eta_k|\langle\alpha_k|\gamma\rangle|^2 \leq 0\end{aligned} \tag{1.B.94}$$

Proof of (iii):

Because linear independence of the states $\{|\beta_k\rangle\}$, any state $|\omega\rangle$ can be written as $|\omega\rangle = \sum_k b_k\,|\beta_k\rangle$, hence

$$E_j\,|\omega\rangle = \sum_k b_kE_j\,|\beta_k\rangle = \sum_k b_k\delta_{jk}\,|\beta_k\rangle = b_j\,|\beta_j\rangle\,, \tag{1.B.95}$$

where we have used property (1.B.93). We also have

$$E_iE_j\,|\omega\rangle = b_jE_i\,|\beta_j\rangle = \delta_{ij}b_i\,|\beta_i\rangle = \delta_{ij}E_i\,|\omega\rangle \tag{1.B.96}$$

Since Eq. (1.B.96) is valid for any state $|\omega\rangle$ we finally have

$$E_iE_j = \delta_{ij}E_i \tag{1.B.97}$$

which is the von Neumann condition when the number of operators is equal to de dimension of the space.

1.C Holevo conditions for linearly independent set of pure states

Let's now specialise in a source of linearly independent pure states $\tilde{\rho}_k = \left|\tilde{\psi}_k\right\rangle\!\left\langle\tilde{\psi}_k\right|$, where $\left|\tilde{\psi}_k\right\rangle \equiv \sqrt{\eta_k}\,|\psi_k\rangle$ (we essentially follow [7, 8]).

In this case there exists a transformation X s.t.

$$\left|\tilde{\psi}_k\right\rangle = X\,|w_k\rangle \rightarrow \quad X_{jk} = \langle\tilde{\rho}_j|\tilde{\psi}_k\rangle \tag{1.C.98}$$

where $\{|w_k\rangle\}_{k=1}^n$ is an orthonormal basis of the space spanned by $\{|\psi_k\rangle\}$, which defines a von Neumann measurement with

$$P_s = \sum_{k=1}^{n} |\langle w_k|\tilde{\psi}_k\rangle|^2 = \sum_k |X_{kk}|^2 \tag{1.C.99}$$

Notice that the gram matrix G with elements $G_{ij} = \langle\tilde{\psi}_i|\tilde{\psi}_j\rangle$ is written as

$$G_{ij} = \langle\tilde{\psi}_i|\tilde{\psi}_j\rangle = \sum_k \langle\tilde{\psi}_i|w_k\rangle\langle w_k|\tilde{\psi}_j\rangle = \sum_k X^\dagger_{ik}X_{kj} \quad , \tag{1.C.100}$$

$$G = X^\dagger X \tag{1.C.101}$$

For a linearly independent set $G > 0$ and X is invertible. The measurement basis is $|w_k\rangle = X^{-1}\left|\tilde{\psi}_k\right\rangle$.

Theorem 1. $G = X^\dagger X$ gives the optimal success probability iff the matrix X satisfies:

$$X_{ii}X^*_{ji} = X_{ij}X^*_{jj} \tag{1.C.102}$$

and

$$Y = XX^\dagger_d > 0; \quad X_d = \mathrm{diag}\{X_{11}, X_{22}, \ldots, X_{nn}\} \tag{1.C.103}$$

Notice that Y has to be strictly positive.

Proof. *Necessity.* Eq. (1.C.102) follows from Eq. (1.5.26). To prove the necessity of (1.C.103) we just prove that none of the X_{ii} terms vanish. The matrix X is non-singular ($\det X \neq 0$) because $G > 0$, hence there is no column or row with all its terms vanishing. Assume $X_{11} = 0$, then there must exist another element of the same row that is non-vanishing, say $X_{12} = \alpha \neq 0$. From Eq. (1.C.102) it follows that $0 = X_{11}X^*_{21} = X_{12}X^*_{22} \rightarrow X_{22} = 0$. Then we could change the roles of the measurement basis elements $|w_1\rangle \leftrightarrow |w_2\rangle$ and obtain $X'_{11} = \alpha \neq 0$. From $|X_{11}|^2 + |X_{22}|^2 = 0 \leq |X'_{11}|^2 + |X'_{22}|^2 \leq |X'_{11}|^2 = |\alpha|^2$ it follows that a measurement with a vanishing probability $p(k|k)$ cannot be optimal. From the Holevo conditions $Y \geq 0$ and from $\det X.X^\dagger_d = \det X \det X_d \neq 0$ it follows $Y > 0$.

Sufficiency $Y = (Y - \left|\tilde{\psi}_k\right\rangle\!\left\langle\tilde{\psi}_k\right|) + \left|\tilde{\psi}_k\right\rangle\!\left\langle\tilde{\psi}_k\right| = Y^k + \left|\tilde{\psi}_k\right\rangle\!\left\langle\tilde{\psi}_k\right|$. The eigenvalues $\{\lambda_a\}$ of Y and $\{\mu_b\}$ of Y^k in ascending order satisfy the interlacing relation

$$\mu_1 \leq \lambda_1 \leq \mu_2 \leq \lambda_2 \leq \cdots \leq \mu_n \leq \lambda_n \tag{1.C.104}$$

Now, $Y > 0 \rightarrow \lambda_1 > 0 \rightarrow \mu_b > 0$ for $b \geq 2$ and since $\langle w_k|\,Y^k\,|w_k\rangle = 0 \rightarrow \mu_1 = 0$ and $Y - \left|\tilde{\psi}_k\right\rangle\!\left\langle\tilde{\psi}_k\right| \geq 0 \;\; \forall k$.

Theorem 2. The square root measurement, i.e. when $X = X^\dagger = S \rightarrow G = S^2$, is optimal iff $S_{ii} = S_{jj}$ for any pair i, j. If G can be written in blocks $G = \bigoplus_t G_t$, S^t_{kk} must constant within each block.

$\leftarrow$ It's direct. If $S = S^\dagger$ and S_{kk} is constant, the conditions of Theorem 1 (1.C.102) and (1.C.103) are trivially satisfied (notice that by construction $S > 0$ and $S_{kk} > 0$ as since they are all equal by hypothesis they cannot be all zero). $\rightarrow$ If S provides the optimal measurement S_{kk} must be equal for all k. We just have to use the first condition (1.C.102) and find a convenient sequence of non-vanishing elements. It is known that for non-singular matrices $\det S \neq 0$ one can always find a sequence of indexes $i_1, i_2, \cdots, i_n$ such that $S_{i_k i_{k+1}} \neq 0$ ($i_{n+1} \equiv i_1$). From Eq. (1.C.102) and hermiticity $S_{ii} = S_{jj} = r$.

Furthermore

$$r = \frac{1}{n}\sum_k S_{kk} = \frac{1}{n}\,\mathrm{tr}\,S = \frac{1}{n}\sum_l \sqrt{\lambda_l} \tag{1.C.105}$$

and

$$P_s = \sum_k |S_{kk}|^2 = nr^2 = \frac{1}{n}\left(\sum_l \sqrt{\lambda_l}\right)^2 \tag{1.C.106}$$

If the a priori probabilities are equal, $\eta_i = 1/n$, it is more convenient to define the Gram matrices without this factor, $G'_{ij} = \langle\psi_i|\psi_j\rangle$. Then $G' = nG$, $\lambda'_l = n\lambda_l$, $X' = \sqrt{n}X$ and $S' = \sqrt{n}S$. Thus

$$P_s = \frac{1}{n}\sum_k |S'_{kk}|^2 = \left(\frac{\sum_l \sqrt{\lambda'_l}}{n}\right)^2 = \langle\sqrt{\lambda'}\rangle^2 \tag{1.C.107}$$

Example 1: Two states equal priors,i.e, where $\eta_i = 1/2$ and $G'_{ij} = \delta_{ij} + c(1-\delta_{ij})$ with $c = |\langle\psi_1|\psi_2\rangle|$. By symmetry the success probabilitiy for each state is the same, whence the square root measurement must be optimal. The eigenvalues are $\lambda_\pm = 1 \pm c$, hence

$$P_s = \left(\frac{\sqrt{1+c}+\sqrt{1-c}}{2}\right)^2 = \frac{1}{2}\left(1+\sqrt{1-c^2}\right) \tag{1.C.108}$$

Example 2: One flip error. A source of n states $\{|\Psi_k\rangle\}$, $k = 1, 2, \cdots, n$, with equal priors $\eta_k = 1/n$. Each state is a tensor product of $n-1$ identical local states $|0\rangle$ and a different local state $|\phi\rangle$ at position k. The Gran matrix has elements $G'_{ij} = \delta_{ij} + c^2(1-\delta_{ij})$ with $c^2 = |\langle 0|\phi\rangle|^2$. The Gram matrix can be written as $G' = (1-c^2)\mathbb{1} + nc^2\,|u\rangle\langle u|$, where $|u\rangle = (1/n)(1, 1, \cdots, 1)$. The eigenvalues are trivially $\lambda_1 = 1 - c^2 + nc^2$, and $\lambda_k = 1 - c^2$ for $k > 1$. Thus

$$P_s = \left[\frac{\sqrt{1+(n-1)c^2}+(n-1)\sqrt{1-c^2}}{n}\right]^2 \tag{1.C.109}$$

Notice the trivial limiting cases: $c = 0 \to P_s = 1$ and $c = 1 \to P_s = 1/n$. For large n one laso has $P_s \simeq 1 - c^2$.

Example 3: Corollary. For general priors Eq. (1.C.106) is a lowerbound, i.e.,

$$P_s^{\mathrm{opt}} \geq \frac{1}{n}\left(\sum_k \sqrt{\lambda}\right)^2 = \frac{\left(\mathrm{tr}\,\sqrt{G}\right)^2}{n} \tag{1.C.110}$$

Check for the case of two states with with unequal priors, η_1 and η_2. The Gram matrix reads:

$$G = \begin{pmatrix} \eta_1 & \sqrt{\eta_1\eta_2}c \\ \sqrt{\eta_1\eta_2}c & \eta_2 \end{pmatrix}$$

$$\langle\sqrt{\lambda}\rangle^2 = \frac{1}{2}\left(\sqrt{\lambda_+}+\sqrt{\lambda_-}\right)^2 = \frac{\lambda_+ + \lambda_- + 2\sqrt{\lambda_+\lambda_-}}{2}$$
$$= \frac{\mathrm{tr}\,G + 2\sqrt{\det G}}{2} = \frac{1+2\sqrt{\eta_1\eta_2(1-c^2)}}{2}$$

Recall

$$P_s^{\mathrm{opt}} = \frac{1+\sqrt{1-4\eta_1\eta_2c^2}}{2}$$

and inequality (1.C.110) holds since $4\eta_1\eta_2 \leq 1$. Well, in fact in this case it is a shitty bound as when $c = 0$ it does not recover $P_s \to 1$.

Proof: The square root measurement, being a particular protocol provides a lower bound. We have

$$P_s \geq \sum_k |S_{kk}|^2 \geq \frac{(\sum_k S_{kk})^2}{n} = \frac{\left(\mathrm{tr}\,\sqrt{G}\right)^2}{n},$$

where the second inequality follows from the Schwarz inequality $|\mathbf{s}|^2|\mathbf{u}|^2 \geq (\mathbf{s}.\mathbf{u})^2$, with $s_k = S_{kk}$ and $u_k = 1/\sqrt{n}$. The bound is worse than the square root measurement and it is saturated iff $S_{ii} = S_{jj}\ \forall i, j$. It has the property though that it is not necessary to compute the square root, just the eigenvalues of G.

1.D Technicalities and duality theorems for SDP

Chapter 2

Quantum Hypothesis Testing Part II: Asymptotic error rates and channel discrimination

2.1 Asymptotics

So far we have studied scenarios where one has limited resources, say one or few copies n of a quantum system, in order to infer some property of the system. The figure of merit, e.g. probability of error, has a non trivial dependence on the number of copies. However one does expect a more and more regular behaviour as the number of copies or resources increases. In this section we will be concerned about the characterization of this large n behaviour.

Knowing the necessary number of copies required to attain a certain high level of confidence has of course practical implications. Also from fundamental point of view, studying the large n scaling allows us to define a consistent n-independent notion of closeness between states. This has a clear advantage over finite n definitions, since as we have already seen it is easy to find instances where

$$P_{\text{err}}(\rho,\sigma) > P_{\text{err}}(\rho',\sigma')$$

but

$$P_{\text{err}}(\rho\otimes\rho,\sigma\otimes\sigma) < P_{\text{err}}(\rho'\otimes\rho',\sigma'\otimes\sigma').$$

As we will shortly see the probability of error in quantum (and classical) hypothesis testing decreases exponentially with the number of copies $P_{\text{err}} \sim \mathrm{e}^{-rn}$ (see Figure 2.1.1). Our aim here will be to find closed-form expressions for the error rates in various quantum settings.

Recap: Before proceeding let us make a short recap on binary (classical) Hypothesis testing: Given a sample x taken from a distribution $P(x)$ provide the optimal decision rule or guess function $g(x) \in \{0,1\}$ according to which the null, $H_0 : Q = P_0$ or the alternative, $H_1 : Q = P_1$, hypothesis are accepted. Equivalently, one can define decision rule as decision regions in sample space: the set of outcomes $A = \{x : g(x) = 0\}$ and its complement A^c are the decision regions for hypothesis H_0 and H_1 respectively.

For every decision rule we can define two types of errors, termed type I (false positive or false alarm) and type II (false negative) or error of first and second kind respectively:

$$\begin{aligned}\alpha &= P(g(x)=1|H_0) = P_0(g(x)=1) = P_0(A^c) \quad \text{Type I error}\\ \beta &= P(g(x)=0|H_1) = P_1(g(x)=0) = P_1(A) \quad \text{Type II error}\end{aligned} \tag{2.1.1}$$

Symmetric Hypothesis testing: follows a bayesian approach, where some prior probabilities $\{\eta_0, \eta_1 = 1-\eta_0\}$ are assigned to each hypothesis so that the total probability of error can be computed

$$P_{\text{err}} = P(A^c|H_0)\eta_0 + P(A|H_1)\eta_1 = \eta_0\alpha + \eta_1\beta \tag{2.1.2}$$

this quantity has to be minimized by choosing the appropriate decision rule. Equivalently we can

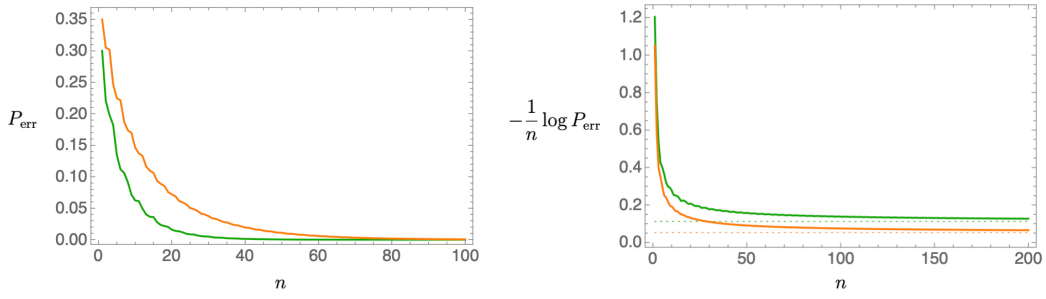


Figure 2.1.1: Error probability for the discrimination between two possible coins with bias $p = 0.9$ versus $q = 0.5$ (green curve) and $p = 0.8$ versus $q = 0.5$ after n tosses. Plot on the right shows that the large-n scaling is exponential, $P_{\text{err}} \sim \mathrm{e}^{-nr}$. The corresponding asymptotic rate r is shown with a dotted line.

maximize the success probability

$$\begin{aligned} P_s &= P(A|H_0)\eta_0 + P(A^c|H_1)\eta_1 \\ &= \sum_x \left(P(g(x)=0,x|H_0)\eta_0 + P(g(x)=1,x|H_1)\eta_1\right) \\ &\leq \sum_x \max\{P(g(x)=0,x|H_0)\eta_0, P(g(x)=1,x|H_1)\eta_1\} \end{aligned} \tag{2.1.3}$$

Clearly the upper bound can be attained by always guessing for the most likely hypothesis: i.e.

$$g(x) = \begin{cases} 0 & \text{if } P(x|H_0)\eta_0 \geq P(x|H_1)\eta_1 \\ 1 & \text{if } P(x|H_0)\eta_0 \leq P(x|H_1)\eta_1 \end{cases} \tag{2.1.4}$$

or more succinctly: $A = \{x : P_0(x)\eta_0 \geq P_1(x)\eta_1\}$ or in terms of the joint probabilities $A = \{x : P(x,H_0) \geq P(x,H_1)\}$.

Therefore the minimum error probability is given by:

$$P_{\text{err}} = \sum_x \min\{P(x|H_0)\eta_0, P(x|H_1)\eta_1\} \tag{2.1.5}$$

Asymmetric Hypothesis testing: In cases where the priors are not provided it is customary to keep track of the probabilities of both types of error (2.1.1). Since there's a trade-off between the two one typically wishes to minimize β the number of false negatives (fatal consequences) leaving some room for positive detections, i.e. $P(A|H_0) = 1-\alpha \geq 1-\epsilon > 0$, i.e. $\alpha < \epsilon$ for some $0 < \epsilon < 1$.

The following theorem singles out the likelihood ratio test as an optimal one in the following sense:

Theorem 2.1.1 (Neyman-Pearson Lemma) *Given a binary hypothesis testing problem with the two hypotheses $Q = P_0$ vs. $Q = P_1$. For $T > 0$ define the acceptance region:*

$$A(T) = \{x : \frac{P_0(x)}{P_1(x)} \geq T\} \tag{2.1.6}$$

and the corresponding error probabilities $\alpha^* = P(A(T)^c|H_0)$ and $\beta^* = P(A(T)|H_1)$. Given any other decision region B and its associated error probabilities α_B and β_B. If $\alpha_B \leq \alpha^*$ then $\beta_B \geq \beta^*$ (and vice versa).

Proof.
Let $\phi_A(x)$ and $\phi_B(x)$ indicator functions, i.e. $\phi_A(x) = (1 \text{ if } x \in A \;\&\; 0 \text{ if } x \notin A)$, then for all x it holds

$$(\phi_A(x) - \phi_B(x))\,(P_0(x) - TP_1(x)) \geq 0 \tag{2.1.7}$$

This can be seen by considering separately the cases $x \in A$ and $x \notin A$. Multiplying out and summing this over the entire space, we obtain

$$0 \leq \sum_x (\phi_A P_0 - T\phi_A P_1 - P_0\phi_B + TP_1\phi_B) = \tag{2.1.8}$$

$$= \sum_{x\in A}(P_0 - TP_1) - \sum_{x\in B}(P_0 + TP_1) \tag{2.1.9}$$

$$= (1-\alpha^*) - T\beta^* - (1-\alpha_B) + T\beta = T(\beta_B - \beta^*) - (\alpha^* - \alpha) \tag{2.1.10}$$

Since $T > 0$, $\beta_B - \beta^* \geq (\alpha^* - \alpha_B)/T \geq 0$ which means that $\alpha_B \leq \alpha^*$ implies $\beta_B \geq \beta^*$. ■

2.1.1 Classical error exponents

(excerpt from Cover and Thomas Chap 11 [9])

Let $X^n = \{X_1, \ldots, X_n\}$ be a sequence of (random) symbols taken from an alphabet $\mathcal{X} = \{a_1, \ldots, a_{|\mathcal{X}|}\}$. A particular instance or realization of this sequence will be denoted by $x^n = \mathbf{x} = \{x_1, \ldots, x_n\}$.

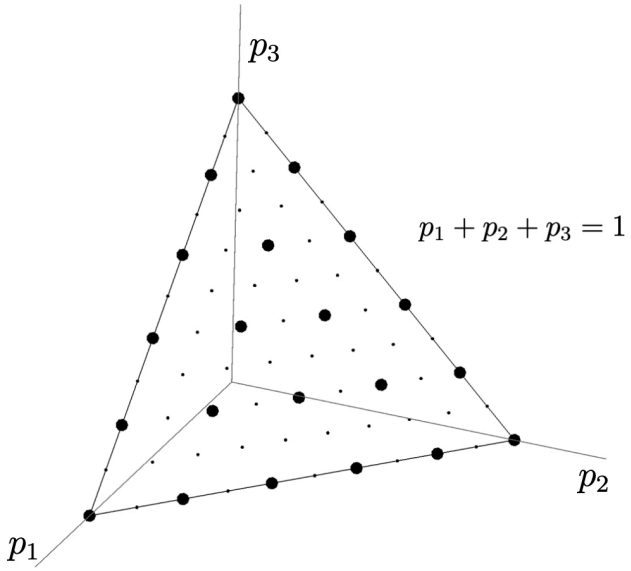


Figure 2.1.2: Probability simplex for alphabet size $m = 3$. Empirical distributions or types $P_{\mathbf{x}} \in \mathcal{P}_n$ for $n = 5$ (large dots) and $n = 10$ (small dots).

Definition 2.1.2 *The* TYPE $P_{\mathbf{x}}$ *(or empirical probability distribution) of a sequence* $x_1, x_2, \ldots, x_n$ *is the relative proportion of occurrences of each symbol of* $\mathcal{X}$*, i.e.,* $P_{\mathbf{x}}(a) = N(a|\mathbf{x})/n$ *for all* $a \in \mathcal{X}$*, where* $N(a|\mathbf{x})$ *is the number of times the symbol* a *occurs in the sequence* $\mathbf{x} \in \mathcal{X}^n$*.*

The PROBABILITY SIMPLEX in $\mathcal{R}^m$ is the set of points $\mathbf{p} = (p_1, \ldots, p_m) \in \mathcal{R}^m$ such that $p_i \geq 0$ and $\sum_{i=1}^m p_i = 1$. The probability simplex is an $(m-1)$-dimensional manifold in m-dimensional space. Note that the empirical distribution $P_{\mathbf{x}}$ is a probability distribution, and therefore is a point on the simplex. Figure 2.1.2 shows the probability simplex for $m = 3$.

Definition 2.1.3 *Let* $\mathcal{P}_n$ *denote the set of types with denominator* n*.*

For example, if $\mathcal{X} = \{0, 1\}$*, the set of possible types with denominator* n *is*

$$\mathcal{P}_n = \left\{ (P(0), P(1)) : \left(\frac{0}{n}, \frac{n}{n}\right), \left(\frac{1}{n}, \frac{n-1}{n}\right), \ldots, \left(\frac{n}{n}, \frac{0}{n}\right) \right\}$$

Figure 2.1.2 shows the types of $n = 5$ and $n = 10$ for an alphabet of size $m = 3$ on the simplex.

Definition 2.1.4 *If* $P \in \mathcal{P}_n$*, the set of sequences of length* n *and type* P *is called the* TYPE CLASS *of* P*, denoted* $T(P)$ *:*

$$T(P) = \{\mathbf{x} \in \mathcal{X}^n : P_{\mathbf{x}} = P\}$$

The number of types is at most **polynomial** in n:

$$|\mathcal{P}_n| \leq (n+1)^{|\mathcal{X}|} \tag{2.1.11}$$

Theorem 2.1.5 *If* $X_1, X_2, \ldots, X_n$ *are drawn i.i.d. according to* $Q(x)$*, the probability of* $\mathbf{x}$ *depends only on its type and is given by*

$$Q^n(\mathbf{x}) = \mathrm{e}^{-n(H(P_{\mathbf{x}}) + D(P_{\mathbf{x}} \| Q))} \tag{2.1.12}$$

Here we have introduced the relative entropy (or Kullback-Leibler divergence) from distribution Q to distribution P as:

$$D\left(P\|Q\right) = \sum_{a\in\mathcal{X}} P(a)\log\frac{P(a)}{Q(a)} \tag{2.1.13}$$

and the Shannon entropy of distribution P as:

$$H(P) = -\sum_{a\in\mathcal{X}} P(a)\log P(a) \tag{2.1.14}$$

Proof.

$$\begin{aligned} Q^n(\mathbf{x}) &= \prod_{i=1}^{n} Q\left(x_i\right) = \prod_{a\in\mathcal{X}} Q(a)^{N(a|\mathbf{x})} = \prod_{a\in\mathcal{X}} Q(a)^{n\cdot P_{\mathbf{x}}(a)} = \prod_{a\in\mathcal{X}} \mathrm{e}^{nP_{\mathbf{x}}(a)\log Q(a)} \\ &= \prod_{a\in\mathcal{X}} \mathrm{e}^{n(P_{\mathbf{x}}(a)\log Q(a) - P_{\mathbf{x}}(a)\log P_{\mathbf{x}}(a) + P_{\mathbf{x}}(a)\log P_{\mathbf{x}}(a))} \\ &= \mathrm{e}^{n\sum_{a\in\mathcal{X}}\left(-P_{\mathbf{x}}(a)\log\frac{P_{\mathbf{x}}(a)}{Q(a)} + P_{\mathbf{x}}(a)\log P_{\mathbf{x}}(a)\right)} = \mathrm{e}^{-n(D(P_{\mathbf{x}}\|Q)+H(P_{\mathbf{x}}))} \end{aligned}$$

■

Notice that this is an **exact** result. As a corollary we get that when $\mathbf{x}$ is in the type class of Q (i.e. typical sequence), then

$$Q^n(\mathbf{x}) = \mathrm{e}^{-nH(Q)} \tag{2.1.15}$$

Theorem 2.1.6 *Size of a type class $T(P)$. For any type $P \in \mathcal{P}_n$,*

$$\frac{1}{(n+1)^{|\mathcal{X}|}}\mathrm{e}^{nH(P)} \leq |T(P)| \leq \mathrm{e}^{nH(P)} \tag{2.1.16}$$

This allows us to write

$$|T(P)| \doteq \mathrm{e}^{nH(P)} \tag{2.1.17}$$

where the symbol $f_n \doteq g_n$ denotes equality to the first order in the exponent, i.e. $\lim_{n\to\infty}\log\frac{f_n}{g_n} = 0$.

Theorem 2.1.7 *Probability of type class. For any $P \in \mathcal{P}_n$ and any distribution Q, the probability of the type class $T(P)$ under Q^n is $\mathrm{e}^{-nD(P\|Q)}$ to first order in the exponent. More precisely,*

$$\frac{1}{(n+1)^{|\mathcal{X}|}}\mathrm{e}^{-nD(P\|Q)} \leq Q^n(T(P)) \leq \mathrm{e}^{-nD(P\|Q)} \tag{2.1.18}$$

Proof.

$$\begin{aligned} Q^n(T(P)) &= \sum_{\mathbf{x}\in T(P)} Q^n(\mathbf{x}) = \sum_{\mathbf{x}\in T(P)} \mathrm{e}^{-n(D(P\|Q)+H(P))} \\ &= |T(P)|\mathrm{e}^{-n(D(P\|Q)+H(P))} \end{aligned}$$

Using this expression together with (2.1.16) we arrive to the desired result. ■

The probability of a type class is asymptotically given by the relative entropy

$$Q^n(T(P)) \doteq \mathrm{e}^{-nD(P\|Q)} \tag{2.1.19}$$

Given an $\epsilon > 0$, we can define a typical set T_Q^ϵ of sequences for the distribution Q^n as

$$T_Q^\epsilon = \{x^n : D\left(P_{x^n}\|Q\right) \leq \epsilon\}$$

Then the probability that x^n is not typical is

$$\begin{aligned} 1 - Q^n\left(T_Q^\epsilon\right) &= \sum_{P:D(P\|Q)>\epsilon} Q^n(T(P)) \leq \sum_{P:D(P\|Q)>\epsilon} \mathrm{e}^{-nD(P\|Q)} \\ &\leq \sum_{P:D(P\|Q)>\epsilon} \mathrm{e}^{-n\epsilon} \leq (n+1)^{|\mathcal{X}|}\mathrm{e}^{-n\epsilon} \\ &= \mathrm{e}^{-n\left(\epsilon - |\mathcal{X}|\frac{\log(n+1)}{n}\right)} \end{aligned}$$

which goes to 0 as $n \to \infty$. Hence, the probability of the typical set T_Q^ϵ goes to 1 as $n \to \infty$. One can also show that empirical distribution P_{x^n} converges to Q. This result has a flavor of the law of large numbers (see exercise below).

Theorem 2.1.8 *Let $X_1, X_2, \ldots, X_n$ be i.i.d. $\sim Q(x)$. Then*

$$\Pr\left\{D\left(P_{x^n}\|Q\right) > \epsilon\right\} \leq \mathrm{e}^{-n\left(\epsilon - |\mathcal{X}|\frac{\log(n+1)}{n}\right)}$$

and consequently, $D\left(P_{x^n}\|Q\right) \to 0$ with probability 1[1].

Exercise 2.1.1. *Prove the following* **Concentration inequalities**

1. *Markov's inequality. For any non-negative random variable X and any $t > 0$, show that*

$$\Pr\{X \geq t\} \leq \frac{\mathbb{E}(X)}{t}$$

Give a random variable that achieves this inequality with equality. It is often a very useful proof method to use a trivial extension of this inequality:
If $f(x)$ is a strictly increasing non-negative function $\Pr\{X \geq t\} = \Pr\{f(X) \geq f(t)\} \leq \frac{\mathbb{E}(f(X))}{f(t)}$.

2. *Chebyshev's inequality. Let Y be a random variable with mean $\mathbb{E}(Y) = \mu$ and variance σ^2. By letting $X = (Y-\mu)^2$, show that for any $\epsilon > 0$,*

$$\Pr\{|Y - \mu| > \epsilon\} \leq \frac{\sigma^2}{\epsilon^2}$$

3. **The weak law of large numbers**. *Let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. random variables with mean μ and variance σ^2. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ be the sample mean. Show that*

$$\Pr\left\{\left|\bar{X}_n - \mu\right| > \epsilon\right\} \leq \frac{\sigma^2}{n\epsilon^2} \tag{2.1.20}$$

Thus $\Pr\left\{\left|\bar{X}_n - \mu\right| > \epsilon\right\} \to 0$ *as $n \to \infty$. This is known as the weak law of large numbers.*

There is also a **strong law of large numbers**, which is harder to proof and gives a stronger sense of convergence: Given i.i.d. sequence with $\mu = \mathbb{E}(X) < \infty$, $\bar{X}_n$ converges almost surely to μ, i.e. $\Pr\{\lim_{n\to\infty} \bar{X}_n = \mu\} = 1$. Furthermore the **central limit theorem** states that the distribution $Z = \sqrt{n}(\bar{X}_n - \mathbb{E}(X))$ goes to a gaussian distribution with zero mean and variance var$X$.

We now give some results that quantify the probability of rare events (**large deviation theory)**. In particular we wish to quantify the probability of a set of sequences (or outcomes) that are fully specified by a region in probability space. In other words, the set necessarily includes all sequences of the same type.

[1] **Convergence of random variables**. We say that a sequence of random variables, $X_1, X_2, \ldots$, converges to a random variable X:
i) *in probability* if for every $\epsilon > 0$, $\Pr\{|X_n - X| > \epsilon\} \to 0$,
ii) *in mean square* if $E\left(X_n - X\right)^2 \to 0$,
iii) *with probability 1* (also called ALMOST SURELY) if $\Pr\{\lim_{n\to\infty} X_n = X\} = 1$.

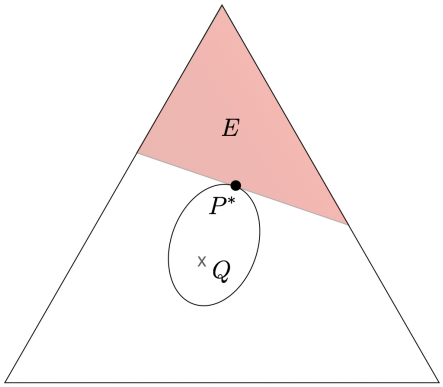


Figure 2.1.3: Probability simplex for alphabet size $m = 3$. Point x shows the true distribution $Q(x)$. E is a set of probability distributions and $P^*(x)$ is the probability distribution belonging to E that is closest to $Q(x)$ according to relative entropy, and which determines the rate $Q^n(E)$ for increasing n.

Theorem 2.1.9 (Sanov's theorem) *Let $X_1, X_2, \ldots, X_n$ be i.i.d. $\sim Q(x)$. Let $E \subseteq \mathcal{P}$ be a set of probability distributions. Then*

$$Q^n(E) := Q^n\left(E \cap \mathcal{P}_n\right) \leq (n+1)^{|\mathcal{X}|} \mathrm{e}^{-nD(P^* \| Q)}$$

where

$$P^* = \arg\min_{P \in E} D(P \| Q) \tag{2.1.21}$$

is the distribution in E that is closest to Q in relative entropy. If, in addition, the set E is the closure of its interior[2] *then*

$$-\frac{1}{n} \log Q^n(E) \to D\left(P^* \| Q\right)$$

Proof.
We only give a prove for an upper bound:

$$\begin{aligned}
Q^n(E) &= \sum_{P \in E \cap \mathcal{P}_n} Q^n(T(P)) \leq \sum_{P \in E \cap \mathcal{P}_n} \mathrm{e}^{-nD(P\|Q)} \\
&\leq \sum_{P \in E \cap \mathcal{P}_n} \max_{P \in E \cap \mathcal{P}_n} \mathrm{e}^{-nD(P\|Q)} = \sum_{P \in E \cap \mathcal{P}_n} \mathrm{e}^{-n \min_{P \in E \cap \mathcal{P}_n} D(P\|Q)} \\
&\leq \sum_{P \in E \cap \mathcal{P}_n} \mathrm{e}^{-n \min_{P \in E} D(P\|Q)} = \sum_{P \in E \cap \mathcal{P}_n} \mathrm{e}^{-nD(P^*\|Q)} \\
&\leq (n+1)^{|\mathcal{X}|} \mathrm{e}^{-nD(P^*\|Q)}
\end{aligned}$$

■

Note that the distribution P^* does not necessarily belong to $\mathcal{P}_n$, however, as we increase n, we can find a distribution in $\mathcal{P}_n$ that is closer a closer to P^*.

As an application of Sanov's theorem suppose that we wish to find the probability to get a sequence $x_1, \ldots, x_n$ that fulfills a set of k linear constrains of the form: $C = \{\frac{1}{n} \sum_{i=1}^n g_j(x_i) \geq b_j\}_{j=1}^k$, which corresponds to a region of the simplex defined by half-planes. Using $P_{\mathbf{x}}(a) = N(a \mid \mathbf{x})/n$ the corresponding set E is defined as

$$E = \left\{ P : \sum_a P(a) g_j(a) \geq b_j, j = 1, 2, \ldots, k \right\} \tag{2.1.22}$$

[2] This condition is necessary to exclude pathological examples like for instance the case were E is the set distributions such that all probabilities are irrational, then $Q^n\left(E \cap \mathcal{P}_n\right) = 0$.

So that $\{\sum_a P_{\mathbf{x}}(a)g_j(a) \geq b_j, j = 1, 2, \ldots, k\}_{\mathbf{x}} = \{P \in E \cap \mathcal{P}_n\}$. To find the closest distribution in E to Q, we minimize $D(P\|Q)$ subject to the constraints C. Using Lagrange multipliers, we construct the functional

$$L(P) = \sum_x P(x) \log \frac{P(x)}{Q(x)} + \sum_j \lambda_j (\sum_x P(x) g_j(x) - b_j) + v(\sum_x P(x) - 1)$$

We then differentiate and calculate the closest distribution to Q to be of the form

$$P^*(x) = \frac{Q(x) \mathrm{e}^{\sum_j \lambda_j g_j(x)}}{\sum_{a \in \mathcal{X}} Q(a) \mathrm{e}^{\sum_j \lambda_j g_j(a)}} \tag{2.1.23}$$

and the constants λ_j are chosen so as to satisfy the constrains.

Let us consider some specific examples:

Exercise 2.1.2. *(Dice) Suppose that we toss a fair die n times; what is the probability that the average of the throws is greater than or equal to 4 ? [you might need to solve numerically some part]*

Exercise 2.1.3. *(Coins) Suppose that we have a fair coin and want to estimate the probability of observing more than 700 heads in a series of 1000 tosses.*

Finally, before entering the proper hypothesis testing results, let us quote a beautiful theorem.

Theorem 2.1.10 (Conditional limit theorem) *Let E be a closed convex subset of $\mathcal{P}$ and let Q be a distribution not in E. Let $X_1, X_2, \ldots, X_n$ be discrete random variables drawn i.i.d. $\sim Q$. Let P^* achieve $\min_{P \in E} D(P\|Q)$. Then*

$$\Pr\left(X_1 = a_1, X_2 = a_2, \ldots, X_m = a_m \mid P_{X^n} \in E\right) \to \prod_{i=1}^{m} P^*\left(a_i\right) \quad \text{in probability} \tag{2.1.24}$$

This holds for fixed m as $n \to \infty$. The result is not true for $m \sim n$, since the conditioning induces correlations.

Now, let us turn back to the asymptotic rates for hypothesis testing. From Neyman-Pearson lemma (Theorem 2.1.1) the optimal guess or acceptance region is given by likelihood ratio test, which we can write as

$$\Lambda(\mathbf{x}) := \log \frac{P_0\left(x_1, x_2, \ldots, x_n\right)}{P_1\left(x_1, x_2, \ldots, x_n\right)} > \log T \tag{2.1.25}$$

For i.i.d.'s we can simplify further the lhs of this inequality

$$\Lambda\left(\mathbf{x}\right) = \log \frac{P_0\left(x_1, x_2, \ldots, x_n\right)}{P_1\left(x_1, x_2, \ldots, x_n\right)} = \sum_{i=1}^{n} \log \frac{P_0\left(x_i\right)}{P_1\left(x_i\right)} \tag{2.1.26}$$

$$= \sum_{a \in \mathcal{X}} n P_{\mathbf{x}}(a) \log \frac{P_0(a)}{P_1(a)} = \tag{2.1.27}$$

$$= \sum_{a \in \mathcal{X}} n P_{\mathbf{x}}(a) \log \frac{P_0(a)}{P_1(a)} \frac{P_{\mathbf{x}}(a)}{P_{\mathbf{x}}(a)} = \tag{2.1.28}$$

$$= \sum_{a \in \mathcal{X}} n P_{\mathbf{x}}(a) \log \frac{P_{\mathbf{x}}(a)}{P_1(a)} - \sum_{a \in \mathcal{X}} n P_{\mathbf{x}}(a) \log \frac{P_{\mathbf{x}}(a)}{P_0(a)} \tag{2.1.29}$$

$$= nD\left(P_{\mathbf{x}} \| P_1\right) - nD\left(P_{\mathbf{x}} \| P_0\right) \tag{2.1.30}$$

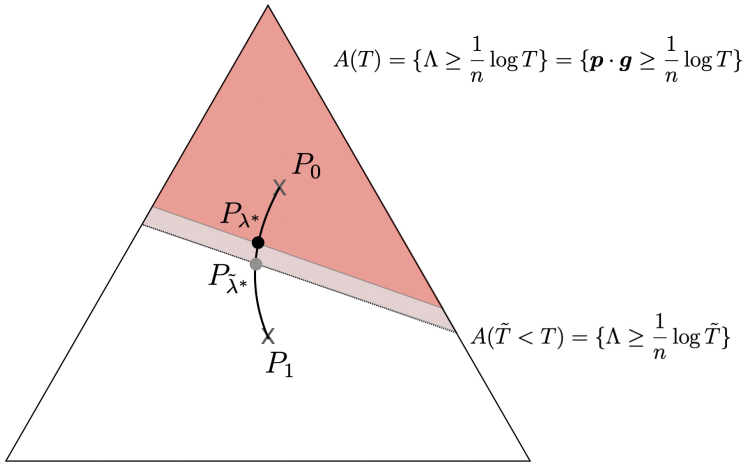


Figure 2.1.4: Illustration of the optimal Neyman-Pearson decision region in 2.1.31 for a T (red area) and $\tilde{T} < T$ (light shaded area) and corresponding closest point $P_{\lambda^*} \in A(T)$ (or $P_{\tilde{\lambda}^*}$) to P_1 associated to the alternative hypothesis. This very same point is also the one that attains the minimum distance from $A(T)^c$ to P_0. According to Sanov's theorem this two distances determine the asymptotic error rates.The Hellinger arc is the curve traced by all points P_λ.

Hence the likelihood test is equivalent to

$$D\left(P_{\mathbf{x}}\|P_1\right) - D\left(P_{\mathbf{x}}\|P_0\right) > \frac{1}{n}\log T \tag{2.1.31}$$

Let A denote the region on which hypothesis 0 is accepted. From (2.1.27) it is clear that the region is defined by a single linear constrain of the type studied in (2.1.22) with $g(a) = \log\frac{P_0(a)}{P_1(a)}$ and $b = \frac{1}{n}\log T$, i.e. using vector notation: $\boldsymbol{p}\cdot\boldsymbol{g} \geq \frac{1}{n}\log T$.

The boundary $\boldsymbol{p}\cdot\boldsymbol{g} = b$ of the region is the set of distributions for which the difference between the "distances" (as measured by the relative entropy[3]) is a constant. This boundary is the analog of the perpendicular bisector in Euclidean geometry, as illustrated in Figure 2.1.4. Since both regions A and A^c are convex we can make use of Sanov's theorem in order to compute the probability of error of first and second kind (2.1.1),

$$\begin{aligned}\alpha &= P_0(A^c) \doteq \mathrm{e}^{-nD(P_0^*\|P_0)} \quad \text{Type I error}\\ \beta &= P_1(A) \doteq \mathrm{e}^{-nD(P_1^*\|P_1)} \quad \text{Type II error}\end{aligned} \tag{2.1.32}$$

where P_0^* is the closest element of A^c to distribution P_0, and P_1^* is the closest element of A to distribution P_1. In order to compute the P_0^* we minimize $D(P\|P_0)$ subject to $\Lambda < \log T$ of the form (2.1.22) with $g(x) = \log\frac{P_0(x)}{P_1(x)}$ and $b = \frac{1}{n}\log T$. Therefore we can use the result of the optimization (2.1.23) and recalling $g(a) = \log\frac{P_0(a)}{P_1(a)}$ we find

$$P_0^*(x) = \frac{P_0(x)\mathrm{e}^{\lambda g(x)}}{\sum_{a\in\mathcal{X}} P_1(a)\mathrm{e}^{\lambda g(a)}} = \frac{P_0(x)^{1-\lambda}P_1(x)^{\lambda}}{\sum_{a\in\mathcal{X}} P_0(a)^{1-\lambda}P_1(a)^{\lambda}} =: P_\lambda \tag{2.1.33}$$

where $\lambda = \lambda^*$ is chosen so that $D(P_{\lambda^*}\|P_0) - D(P_{\lambda^*}\|P_1) = \frac{1}{n}\log T$. From the symmetry of expression (2.1.33), it is clear that $P_1^* = P_0^*$ (see also Figure 2.1.4) and that the probabilities of error behave exponentially with exponents given by the relative entropies $D(P^*\|P_0)$ and $D(P^*\|P_1)$, i.e.

$$\begin{aligned}\alpha &= P_0(A^c) \doteq \mathrm{e}^{-nD(P^*\|P_0)} \quad \text{Type I error}\\ \beta &= P_1(A) \doteq \mathrm{e}^{-nD(P^*\|P_1)} \quad \text{Type II error}\end{aligned} \tag{2.1.34}$$

[3]The relative entropy, a.k.a Kullback-Leibler distance, is not a proper (metric) distance as it is not symmetric and does not satisfy the triangle inequality. See Section **??**

Also note from equation (2.1.33) that P_λ with $\lambda \in [0,1]$ describes a geodesic curve in the simplex, often called HELLINGER ARC in the literature, which interpolates between $P_{\lambda\to 0} = P_0$ and $P_{\lambda\to 1} = P_1$.

To end this section we will calculate the best error exponent when one of the two types of error goes to zero arbitrarily slow, but before that we still need a couple of preliminary theorems and definitions.

Theorem 2.1.11 *Let $X_1, X_2, \dots X_n$ be a sequence of random variables drawn i.i.d. according to $P_0(x)$, and let $P_1(x)$ be any other distribution on the same alphabet $\mathcal{X}$. Then the log–likelihood per number of sample concentrates around the relative entropy:*

$$\frac{1}{n}\Lambda(\mathbf{x}) = \frac{1}{n}\log\frac{P_0(x_1, x_2, \dots, x_n)}{P_1(x_1, x_2, \dots, x_n)} \to D(P_{\mathbf{x}}\|P_0) \quad \textit{in probability} \tag{2.1.35}$$

PROOF.
It follows directly from the weak law of large numbers (2.1.20) starting from (2.1.27)

$$\frac{1}{n}\Lambda(\mathbf{x}) = \sum_{a\in\mathcal{X}} P_{\mathbf{x}}(a)\log\frac{P_0(a)}{P_1(a)} = \sum_{a\in\mathcal{X}} P_{\mathbf{x}}(a)g(a) \tag{2.1.36}$$

$$\to \mathbb{E}_{P_0}(g(a)) = D(P_0\|P_1) \tag{2.1.37}$$

■

Definition 2.1.12 *For a fixed n and $\epsilon > 0$, a sequence $(\mathbf{x}) \in \mathcal{X}^n$ is said to be relative entropy typical if and only if*

$$D(P_0\|P_1) - \epsilon \le \frac{1}{n}\log\frac{P_0(\mathbf{x})}{P_1(\mathbf{x})} \le D(P_0\|P_1) + \epsilon \tag{2.1.38}$$

The set of relative entropy typical sequences is called the **relative entropy typical set** $A_\epsilon^{(n)}(P_0\|P_1)$.

Theorem 2.1.13 *The relative entropy typical set satisfies the following properties:*

1. *For $(x_1, x_2, \dots, x_n) \in A_\epsilon^{(n)}(P_0\|P_1)$,*

$$P_0(\mathbf{x})\mathrm{e}^{-n(D(P_0\|P_1)+\epsilon)} \le P_1(\mathbf{x}) \le P_0(\mathbf{x})\mathrm{e}^{-n(D(P_0\|P_1)-\epsilon)}$$

2. *$P_0\left(A_\epsilon^{(n)}(P_0\|P_1)\right) > 1-\epsilon$, for n sufficiently large.*

3. *$P_1\left(A_\epsilon^{(n)}(P_0\|P_1)\right) < \mathrm{e}^{-n(D(P_0\|P_1)-\epsilon)}$*

4. *$P_1\left(A_\epsilon^{(n)}(P_0\|P_1)\right) > (1-\epsilon)\mathrm{e}^{-n(D(P_0\|P_1)+\epsilon)}$, for n sufficiently large.*

PROOF.
The proof of property 1 follows directly from the definition of the relative entropy typical set. The second property follows from Theorem 2.1.11. To prove the third property, we write

$$\begin{aligned} P_1\left(A_\epsilon^{(n)}(P_0||P_1)\right) &= \sum_{\mathbf{x}\in A_\epsilon^{(n)}(P_0\|P_1)} P_1(\mathbf{x}) \le \sum_{\mathbf{x}\in A_\epsilon^{(n)}(P_0\|P_1)} P_0(\mathbf{x})\mathrm{e}^{-n(D(P_0\|P_1)-\epsilon)} \\ &= \mathrm{e}^{-n(D(P_0\|P_1)-\epsilon)} \sum_{\mathbf{x}\in A_\epsilon^{(n)}(P_0\|P_1)} P_0(\mathbf{x}) = \mathrm{e}^{-n(D(P_0\|P_1)-\epsilon)} P_0\left(A_\epsilon^{(n)}(P_0\|P_1)\right) \\ &\le \mathrm{e}^{-n(D(P_0\|P_1)-\epsilon)} \end{aligned}$$

where the first inequality follows from property 1, and the second inequality follows from the fact that the probability of any set under P_0 is less than 1 .

To prove the lower bound on the probability of the relative entropy typical set, we use a parallel argument with a lower bound on the probability:

$$\begin{aligned} P_1\left(A_\epsilon^{(n)}\left(P_0\|P_1\right)\right) &= \sum_{\mathbf{x}\in A_\epsilon^{(n)}(P_0\|P_1)} P_1(\mathbf{x}) \geq \sum_{\mathbf{x}\in A_\epsilon^{(n)}(P_0\|P_1)} P_0(\mathbf{x})\mathrm{e}^{-n(D(P_0\|P_1)+\epsilon)} \\ &= \mathrm{e}^{-n(D(P_0\|P_1)+\epsilon)} \sum_{\mathbf{x}\in A_\epsilon^{(n)}(P_0\|P_1)} P_0(\mathbf{x}) = \mathrm{e}^{-n(D(P_0\|P_1)+\epsilon)} P_0\left(A_\epsilon^{(n)}\left(P_0||P_1\right)\right) \\ &\geq (1-\epsilon)\mathrm{e}^{-n(D(P_0\|P_1)+\epsilon)} \end{aligned}$$

■

Lemma 2.1.14 *Let $B_n \subset \mathcal{X}^n$ be any set of sequences $x_1, x_2, \ldots, x_n$ such that $P_0(B_n) > 1-\epsilon$. Let P_1 be any other distribution such that $D(P_0\|P_1) < \infty$. Then $P_1(B_n) > (1-2\epsilon)\mathrm{e}^{-n(D(P_0\|P_1)+\epsilon)}$*

Proof.
For simplicity, we will denote $A_\epsilon^{(n)}(P_0||P_1)$ by A_n. Since $P_0(B_n) > 1-\epsilon$ (by hypothesis) and $P_0(A_n) > 1-\epsilon$ (by Theorem 2.1.13), we have, by the union bound, $P_0(A_n^c \cup B_n^c) < 2\epsilon$, or equivalently, $P_0(A_n \cap B_n) > 1-2\epsilon$. Thus,

$$\begin{aligned} P_1(B_n) \geq P_1(A_n \cap B_n) &= \sum_{x^n \in A_n \cap B_n} P_1(x^n) && (2.1.39) \\ &\geq \sum_{x^n \in A_n \cap B_n} P_0(x^n)\,\mathrm{e}^{-n(D(P_0\|P_1)+\epsilon)} = \mathrm{e}^{-n(D(P_0\|P_1)+\epsilon)} \sum_{x^n \in A_n \cap B_n} P_0(x^n) \\ &= \mathrm{e}^{-n(D(P_0\|P_1)+\epsilon)} P_0(A_n \cap B_n) && (2.1.40) \\ &\geq \mathrm{e}^{-n(D(P_0\|P_1)+\epsilon)}(1-2\epsilon) && (2.1.41) \end{aligned}$$

where the second inequality follows from the properties of the relative entropy typical sequences (Theorem 2.1.13) and the last inequality follows from the union bound above.

■

We now consider the problem of testing two hypotheses, P_0 vs. P_1. We hold one of the probabilities of error bounded by a fix ϵ and attempt to minimize the other probability of error. We show that the relative entropy is the best exponent in probability of error.

Theorem 2.1.15 (Stein Lemma) *Let $X_1, X_2, \ldots, X_n$ be i.i.d. $\sim Q$. Consider the hypothesis test between two alternatives, $Q = P_0$ and $Q = P_1$, where $D(P_0\|P_1) < \infty$. Let $A_n \subseteq \mathcal{X}^n$ be an acceptance region for hypothesis H_0. Let the probabilities of error be*

$$\alpha_n = P_0^n(A_n^c), \quad \beta_n = P_1^n(A_n)$$

and for $0 < \epsilon < \frac{1}{2}$, define

$$\beta_n^\epsilon = \min_{A_n \subseteq \mathcal{X}^n} \beta_n \text{ such that } \alpha_n < \epsilon$$

then the optimal error exponent is:

$$-\lim_{n\to\infty} \frac{1}{n} \log \beta_n^\epsilon = D(P_0\|P_1)$$

Proof.
We prove this theorem in two parts. In the first part (*direct part, or achievability*) we exhibit a sequence of sets A_n for which the probability of error β_n goes exponentially to zero as $D(P_0\|P_1)$. In the second part (*converse part, or optimality*) we show that no other sequence of sets can have a lower exponent in the probability of error. For the first part, we choose as the sets $A_n = A_\epsilon^{(n)}(P_0||P_1)$. As proved in

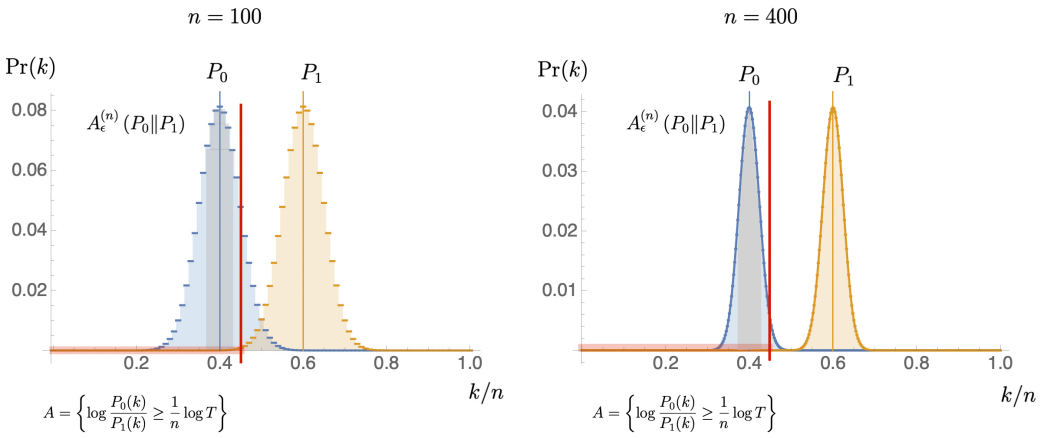


Figure 2.1.5: Illustration of likelihood test for two coins: P_0 corresponding to a probability of heads of $p_A = 0.4$, and P_1 with probability $p_B = 0.6$ for a sequence of $n = 100$ tosses (left) and $n = 400$ (right). They types, or empirical probability distributions, are fully characterized by the number of heads k: $\mathcal{P}_n = \{(\frac{k}{n}, \frac{n-k}{n})\}_{k=1}^{n}$. The probability concentrates around the type $p = k/n = p_A$ for the first coin and $p = k/n = p_B$ for the second. The optimal Neyman-Pearson decision is region marked in red, while the gray shaded region indicates the relative entropy typical set and its probability under H_0.

2.1.13, this sequence of sets has $P_0\left(A_n^c\right) < \epsilon$ for n large enough. Also,

$$-\lim_{n\to\infty} \frac{1}{n} \log P_1\left(A_n\right) \geq D\left(P_0 \| P_1\right) - \epsilon$$

from property 3 of Theorem 2.1.13. Thus, the relative entropy typical set satisfies the bounds of the lemma.

To show that no other sequence of sets can to better (optimality), consider any sequence of sets B_n with $P_0\left(B_n\right) > 1 - \epsilon$. By Lemma 2.1.14, we have $P_1\left(B_n\right) > (1 - 2\epsilon)\mathrm{e}^{-n(D(P_0\|P_1)+\epsilon)}$, and therefore

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n} \log P_1\left(B_n\right) &> -\left(D\left(P_0 \| P_1\right) + \epsilon\right) + \lim_{n\to\infty} \frac{1}{n} \log(1 - 2\epsilon) \\ &= -\left(D\left(P_0 \| P_1\right) + \epsilon\right)\end{aligned}$$

Thus, no other sequence of sets has a probability of error exponent better than $D\left(P_0 \| P_1\right)$. Thus, the set sequence $A_n = A_\epsilon^{(n)}\left(P_0 \| P_1\right)$ is asymptotically optimal in terms of the exponent in the probability. ■

Note that the relative entropy typical set, although asymptotically optimal (i.e., achieving the best asymptotic rate), is not the optimal set for any fixed hypothesis-testing problem. The optimal set that minimizes the probabilities of error is that given by the Neyman-Pearson lemma (theorem 2.1.1). This and other asymptotic features of the i.i.d. probability distributions are illustrated in figure 2.1.5.

We next move to the **symmetric hypothesis testing** scenario where we are given a sequence $X_1, X_2, \ldots, X_n$ i.i.d. drawn from $Q(x)$. We have two hypotheses: $Q = P_0$ with prior probability π_0 and $Q = P_1$ with prior probability π_1 and the total probability of error is

$$P_{\text{err}}^{(n)} = \pi_0 \alpha_n + \pi_1 \beta_n$$

and our purpose is to find the error exponent

$$C^* = \lim_{n\to\infty} -\frac{1}{n} \log \min_{A_n \subseteq \mathcal{X}^n} P_{\text{err}}^{(n)}. \tag{2.1.42}$$

Theorem 2.1.16 (Chernoff) *The best achievable exponent in the Bayesian probability of error is C^*, where*

$$C^* = D\left(P_{\lambda *} \| P_0\right) = D\left(P_{\lambda^*} \| P_1\right)$$

with

$$P_\lambda = \frac{P_0^\lambda(x) P_1^{1-\lambda}(x)}{\sum_{a \in \mathcal{X}} P_0^\lambda(a) P_1^{1-\lambda}(a)}$$

and λ^* the value of λ such that

$$D\left(P_{\lambda^*} \| P_0\right) = D\left(P_{\lambda^*} \| P_1\right)$$

Or equivalently,

$$C^* = C_{\mathrm{Ch}}\left(P_0, P_1\right) := -\min_{0 \le \lambda \le 1} \log \sum_{a \in \mathcal{X}} P_0^\lambda(a) P_1^{1-\lambda}(a)$$

Proof.
We have already presented the basic ingredients for the proof. We have shown that the optimum test is a likelihood ratio test (Theorem 2.1.1 and equation (2.1.31)). The test divides the probability simplex into regions corresponding to hypothesis H_0 (A) and hypothesis H_1 (A^c), respectively. P_λ^* that lies in the border minimizes the distances from A to P_1 and at the same time from A^c to P_0. Hence from Sanov's theorem we get (2.1.34):

$$\alpha_n = P_0^n\left(A^c\right) \doteq \mathrm{e}^{-nD(P_{\lambda *}||P_0)} \quad \beta_n = P_1^n(A) \doteq \mathrm{e}^{-nD(P_{\lambda^*}\|P_1)}$$

where λ^* actually depends on the decision region (value of T) that we need to optimize, so we drop the $*$ fro the moment and still refer to it as λ.

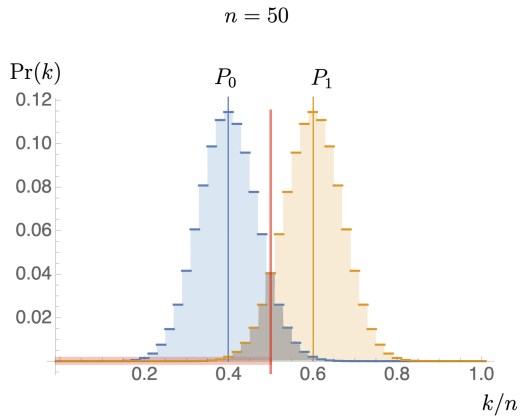


Figure 2.1.6: Illustration of likelihood test for two coins: P_0 corresponding to a probability of heads of $p_A = 0.4$, and P_1 with probability $p_B = 0.6$ for a sequence of $n = 100$ tosses (left) and $n = 400$ right. The optimal decision region for symmetric hypothesis (with equal priors) is marked in red. The total probability of error is given by gray shaded are, which in turn is dominated by the peak value at the crossing of both distributions.

In the Bayesian case, the overall probability of error is the weighted sum of the two probabilities of error,

$$P_{\mathrm{err}} \doteq \pi_1 \mathrm{e}^{-nD(P_\lambda \| P_1)} + \pi_2 \mathrm{e}^{-nD(P_\lambda \| P_2)} \doteq \mathrm{e}^{-n \min\{D(P_\lambda \| P_1), D(P_\lambda \| P_2)\}}$$

since the exponential rate is determined by the worst exponent. Since $D\left(P_\lambda || P_1\right)$ increases with λ and $D\left(P_\lambda \| P_2\right)$ decreases with λ, the maximum value of the minimum of $\{D\left(P_\lambda \| P_1\right), D\left(P_\lambda \| P_2\right)\}$ is attained when they are equal. Hence, we choose λ so that

$$D\left(P_{\lambda^*} \| P_1\right) = D\left(P_{\lambda^*} || P_2\right), \tag{2.1.43}$$

arriving to the desired result:

$$P_{\mathrm{err}} \doteq \mathrm{e}^{-nD(P_{\lambda^*}\|P_1)} \tag{2.1.44}$$

Finally, we note that condition (2.1.43) is equivalent to the optimality condition in the minimization of

$$C_{\mathrm{Ch}}(P_0, P_1) := -\min_{0\le\lambda\le 1} \log \sum_{a\in\mathcal{X}} P_0^{\lambda}(a) P_1^{1-\lambda}(a) \tag{2.1.45}$$

Indeed, if we define[4]

$$\phi(\lambda; P_0, P_1) := \log \sum_{a\in\mathcal{X}} P_0^{\lambda}(a) P_1^{1-\lambda}(a) \tag{2.1.46}$$

Then imposing $\frac{\mathrm{d}}{\mathrm{d}\lambda}\phi(\lambda; P_0, P_1) = 0$ we get

$$\begin{aligned}
0 &= \sum_a P_0^{\lambda}(a) P_1^{1-\lambda}(a) \log P_0(a) - \sum_a P_0^{\lambda}(a) P_1^{1-\lambda}(a) \log P_1(a) \\
0 &= \sum_a \frac{P_0^{\lambda}(a) P_1^{1-\lambda}(a)}{\sum_a P_0^{\lambda}(a) P_1^{1-\lambda}(a)} \log P_0(a) - \sum_a \frac{P_0^{\lambda}(a) P_1^{1-\lambda}(a)}{\sum_a P_0^{\lambda}(a) P_1^{1-\lambda}(a)} \log P_1(a) \\
0 &= \sum_a P_\lambda(a) \log P_0(a) - \sum_a P_\lambda(a) \log P_1(a) \\
0 &= \sum_a P_\lambda(a) \log \frac{P_0(a)}{P_\lambda(a)} - \sum_a P_\lambda(a) \log \frac{P_1(a)}{P_\lambda(a)} \\
0 &= D(P_\lambda \| P_0) - D(P_\lambda \| P_1)
\end{aligned}$$

■

An alternative way to prove the converse part of this theorem (optimality) is to use the following upper-bound to the probability of error as written in (2.1.5):

$$\begin{aligned}
P_{\mathrm{err}} &= \sum_x \min\{P_0(x)\eta_0, P_1(x)\eta_1\} \le \sum_x \min_{0\le\lambda\le 1} (P_0(x)\eta_0)^{\lambda} (P_1(x)\eta_1)^{1-\lambda} \\
&\le \min_{0\le\lambda\le 1} \sum_x (P_0(x)\eta_0)^{\lambda} (P_1(x)\eta_1)^{1-\lambda} \qquad (2.1.47) \\
&= \min_{0\le\lambda\le 1} \eta_0^{\lambda}\eta_1^{1-\lambda} \sum_x P_0(x)^{\lambda} P_1(x)^{1-\lambda} \qquad (2.1.48)
\end{aligned}$$

where we have used that for $a, b > 0$, $\min\{a, b\} \le a^{\lambda} b^{1-\lambda}$ for all $0 \le \lambda \le 1$ [5]. This upper bound holds for general distributions P_0, P_1. In the case of an i.i.d. $P_k(\mathbf{x}) = \prod_{i=1}^n P_k(x_i)$ it simplifies further:

$$\begin{aligned}
P_{\mathrm{err}}^{(n)} &\le \pi_0^{\lambda}\pi_1^{1-\lambda} \sum_{\mathbf{x}} \prod_{i=1}^n P_0^{\lambda}(x_i) P_1^{1-\lambda}(x_i) = \pi_0^{\lambda}\pi_1^{1-\lambda} \prod_{i=1}^n \sum_{x_i\in\mathcal{X}} P_0^{\lambda}(x_i) P_1^{1-\lambda}(x_i) \\
&= \pi_0^{\lambda}\pi_1^{1-\lambda} \left(\sum_{a\in\mathcal{X}} P_0^{\lambda}(a) P_1(a)^{1-\lambda} \right)^n \qquad (2.1.49) \\
&\le \left(\sum_{a\in\mathcal{X}} P_0^{\lambda}(a) P_1(a)^{1-\lambda} \right)^n \qquad (2.1.50)
\end{aligned}$$

Hence, we have

$$\frac{1}{n} \log P_{\mathrm{err}}^{(n)} \le \log \sum P_0^{\lambda}(x) P_1^{1-\lambda}(x)$$

Since this is true for all λ, we can take the minimum over $0 \le \lambda \le 1$ resulting in the Chernoff information bound. This proves that the exponent is no better than $C_{\mathrm{Ch}}(P_0, P_1)$.

Note that the Bayesian error exponent does not depend on the actual value of π_0 and π_1, as long as they are nonzero. Essentially, the effect of the prior is washed out for large sample sizes.

[4] This function is closely related to the so-called Rényi relative entropy

[5] This follows from Young's inequality that relates the geometric and arithmetic means: $a^{\lambda}b^{1-\lambda} \ge \lambda a + (1-\lambda)b \ge \min\{a, b\}$. The first ineq. (Young's) is a direct consequence of the concavity of the log: $\log(a^{\lambda}b^{1-\lambda}) = \lambda \log a + (1-\lambda)\log b \le \log(\lambda a + (1-\lambda)b)$.

Also notice that the exponential error rate given by Stein Lemma is larger (smaller error β_n) than that given by the Chernoff Information. The reason being that in the former, the asymmetric scenario, $\alpha_n \leq \epsilon$ for a fixed ϵ while in the latter, bayesian scenario, both types of error must decay exponential. The following theorem finds the optimal decay rate for type-II error, when the type-I is constrained to decay at least at given exponential rate.

For this purpose, we can define the error rates (if they exist) for type-I and type-II errors, respectively, given a sequence of guess functions or decision regions $A = \{A^{(n)}\}$:

$$\alpha_R(A) := \lim_{n\to\infty} -\frac{1}{n}\log \alpha_n(\phi), \quad \beta_R(A) := \lim_{n\to\infty} -\frac{1}{n}\log \beta_n(\phi) \tag{2.1.51}$$

Theorem 2.1.17 (Hoeffding bound) *(as formulated in [10]) Given two mutually absolutely continuous (have the same support) distributions P_0 and P_1. Then for each $r > 0$ there exists a sequence A of decision regions A_n such that the rate limits of the type-II and type-I errors behave like $\beta_R(\phi) \geq r$ and*

$$\alpha_R(\phi) = e(r) = \sup_{0\leq\lambda\leq 1} \frac{-\lambda r - \phi(\lambda; P_0, P_1)}{1-\lambda} \tag{2.1.52}$$

$$= \sup_{0\leq\lambda\leq 1} \frac{-\lambda r - \log \sum P_0^{\lambda}(x) P_1^{1-\lambda}(x)}{1-\lambda} \tag{2.1.53}$$

Moreover, for any sequence A such that $\alpha_R(A)$ and $\beta_R(A)$ both exist, the relation $\beta_R(A) > r$ implies $\alpha_R(A) \leq e(r)$

More graphically, Theorem 2.1.17 claims that for all possible decision functions A, the points $(\beta_R(A), \alpha_R(A))$ cannot be above the graph of $e(r)$ over $r > 0$ (see Figure 2.1.7). Note that due to Stein lemma (in particular the converse proof), the case where $\beta_R(A) > r \geq D(P_0\|P_1)$, where $e(r) = 0$, corresponds to $\alpha_n \to 1$ (which still results in a zero α_R rate!), rather than $\alpha_n \to 0$.

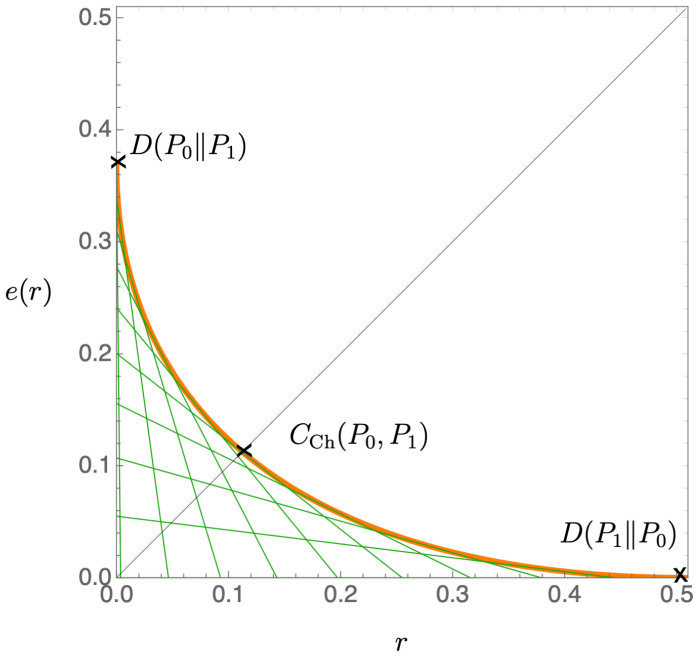


Figure 2.1.7: Optimal type I error decay rate $e(r)$ that can be attained when type II errors is constrained to decay faster than $\beta_n \doteq \mathrm{e}^{-nr}$, for the distributions $P_0 = \{0.9, .1\}$ and $P_1 = \{0.5, 0.5\}$. The various thin lines correspond to the argument in the supremum of (2.1.53) for various values of λ. For each r, there is a value of $0 \leq \lambda \leq 1$, i.e. one of the thin lines, that attains the maximum value, $e(r)$, at this point.The relative entropy is the optimal rate that can be attained (Stein Lemma) under very weak constrains on β. The Chernoff quantity is obtained when both error rates are equal.

2.1.2 Quantum error exponents

We now finally move to the quantum realm. Quantum states are in general not distinguishable. Moreover in the generic case that both states have a common support, for any quantum measurement that we might do all measurement outcomes could be produced by either states. This means that any decision that is made based on the measurement outcome will entail some non-zero probability of error. The question is now what is the optimal quantum measurement that minimizes the probability of making an error.

Matrix Analysis Intermezzo

Matrix analysis is a field on its own, that has found a lot of use in quantum information theory. We quote some useful matrix analysis definitions and tools that will prove handy. For a comprehensive look at the topic see [11, 12, 13].

1. The *spectral decomposition* of a normal operator $AA^\dagger = A^\dagger A$ is
$$A = \sum_i a_i \Pi_i^A \tag{2.1.54}$$
where $\operatorname{spec} A := \{a_i \in \mathbb{C}\}$ are its eigenvalues and $\{\Pi_i\}$ are orthogonal projectors, $\Pi_i\Pi_j = \delta_{ij}\Pi_i$, on the invariant subspaces of A. Moreover if A is hermitian ($A = A^\dagger$) then $\operatorname{spec} A := \{a_i \in \mathbb{R}\}$.

2. $A \geq 0$ is a *positive semi-definite* operator if $\forall\, |\psi\rangle\ \langle\psi| A |\psi\rangle \geq 0$ or equivalently its eigenvalues $a_i \geq 0$ (positiveness implies hermicity).

3. $A > 0$ is a *positive definite*: same as above but with strict inequality.

4. We will write $A \geq B$ if $A - B \geq 0$.

5. The *absolute value* of any linear operator B is $|B| := \sqrt{B^\dagger B}$, where we use the positive square root. For an hermitian matrix it holds $|A| = \sum_i |a_i| \Pi_i^A$.

6. Any hermitian operator admits a Jordan decomposition
$$A = A_+ - A_-, \text{ with } A_\pm > 0 \tag{2.1.55}$$
The *positive and negative parts of* A can be written in terms of its spectral decomposition (2.1.54) as
$$A_+ = \sum_{\{i:a_i>0\}} a_i \Pi_i^A = \frac{1}{2}(|A| + A) \tag{2.1.56}$$
$$A_- = \sum_{\{i:a_i<0\}} |a_i| \Pi_i^A = \frac{1}{2}(|A| - A) \tag{2.1.57}$$

7. A function $f : I \to \mathbb{R}$ can be extended over hermitian operators A with $\operatorname{spec} A \subset I \subset \mathbb{R}$ by means of their spectral decomposition: $f(A) = \sum_i f(a_i)\Pi_i^A$.

8. A function $f : I \to \mathbb{R}$ is called *operator monotone*, if for any hermitian operators A, B with $\operatorname{spec} A, \operatorname{spec} B \subset I \subset \mathbb{R}$ we have $f(A) \leq f(B)$ if $A \leq B$. Some known operator monotone functions in $[0, \infty]$ $f(t) = t^\lambda$ with $0 < \lambda \leq 1$, and in $(0, \infty)$ $f(t) = \log t$ and $f(t) = -1/t$ (see exercise 2.1.5). Of course one can generate new operator monotone functions by composing monotone operators functions (paying attention to the domains of validity). For instance $f(t) = -t^{-\lambda}$ for $0 < \lambda \leq 1$ is operator monotone since $-1/t$ and t^λ are.

9. A function $f : I \to \mathbb{R}$ is called *operator convex*, if for any hermitian operators A, B with $\operatorname{spec} A$, $\operatorname{spec} B \subset I \subset \mathbb{R}$ and $\lambda \in [0, 1]$ we have
$$f(\lambda A + (1-\lambda)B) \leq \lambda f(A) + (1-\lambda) f(B).$$

$f(t)$ is *operator concave* if $-f(t)$ is operator convex. For continuous functions $f(t)$ mapping $[0,\infty)$ onto itself: $f(t)$ operator monotone *iff* $f(t)$ is operator concave *iff* $g(t) = 1/f(t)$ is operator convex *iff* $h(t) = t/f(t)$ is operator monotone. On $(0,\infty)$ $f(t) = 1/t$ and $f(t) = -\log t$ are operator convex and on $[0,\infty)$ $f(t) = t\log t$ is operator convex.

10. Given an hermitian operator A we will denote by $\{A > 0\}$ the projector on the support of A_+, i.e.

$$\{A > 0\} = \sum_{\{i:a_i>0\}} \Pi_i^A \tag{2.1.58}$$

That is, we can write $A_+ = A\{A > 0\} = \{A > 0\}A$. By property 4 this can be extended to general matrix inequalities, e.g. $\{A - B > 0\}$ or $\{f(A) > 0\}$.

11. Trace of positive part can cast as variational (SDP) optimization (prove it!)

$$\operatorname{tr} A_+ = \max_{0\le E\le \mathbb{1}} \operatorname{tr}(AE) = \max_{\Pi:\Pi^2=\Pi} \operatorname{tr}(A\Pi) \tag{2.1.59}$$

12. The *trace-norm* is defined as

$$||A||_1 = \operatorname{tr}|A| = \operatorname{tr}\sqrt{A^\dagger A} = \sum_i |a_i| = \operatorname{tr} A_+ + \operatorname{tr} A_- \tag{2.1.60}$$

and hence

$$\operatorname{tr} A_+ = \frac{1}{2}(||A||_1 + \operatorname{tr} A) \tag{2.1.61}$$

13. Pseudo- or generalized inverse. If a given hermitian operator A does not have full support, i.e. it has some zero eigenvalues or more succinctly $\{A \neq 0\} \neq \mathbb{1}$, it cannot be inverted. Nevertheless it is convenient to define a *pseudo-inverse (or generalized inverse)* that does the job on the relevant subspace: $A^{\ominus 1} = \sum_{i:|a_i|>0} \frac{1}{a_i} P_i^A$ satisfying $A^{\ominus 1}A = AA^{\ominus 1} = \{A \neq 0\}$.

Exercise 2.1.4. *$f(t) = t^p$ for $p > 1$ is not operator monotone. Show it for $f(t) = t^2$ by giving a counter example: use*

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

to show that $A \le B$ but $A^2 \not\le B^2$.

Similarly, use the matrices

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}$$

and $\lambda = 1/2$ to show that $f(t) = t^3$ is not operator convex. This examples show that functions that are montonone (convex) as real functions need not be operator monotone (operator convex).

Exercise 2.1.5. *Let $A, B \in \mathcal{B}(\mathcal{H})$ be hermitian with $A \le B$.*

(a) Show that $XAX^+ \le XBX^\dagger$ for any $X \in \mathcal{B}(\mathcal{H})$.

(b) Show that $f(t) = 1/t$ is order-reversing for commuting, strictly positive operators: i.e. if $0 < A \le B$ with $[A, B] = 0$, then $A^{-1} \ge B^{-1}$.

(c) Use (a) and (b) to show that $f(t) = -t^{-1}$ is operator monotone.

(d) Use (c) to show that $f(t) = -(t+s)^{-1}$ is operator monotone for any $s > 0$

(e) Use (d) and the integral representation

$$\log t = \int_0^\infty \mathrm{d}s((1+s)^{-1} - (t+s)^{-1}) \tag{2.1.62}$$

to show that $f(t) = \log t$ is operator monotone.

Theorem 2.1.18 (Quantum Neyman-Pearson Lemma) *Given a binary hypothesis testing problem with the two hypotheses: ρ (H_0) vs. σ (H_1). For $T > 0$ define a two outcome POVM $E = \{E_0, E_1 = \mathbb{1} - E_0\}$ with*

$$E_0(T) = \{\rho - T\sigma > 0\} \tag{2.1.63}$$

and the corresponding error probabilities $\alpha^* = P(E_1|H_0) = \mathrm{tr}\,(E_1(T)\rho)$ and $\beta^* = P(E_0|H_1) = \mathrm{tr}\,[E_0(T)\sigma]$. Given any measurement $F = \{F_0, F_1\}$ and its associated error probabilities α_F and β_F. If $\alpha_F \leq \alpha^*$ then $\beta_F \geq \beta^*$ (and vice versa).

Proof.

$$\begin{aligned}\mathrm{tr}[E_0(\rho - T\sigma)] &= \mathrm{tr}[\{\rho - T\sigma) > 0\}(\rho - T\sigma)] &(2.1.64)\\ &= \mathrm{tr}[(\rho - T\sigma)_+] \geq \mathrm{tr}[F_0(\rho - T\sigma)] &(2.1.65)\end{aligned}$$

In the last inequality we use (11) in the matrix analysis intermezzo (see Section 2.1.2). We can rewrite the above inequality in terms of the defined error probabilities

$$\begin{aligned}(1 - \alpha^*) - T\beta^* &\geq (1 - \alpha_F) - T\beta_F &(2.1.66)\\ \alpha^* + T\beta^* &\leq \alpha_F + T\beta_F &(2.1.67)\end{aligned}$$

and since $T > 0$, $\beta_B - \beta^* \geq (\alpha^* - \alpha_B)/T$ which means that if $\alpha_B \leq \alpha^*$ the RHS is positive, which implies $\beta_B \geq \beta^*$. ■

The projective measurement $E_0(T) = \{\rho - T\sigma > 0\}$ from Theorem 2.1.18 is the quantum analog of the likelihood test which defined the acceptance region (for hypothesis H_0)

$$A = \left\{\mathbf{x} : \frac{P_0(\mathbf{x})}{P_1(\mathbf{x})} > T\right\} = \{\mathbf{x} : P_0(\mathbf{x}) - TP_1(\mathbf{x}) > 0\}$$

Note also from (2.1.67) that such class of measurements $E_0(T)$ is optimal when one wishes to minimize the linear combination of errors $\alpha^* + T\beta^*$, as for instance in **symmetric hypothesis testing** taking the parameter T to be the ratio of priors: $T = \eta_1/\eta_0$. Indeed for this case we recover the famous Helstrom bound:

$$\begin{aligned}P_{\text{err}} &= \eta_0\,\mathrm{tr}(E_1\rho) + \eta_1\,\mathrm{tr}(E_0\sigma) = \eta_0\,\mathrm{tr}[(\mathbb{1} - E_0)\rho] + \eta_1\,\mathrm{tr}(E_0\sigma) = \eta_0 + \mathrm{tr}[E_0(\eta_0\rho - \eta_1\rho_1)] = \\ &= \eta_0 - \mathrm{tr}[\{\eta_0\rho - \eta_1\sigma > 0\}(\eta_0\rho - \eta_1\sigma)] = \eta_0 + \mathrm{tr}[(\eta_0\rho - \eta_1\sigma)_+] = \frac{1}{2}(1 - ||\eta_0\rho - \eta_1\sigma||_1) \qquad (2.1.68)\end{aligned}$$

where we used (11) together with $\mathrm{tr}(\eta_0\rho - \eta_1\sigma) = \eta_0 - \eta_1$.

In **asymmetric hypothesis testing** the task is to distinguish between ρ or σ minimizing the type-II error under the constraint that the type-I error remains bounded, $\alpha \leq \varepsilon$:

$$\beta^\epsilon := \min_{0 \leq E_0 \leq \mathbb{1}} \{\mathrm{tr}(\sigma E_0) : \mathrm{tr}(\rho E_0) \geq 1 - \varepsilon\} \tag{2.1.69}$$

The quantum Neyman-Pearson lemma guarantees that the optimal E_0^* will be a projector of the form $\{\rho - T\sigma > 0\}$ for some $T > 0$.

Exercise 2.1.6. *The above quantity β^ϵ has a clear operational meaning is directly related to a* **(one-shot)** *distinguishability measure [14] the* hypothesis testing relative entropy*:*

$$D_H^\varepsilon(\rho\|\sigma) := -\log \beta^\epsilon. \tag{2.1.70}$$

Prove the data-processing inequality for the hypothesis testing relative entropy: i.e. for a quantum channel $\mathcal{N}$ and states ρ, σ,

$$D_H^\varepsilon(\rho\|\sigma) \geq D_H^\varepsilon(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))$$

Hint: Prove unitary invariance and monotonicity under tracing out.

The error exponents for the symmetric hypothesis testing could in principle be obtained from (2.1.68)

$$P_{\text{err}}^{(n)} = \frac{1}{2}(1 - ||\eta_0\rho^{\otimes n} - \eta_1\sigma^{\otimes n}||_1) \tag{2.1.71}$$

Similarly one could tackle the asymmetric setting by replacing in (2.1.69) ρ and σ by their n-copy versions. However, computing such expressions for arbitrary n and taking the limit $n \to \infty$ is impossible in practice. Instead we follow a similar approach than in the classical case, namely provide upper and lower bounds and prove that they converge in the asymptotic limit. However, we will have to rely on completely different tools.

We start by the symmetric (bayesian) setting, prove it in some detail, and we will only state the main results for the asymmetric setting (Stein & Hoeffding).

Lemma 2.1.19 *For any positive semidefinite operators A and B it holds*

$$\operatorname{tr} A\{A \leq B\} + \operatorname{tr} B\{A > B\} \leq \operatorname{tr} A^{1-\lambda}B^{\lambda} \tag{2.1.72}$$

Proof.
[Proof of Lemma 2.1.19] We start by rewriting the lhs as

$$\begin{aligned}\operatorname{tr} A\{A \leq B\} + \operatorname{tr} B\{A > B\} &= \operatorname{tr}[A(\mathbb{1} - \{A > B\})] + \operatorname{tr} B\{A > B\} \\ &= \operatorname{tr} A - \operatorname{tr}(A - B)\{A - B > 0\}\end{aligned} \tag{2.1.73}$$

Since $A - B \leq (A - B)_+$, we have $A \leq B + (A - B)_+$. Similarly, the inequality $B + (A - B)_+ \geq B$ holds. Hence, the matrix monotonicity of $f(t) = t^{\lambda}$ (see property 8)

$$A^{\lambda} \leq (B + (A - B)_+)^{\lambda} \tag{2.1.74}$$

$$(B + (A - B)_+)^{\lambda} - B^{\lambda} \geq 0 \tag{2.1.75}$$

$$(B + (A - B)_+)^{1-\lambda} \geq B^{1-\lambda} \tag{2.1.76}$$

Hence,

$$\begin{aligned}&\operatorname{tr} A - \operatorname{tr} A^{1-\lambda}B^{\lambda} = \operatorname{tr} A^{1-\lambda}(A^{\lambda} - B^{\lambda}) \\ &\leq \operatorname{tr} A^{1-\lambda}\left((B + (A - B)_+)^{\lambda} - B^{\lambda}\right) \\ &\leq \operatorname{tr} (B + (A - B)_+)^{1-\lambda}\left((B + (A - B)_+)^{\lambda} - B^{\lambda}\right) \\ &= \operatorname{tr} (B + (A - B)_+) - \operatorname{tr} (B + (A - B)_+)^{1-\lambda}B^{\lambda} \\ &\leq \operatorname{tr} (B + (A - B)_+) - \operatorname{tr} B^{1-\lambda}B^{\lambda} \\ &= \operatorname{tr} B + \operatorname{tr}(A - B)_+ - \operatorname{tr} B = \operatorname{tr}(A - B)_+\end{aligned} \tag{2.1.77}$$

where the inequalities (2.1.74),(2.1.75) and (2.1.76) have been used in that order. Thus applying the inequlaity (2.1.77) to (2.1.73) we find the desired results:

$$\operatorname{tr} A\{A \leq B\} + \operatorname{tr} B\{A > B\} \leq \operatorname{tr} A^{1-\lambda}B^{\lambda} \tag{2.1.78}$$

■

Theorem 2.1.20 (Quantum Chernoff bound) *The best achievable exponent in the Bayesian probability of error in discriminating ρ and σ for any fixed priors η_0 and η_1 is*

$$C^*(\rho, \sigma) = -\lim_{n\to\infty} \frac{1}{n} \log P_{\text{err}}^{(n)} = C_{\text{Ch}}(\rho, \sigma) := -\min_{0\leq\lambda\leq 1} \log \operatorname{tr} \rho^{\lambda}\sigma^{1-\lambda} \tag{2.1.79}$$

PROOF.
The direct or **achievability part** is a direct consequence of Lemma 2.1.19. Indeed starting from (2.1.68) and taking $A = \eta_0\rho$ and $B = \eta_1\sigma$ with $E_0 = \{\eta_0\rho - \eta_1\sigma > 0\}$ we have

$$P_{\text{err}} = \eta_0 \operatorname{tr}(E_1\rho) + \eta_1 \operatorname{tr}(E_0\sigma) \leq \operatorname{tr}(\eta_0\rho)^{1-\lambda}(\eta_1\sigma)^{\lambda} = \eta_0^{1-\lambda}\eta_1^{\lambda} \operatorname{tr}(\rho^{1-\lambda}\sigma^{\lambda}) \tag{2.1.80}$$

which provides a general (non-asympotic) upper bound to the probability or error. A clear advantage of this bound is that it is easily computable for n-copy states,

$$\begin{aligned} P_{\text{err}}^{(n)} &= \eta_0 \operatorname{tr}(E_1\rho^{\otimes n}) + \eta_1 \operatorname{tr}(E_0\sigma^{\otimes n}) && (2.1.81)\\ &\leq \eta_0^{1-\lambda}\eta_1^{\lambda} \operatorname{tr}[\rho^{\otimes n})^{1-\lambda}(\sigma^{\otimes n})^{\lambda}] && (2.1.82)\\ &\leq \eta_0^{1-\lambda}\eta_1^{\lambda}(\operatorname{tr}(\rho^{1-\lambda}\sigma^{\lambda}))^n && (2.1.83) \end{aligned}$$

where we have used that $(A \otimes B)^{\lambda} = A^{\lambda} \otimes B^{\lambda}$ and $\operatorname{tr}(A \otimes B) = (\operatorname{tr} A)(\operatorname{tr} B)$. Now it is immediate to obtain a lower bound on the error rate

$$C^*(\rho, \sigma) = -\lim_{n\to\infty} \frac{1}{n} \log P_{\text{err}}^{(n)} \geq -\log \inf_{0\leq\lambda\leq 1} \operatorname{tr}(\rho^{1-\lambda}\sigma^{\lambda}) \tag{2.1.84}$$

where we have used that the inequality holds for all $0 \leq \lambda \leq 1$ to write down the tightest upper-bound to the error rate.

We now provide the **converse or optimality part** of the proof. That is we need to prove that all possible sequence of measurements $\{E_0^n\}$ lead to a worse (smaller) error rate than the Chernoff distance. The proof, is based on relating the quantum to the classical case by using a special mapping from a pair of density matrices (ρ, σ) to a pair of probability distributions (P, Q): Given ρ and σ written in their respective diagonal basis

$$\rho = \sum_{i=1}^{d} p_i \, |e_i\rangle\langle e_i| \quad \sigma = \sum_{j=1}^{d} q_j \, |f_j\rangle\langle f_j|$$

we map these density operators to the the "pinched" probability distributions:

$$P_{i,j} = p_i |\langle e_i|f_j\rangle|^2, \quad Q_{i,j} = q_j |\langle e_i|f_j\rangle|^2 \tag{2.1.85}$$

with $1 \leq i, j \leq d$. $P_{i,j}$ can be understood as the probability that results from measuring ρ in its own eigenbasis followed by a measurement in the eigenbasis of σ and getting the outcomes corresponding to $|e_i\rangle$ and $|f_j\rangle$. Similarly $Q_{i,j}$ corresponds to the outcome probability resulting from measuring σ in its own eigenbasis followed by a measurement in the eigenbasis of ρ and getting the jth and ith outcome respectively. Notice that the order of measurements is different in each case (but not the order of the indices), so P_{ij}, Q_{ij} do not represent the probability distributions that could correspond to a particular measurement strategy. Nevertheless, they are proper distributions mathematically speaking and that's all we will need.

In particular, $P_{i,j}$ and $Q_{i,j}$ defined in (2.1.85) enjoy a very nice property for many derived scalar quantities

$$\sum_{i,j} P_{i,j}^{1-\lambda} Q_{i,j}^{\lambda} = \sum_{ij} (p_i |\langle e_i|f_j\rangle|^2)^{1-\lambda} (q_j |\langle e_i|f_j\rangle|^2)^{\lambda} = \sum_{i,j} p_i^{1-\lambda} q_j^{\lambda} |\langle e_i|f_j\rangle|^2 = \operatorname{tr} \rho^{1-\lambda}\sigma^{\lambda} \tag{2.1.86}$$

$$D(P\|Q) = \ldots [\text{similarly}] \ldots = D(\rho\|\sigma) \tag{2.1.87}$$

Here, $D(\rho\|\sigma)$ is the QUANTUM RELATIVE ENTROPY defined as

$$D(\rho\|\sigma) := \begin{cases} \operatorname{tr}[\rho(\log\rho - \log\sigma)], & \text{if } \operatorname{supp}\rho \leq \operatorname{supp}\sigma \\ +\infty, & \text{otherwise} \end{cases} \tag{2.1.88}$$

In addition, from (2.1.5) we know that the minimum error probability for the discrimination of the classical distributions is P, Q can be written as

$$P_{\text{err}}(P, Q) = \sum_{i,j} \min\{P_{ij}\eta_0, Q_{ij}\eta_1\} = \sum_{i,j} \min\{p_i\eta_0, q_j\eta_1\} |\langle e_i|f_j\rangle|^2 \tag{2.1.89}$$

Now we can actually relate the rhs of this equation to the quantum error probability obtained by a projective measurement[6] $E_0 = \Pi$ with $\Pi^2 = \Pi$:

$$P_{\text{err}}(\rho,\sigma) = \eta_0 \operatorname{tr}[(\mathbb{1}-\Pi)\rho] + \eta_1 \operatorname{tr}(\Pi\sigma) = \sum_{i,j} p_i |\langle e_i|(\mathbb{1}-\Pi)|f_i\rangle|^2 + \sum_{i,j} q_j |\langle e_i|\Pi|f_i\rangle|^2 = \tag{2.1.90}$$

$$\geq \sum_{i,j} \min\{p_i,q_j\}(|\langle e_i|(\mathbb{1}-\Pi)|f_i\rangle|^2 + |\langle e_i|\Pi|f_i\rangle|^2) \tag{2.1.91}$$

$$\geq \frac{1}{2}\sum_{i,j} \min\{p_i,q_j\}(|\langle e_i|(\mathbb{1}-\Pi)|f_i\rangle| + |\langle e_i|\Pi|f_i\rangle|)^2 \tag{2.1.92}$$

$$= \frac{1}{2}\sum_{i,j} \min\{p_i,q_j\}|\langle e_i|f_i\rangle|^2 \tag{2.1.93}$$

where in (2.1.90) we have used

$$\operatorname{tr}(\sigma\Pi) = \operatorname{tr}(\sigma\Pi^2) = \operatorname{tr}[\sigma\Pi(\sum_i |e_i\rangle\langle e_i|)\Pi] = \sum_j q_j |\langle e_i|\Pi|f_i\rangle|^2$$

and similarly for $\operatorname{tr}[\rho(\mathbb{1}-\Pi)] = \sum_{i,j} p_i |\langle e_i|(\mathbb{1}-\Pi)|f_i\rangle|^2$. In (2.1.90) we have used that[7] for $a,b\in\mathbb{R}$, $a^2+b^2 \geq \frac{1}{2}(a+b)^2$.

Putting this together with (2.1.89) we arrive at the elegant result,

$$P_{\text{err}}(\rho,\sigma) \geq \frac{1}{2} P_{\text{err}}(P,Q) \tag{2.1.94}$$

At this point we only need to see what happens to the n-copy version on this inequality. Quite crucially, it is straightforward to check that pinched probabilities $P^{(n)}$, $Q^{(n)}$ corresponding to the n-copy states $(\rho^{\otimes},\sigma^{\otimes n})$ are just i.i.d. versions of the ones corresponding to (ρ,σ):

$$P^{(n)}(i_1,j_1,\dots i_n,j_n) = \prod_{k=1}^n P(i_k,j_k) \tag{2.1.95}$$
$$Q^{(n)}(i_1,j_1,\dots i_n,j_n) = \prod_{k=1}^n Q(i_k,j_k).$$

Hence,

$$P_{\text{err}}(\rho^{\otimes},\sigma^{\otimes n}) \geq \frac{1}{2} P_{\text{err}}(P^{(n)},Q^{(n)}) \tag{2.1.96}$$

Now we can obtain the desired bound on the error rate by taking the log dividing by n and taking the $n\to\infty$ limit

$$C(\rho,\sigma) := -\lim_{n\to\infty}\frac{1}{n}\log P_{\text{err}}(\rho^{\otimes},\sigma^{\otimes n}) \geq -\lim_{n\to\infty}\frac{1}{n}\log P_{\text{err}}(P^{(n)},Q^{(n)}) \tag{2.1.97}$$

$$= C_{\text{Ch}}(P,Q) = -\inf_{0\leq\lambda\leq 1}\log\sum_{i,j} P_{ij}^{\lambda} Q_{ij}^{1-\lambda} \tag{2.1.98}$$

$$= -\log\inf_{0\leq\lambda\leq 1}\operatorname{tr}\rho^{1-\lambda}\sigma^{\lambda} \tag{2.1.99}$$

where in (2.1.98) we have used the classical Chernoff bound (2.1.45) owing to the fact that the joint distributions $P^{(n)}$ and $Q^{(n)}$ are i.i.d.; and in the last equality we have used property (2.1.86) of the pinched distributions. This concludes the proof as we have shown that the upper and lower bounds coincide. ∎

[6] We know from the quantum Neyman-Pearson lemma (Theorem 2.1.18) that we w.l.o.g. can optimize P_{err} over projectors

[7] This can be readily seen by writing obtained $0 \leq (a-b)^2 = a^2+b^2-2ab \implies 2ab \leq a^2+b^2$ which after adding a^2+b^2 on both sides becomes $2(a+b)^2 \leq a^2+b^2$

Theorem 2.1.21 (Quantum Stein's Lemma [15, 16]) *The best achievable exponent in the binary asymmetric hypothesis testing with ρ and σ is*

$$\beta_R := \lim_{n\to\infty} -\frac{1}{n}\log\beta_n^\epsilon = D(\rho\|\sigma) \tag{2.1.100}$$

where the

$$\beta_n^\epsilon := \min_{0\le E_0\le\mathbb{1}}\{\mathrm{tr}(\sigma^{\otimes n}E_0) : \mathrm{tr}(\rho^{\otimes n}E_0)\ge 1-\varepsilon\}$$

and where $D(\rho\|\sigma)=\mathrm{tr}[\rho(\log\rho-\log\sigma)]$) is the quantum relative entropy.

Theorem 2.1.22 (Quantum Hoeffding bound) *Given two faithful states (equal support) $\rho>0$ and $\sigma>0$, then for each $r>0$ there exists a sequence of measurements $E=\{E_0^n\}$ such that the rate limits of the type-II and type-I errors behave like $\beta_R(\phi)\ge r$ and*

$$\alpha_R(E)=e_Q(r)=\sup_{0\le\lambda\le 1}\frac{-\lambda r-\log\mathrm{tr}\,\rho^\lambda\sigma^{1-\lambda}}{1-\lambda} \tag{2.1.101}$$

If $\beta_R(\phi)>r\ge D(\rho\|\sigma)$ then, as expected $e_Q(r)=0$, but actually α_n converges to one exponentially fast (at a rate given by the so-called strong converse exponents, see e.g. [17]).

Theorem 2.1.23 (Quantum Sanov theorem) *Let the null hypothesis H_0 correspond to a family Γ of density operators on $\mathcal{H}$. Let the alternative hypothesis H_1 be represented by a single density operator σ. Then there exists a sequence of orthogonal projections $\Pi=\{\Pi^n\}$ such that for all $\rho\in\Gamma$ the corresponding type-I error vanishes asymptotically, i.e.*

$$\lim_{n\to\infty}\mathrm{Tr}\left[\rho^{\otimes n}\Pi_n\right]=0 \tag{2.1.102}$$

while the type-II error rate limit $\beta_R(\Pi)$ is equal to the relative entropy distance from Γ to σ :

$$S(\Gamma\|\sigma):=\inf_{\rho\in\Gamma}D(\rho\|\sigma)$$

Moreover $S(\Gamma\|\sigma)$ is the upper bound on type-II error (upper) rate limit, for any sequence Π of POVMs satisfying the constraint (2.1.102).

This way of "quantizing" Sanov's theorem provides a solution to a composite hypothesis testing. When only the null hypothesis is composite. The problem becomes much more involved when the alternative hypothesis (or both) are composite (see [18]).

We remark here, that although we have derived ultimate quantum bounds, the proofs are not constructive, i.e. they do not provide explicit measurement sequences that attain such limits. The optimal test to discriminate $\rho^{\otimes n}$ and $\sigma^{\otimes n}$ makes use of joint measurements (not implementable by local measurements on every copy). However, in any case, the permutational symmetry of n-copy states guarantees that the optimal collective measurement can be implemented efficiently (with a polynomial-size circuit). Collective measurements dominate local ones even in the asymptotic sense: if one restricts the measurement sequence $\{E_n\}$ the obtained error rates are in general lower the ultimate quantum limits computed here. For Stein's setting Hayashi [19] gave an explicit scheme to attain the asymptotic rate given by the quantum relative entropy. For the other settings it remains an open problem.

There is a simple explicit construction of POVM, valid for binary as well as multiple hypothesis testing, that does a good enough job to assess the performance of some tasks. This is the so-called square-root or "pretty-good" measurement. The basic idea is to pick the POVM E_i proportional to the state it ought to detect $E_i=\mu_i\rho_i$, so as to guarantee a good enough overlap $p(E_i|\rho_i)=\mu_i\,\mathrm{tr}(\rho_i E_i)$. However, in general such a POVM will not fulfill the completeness relation $\sum_i E_i=\mathbb{1}$. The pretty good measurement takes this idea while avoiding this nuisance.

Definition 2.1.24 (Pretty good or square root measurement) *For any ensemble of states* $\{\rho_i\}_{i=1}^m$ *with prior probability distribution* $\{\eta_i\}_{i=1}^m$ *and ensemble average* $\Omega = \sum_{i=1}^m \eta_i \rho_i$ *one can define a POVM called the pretty good measurement, with elements:*

$$E_i = \eta_i \Omega^{-1/2} \rho_i \Omega^{-1/2} \text{ for } i = 1, \ldots, m \tag{2.1.103}$$

It is immediate to check that this is indeed a valid POVM i.e. $E_i \geq 0$ and $\sum_{i=1}^m E_i = \mathbb{1}$. The pretty good measurement is optimal for sets of symmetric states and for pure states it coincides with the least-squares measurement [20]. In addition, it gives a powerful bound on the optimal success probability [21]

$$\left(P_{\text{succ}}^{\text{opt}}\right)^2 \leq P_{\text{succ}}^{\text{PGM}} \leq P_{\text{succ}} \tag{2.1.104}$$

Or in terms of the error probability

$$\frac{1}{2} P_{\text{err}}^{\text{PGM}} \leq \frac{1}{2}(1 - \left(1 - P_{\text{err}}^{\text{opt}}\right)^2) \leq P_{\text{err}}^{\text{opt}} \leq P_{\text{err}}^{\text{PGM}} \tag{2.1.105}$$

In particular, the pretty good measurement attains the optimal asymptotic error exponent.

Finally, we also give a useful decoupling upper bound for the multiple-hypothesis error probability in terms of pairwise state-distinguishability measures.

$$P_{\text{err}}^{\text{PGM}} \leq \frac{1}{2} \sum_{(i,j):i \neq j} \sqrt{\eta_i \eta_j} F\left(\rho_i, \rho_j\right). \tag{2.1.106}$$

where the fidelity is given by

$$F(\rho, \sigma) := ||\sqrt{\rho}\sqrt{\sigma}||_1 \tag{2.1.107}$$

2.2 Discrimination of channels

We now move to hypothesis testing in cases where each hypothesis is characterized by a quantum channel (instead of state). For instance, in quantum sensing, the presence of nearby mass might alter the dynamics of a quantum sensor and therefore detecting the presence of the mass amounts to distinguishing between the default and perturbed dynamics of the system (two quantum channels). As compared to the discrimination of quantum states, finding the optimal quantum strategy to discriminate between two channels has the complication that in addition to the optimization over quantum measurements one needs to optimize over possible input states (see figure 2.2.1). Before formalizing the problem we will summarize the basic tools to describe quantum channels and quantum testers.

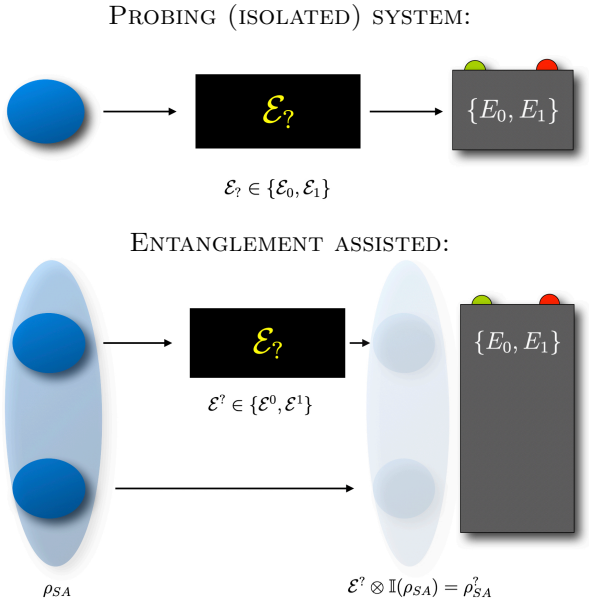


Figure 2.2.1: Schematic representation of a quantum channel discrimination strategy, with (bottom) and without (top) use of entangled auxiliary system.

2.2.1 Recap on Bipartite Pure States systems as Matrices

As will be shown later there is an extremely useful characterization of quantum channels in terms of bipartite quantum states, the so-called Choi-Jamiolkowski states (see 2.2.2). For this, and other purposes is it very convenient to represent bipartite pure states as matrices and vice versa.

Given bipartite state $|\Psi\rangle_{AB} \in \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$:

$$|\Psi\rangle_{AB} = \sum_{i,j} C_{ij} |i\rangle_A |j\rangle_B \equiv |C\rangle \tag{2.2.108}$$

Where $\{|i\rangle\}_{i=1}^{d_A}$ and $\{|j\rangle\}_{j=1}^{d_B}$ are basis of $\mathcal{H}_\mathcal{A}$ and $\mathcal{H}_\mathcal{B}$ respectively. We will assign the $d_A d_B$ components C_{ij} to a $d_A \times d_B$ matrix C and our state will be labeled as $|C\rangle$. This operation of switching from vector representation of bipartite state to a matrix can be understood as "flipping" the $|j\rangle$ into a ket $\langle j|$, turning $|\Psi\rangle_{AB} = \sum_{i,j} C_{ij} |i\rangle_A |j\rangle_B$ into $\hat{C} = \sum_{i,j} C_{ij} |i\rangle_A \langle j|_B$.

Now note that if we applying a given operator A on the first subsystem in corresponds to a simple

matrix multiplication

$$(A\otimes\mathbb{1})\left|C\right\rangle=\left(\sum_{m,n}A_{mn}\left|m\right\rangle\!\left\langle n\right|\otimes\mathbb{1}\right)\sum_{i,j}C_{ij}\left|i\right\rangle\left|j\right\rangle=\sum_{\substack{i,j\\m,n}}A_{mn}C_{ij}\left|m\right\rangle\underbrace{\left\langle n|i\right\rangle}_{=\delta_{ni}}\left|j\right\rangle=$$
$$=\sum_{i,j,m}A_{mi}C_{ij}\left|m\right\rangle\left|j\right\rangle=\sum_{m,j}(AC)_{jm}\left|m\right\rangle\left|j\right\rangle=\left|AC\right\rangle \tag{2.2.109}$$

Similarly the operation of B on the second subsystem leads to:

$$(\mathbb{1}\otimes B)\left|C\right\rangle=\left(\mathbb{1}\otimes\sum_{r,s}B_{rs}\left|r\right\rangle\!\left\langle s\right|\right)\sum_{i,j}C_{ij}\left|i\right\rangle\left|j\right\rangle=\sum_{\substack{i,j\\r,s}}C_{ij}B_{rs}\left|i\right\rangle\left|r\right\rangle\underbrace{\left\langle s|j\right\rangle}_{=\delta_{sj}}=$$
$$=\sum_{i,j,r}C_{ij}B_{rj}\left|i\right\rangle\left|r\right\rangle=\sum_{i,j,r}C_{ij}(B_{jr})^{T}\left|i\right\rangle\left|r\right\rangle=\sum_{i,r}(CB^{T})_{ir}\left|i\right\rangle\left|r\right\rangle=\left|CB^{T}\right\rangle \tag{2.2.110}$$

And putting it all together we can write

$$(A\otimes B)\left|C\right\rangle=\left|ACB^{T}\right\rangle \tag{2.2.111}$$

The reduced states (or marginals) of $\left|C\right\rangle$ when one of the sub-systems is traced also reduces to a very simple matrix multiplication:

$$\rho_A=\mathrm{tr}_B(\rho_{AB})=\mathrm{tr}_B\left|C\right\rangle\!\left\langle C\right|=\mathrm{tr}_B\left(\sum_{\substack{i,j\\i',j'}}C_{ij}C^{*}_{i'j'}\left|i\right\rangle_A\!\left\langle i'\right|\otimes\left|j\right\rangle_B\!\left\langle j'\right|\right)=\sum_{\substack{i,j\\i',j'}}C_{ij}C^{*}_{i'j'}\left|i\right\rangle_A\!\left\langle i'\right|\left\langle j'|j\right\rangle_B$$
$$=\sum_{i,j,i'}C_{ij}C^{*}_{i'j}\left|i\right\rangle_A\!\left\langle i'\right|=\sum_{i,i',j}C_{ij}C^{\dagger}_{ji'}\left|i\right\rangle_A\!\left\langle i'\right|=\sum_{i,i'}(CC^{\dagger})_{ii'}\left|i\right\rangle\!\left\langle i'\right|=CC^{\dagger} \tag{2.2.112}$$

$$\rho_B=\mathrm{tr}_A(\rho_{AB})=\mathrm{tr}_A\left|C\right\rangle\!\left\langle C\right|=\mathrm{tr}_A\left(\sum_{\substack{i,j\\i',j'}}C_{ij}C^{*}_{i'j'}\left|i\right\rangle_A\!\left\langle i'\right|\otimes\left|j\right\rangle_B\!\left\langle j'\right|\right)=\sum_{\substack{i,j\\i',j'}}C_{ij}C^{*}_{i'j'}\left\langle i'|i\right\rangle_A\left|j\right\rangle_B\!\left\langle j'\right|$$
$$=\sum_{i,j,j'}C_{ij}C^{*}_{ij'}\left|j\right\rangle_B\!\left\langle j'\right|=\sum_{i,j,j'}C^{T}_{ji}C^{*}_{ij'}\left|j\right\rangle_A\!\left\langle j'\right|=C^{T}C^{*}=(C^{\dagger}C)^{*} \tag{2.2.113}$$

That is, we can readily obtain the reduced states directly form the matrix C, by multiplying it by its adjoint. A first very interesting result follows from this relation and the singular value decomposition of C.

For any $d_A\times d_B$ matrix C there exists be a pair of orthonormal matrices U and V such that $C=UDV^{T}$, where D a diagonal matrix with positive (and ordered) eigenvalues: $\lambda_1\geq\lambda_2\ldots\lambda_r>0$. Therefore:

$$\left|\psi\right\rangle_{AB}=\left|C\right\rangle=\left|UDV^{T}\right\rangle=(U\otimes V)\left|D\right\rangle=(U\otimes V)\sum_{k}\lambda_k\left|k\right\rangle\left|k\right\rangle\equiv\sum_{k=1}^{r}\lambda_k\left|a_k\right\rangle\left|b_k\right\rangle \tag{2.2.114}$$

Where $\{\left|a_k\right\rangle\equiv U\left|k\right\rangle\}_{k=1}^{d_A}$ and $\{\left|b_k\right\rangle\equiv V\left|k\right\rangle\}_{k=1}^{d_B}$ are an orthonormal basis of $\mathcal{H}_{\mathcal{A}}$ and $\mathcal{H}_{\mathcal{B}}$ given by the columns of U and V respectively[8].This new way of defining $\left|\psi\right\rangle_{AB}$ is known as *Schmidt Decomposition* and $\{\lambda_k\}_{k=1}^{r}$ as the *Schmidt Coefficients*. All entanglement properties, since those are invariant under local unitaries, are only a function of these coefficients.

Moreover using (2.2.112) and (2.2.113) we find that given a pure bipartite state $\mathrm{tr}_A(\rho_{AB})=\left|C\right\rangle\!\left\langle C\right|$ characterized by $C=UDV^{T}$ its reductions are written as:

$$\rho_A=\mathrm{tr}_B\left|C\right\rangle\!\left\langle C\right|=CC^{\dagger}=UD(V^{\dagger}V)^{*}DU^{\dagger}=UD^2U^{\dagger}=\sum_{k}\lambda_k^2\left|a_k\right\rangle\!\left\langle a_k\right| \tag{2.2.115}$$

$$\rho_B=\mathrm{tr}_A\left|C\right\rangle\!\left\langle C\right|=(C^{\dagger}C)^{*}=VD(U^{\dagger}U)^{*}DV^{\dagger}=VD^2V^{\dagger}=\sum_{k}\lambda_k^2\left|b_k\right\rangle\!\left\langle b_k\right| \tag{2.2.116}$$

[8]Defining such matrices in the canonical basis $\{\left|k\right\rangle\}$

We hence see that both reductions have the same spectra (eigenvalues), which coincide with the Schmidt Coefficients, while their eigenbasis give the Schmidt basis $\{|a_k\rangle\}$ and $\{|b_k\rangle\}$. So, if the marginals don't have degenarate spectra $\lambda_i \neq \lambda_j$ one can readily construct their Schmidt decomposition from the marginals. For degenerate eigenvalues the choice of basis of ρ_A and ρ_B and pairing is not unique, and one needs to find the Schmidt decomposition more explicitly.

Note that you can also revert this relation (in a non-unique way). That is, given an arbitrary mixed state $\rho \in \mathcal{H}_\mathcal{A}$ it is always possible to find a pure state $|\Psi\rangle$ of an extended system AB such that its reduction or marginal is equal to ρ:

$$\rho_A = \mathrm{tr}_B |\Psi\rangle\langle\Psi| = \rho \tag{2.2.117}$$

Such a state $|\Psi_\rho\rangle = |\Psi\rangle_{AB}$ is called a **purification of** ρ. An immediate way to construct such a purification of ρ is to use $\rho_A = CC^\dagger$ and pick $C = \sqrt{\rho}$. That is, following our notation: $|\Psi_\rho\rangle = \left|\sqrt{\rho}\right\rangle$. Note that indeed such a purification is not unique as any unitary acting on B will not affect ρ_A. If ρ is non-degenerate, all purifications must be of this form $|\Psi\rangle = \mathbb{1} \otimes V \left|\sqrt{\rho}\right\rangle$, i.e. purifications are unique modulo unitaries on the ancillary system B.

This leads to the **GHJW theorem** that states that there are infinitely many ensembles compatible with a mixed state:

$$\rho = \sum_{i=1}^{r} \left|\tilde{\psi}_i\right\rangle\!\left\langle\tilde{\psi}_i\right| = \sum_{k=1}^{m} \left|\tilde{\phi}_k\right\rangle\!\left\langle\tilde{\phi}_k\right| \quad \Leftrightarrow \quad \left|\tilde{\phi}_k\right\rangle = \sum_l V_{ki} \left|\tilde{\psi}_i\right\rangle \tag{2.2.118}$$

This theorem can be readily be understood by interpreting both ensembles as emerging from different measurements on the ancillary system of a purification of ρ: $\left|\sqrt{\rho}\right\rangle = \sum_i \left|\tilde{\psi}_i\right\rangle_A |i\rangle_B$. The first (minimal) ensemble is realized by measuring in the Schmidt basis $|i\rangle_B$, while the second is obtained by measuring in $|f_k\rangle_B = V^\dagger |k\rangle_B$,where the basis of system B can be extended beyond r, $d_B = m$ with $m \geq r$. That is

$$\left|\tilde{\phi}_k\right\rangle_A = {}_B\langle f_k|\sqrt{\rho}\rangle_{AB} = \sum_i {}_B\langle k| V |i\rangle_B \left|\tilde{\psi}_i\right\rangle_A = \sum_i V_{ki} \left|\tilde{\psi}_i\right\rangle_A$$

2.2.2 Recap: Choi-Jamiolkowski Isomorphism

With the notation seen in the previous section we are ready to take a step further in the description of quantum channels. Concretely, we are going to introduce an alternative way of representing the action of a linear map $\Lambda(\rho_{\text{in}}) = \rho'_{\text{out}}$, and in particular a full characterization of completely positive trace-preserving (CPTP) maps. This result, known as *Choi-Jamiolkowski state Isomorphism*, establishes a one-to-one correspondence between quantum maps and quantum state (its *Choi-Jamiolkowski state*).

We start by noticing that the action of a linear a map, $\Lambda : \mathcal{B}(\mathcal{H}_{\text{in}}) \to \mathcal{B}(\mathcal{H}_{\text{out}})$[9], can be fully characterized (mathematically) by its action on an operator basis, e.g. $\Lambda : |i\rangle\langle j| \to \Lambda(|i\rangle\langle j|)$ for all $i, j \in [1, d_{\text{in}}]$, and where $\Lambda(|i\rangle\langle j|) \in \mathcal{B}(\mathcal{H}_{\text{out}})$. So that for any input state state $\rho = \sum_{ij} \rho_{ij} |i\rangle\langle j| \in \mathcal{B}(\mathcal{H}_{\text{in}})$ we can compute $\rho' = \sum_{ij} \rho_{ij} \Lambda(|i\rangle\langle j|)$.

It is immediate to see that this very same information can be effectively encoded in bipartite state ρ'_{SA} involving the input system (S) and an ancillary system (A).

$$\rho'_{SA} = \Lambda \otimes \mathbb{1}(\left|\Phi^+\right\rangle\!\left\langle\Phi^+\right|) = \frac{1}{d_{\text{in}}} \Lambda \otimes \mathbb{1}(\sum_{ij} |i\rangle\langle j| \otimes |i\rangle\langle j|) = \tag{2.2.119}$$

$$= \frac{1}{d_{\text{in}}} \sum_{ij} \Lambda(|i\rangle\langle j|) \otimes |i\rangle\langle j| \in \mathcal{B}(\mathcal{H}_{\text{out}} \otimes \mathcal{H}_{\text{in}}) \tag{2.2.120}$$

That is, the so-called *Choi-state* of the channel Λ can be obtained by sending through the channel one party of the maximally entangled state $|\Phi^+\rangle = \frac{1}{\sqrt{d_{\text{in}}}} \sum_{i=1}^{d_{\text{in}}} |ii\rangle$. Notice that that operator ρ'_{SA} is a proper state (positive and unit trace) if Λ is CPTP. Also, note that we have allowed for the possibility that the input and output hilbert spaces of the map differ.

[9] The set of all bounded linear operators on $\mathcal{H}$ is denoted by $\mathcal{B}(\mathcal{H})$

For every Choi state ρ'_{SA} defined above we can readily obtain the action of the map Λ on an arbitrary state $\rho \in \mathcal{B}(\mathcal{H}_{\text{in}})$, as

$$\text{tr}_{\text{in}}\left[(\mathbb{1} \otimes \rho^T) J_\Lambda\right] = \Lambda(\rho) \in \mathcal{B}(\mathcal{H}_{\text{out}}) \tag{2.2.121}$$

where the $J_\Lambda := d_{\text{in}} \rho_{SA}$ is the so-called Choi operator or process matrix:

$$J_\Lambda := \sum_{i,j=1}^{d_{\text{in}}} \Lambda\left(|i\rangle\langle j|\right) \otimes |i\rangle\langle j| \in \mathcal{B}(\mathcal{H}_{\text{out}} \otimes \mathcal{H}_{\text{in}}) \tag{2.2.122}$$

More explicitly,

$$\text{tr}_{\text{in}}\left[(\mathbb{1} \otimes \rho^T) \sum_{i,j=1}^{d_{\text{in}}} \Lambda(|i\rangle\langle j|) \otimes |i\rangle\langle j|\right] = \text{tr}_{\text{in}}\left[\sum_{i,j=1}^{d_{\text{in}}} \Lambda(|i\rangle\langle j|) \otimes \rho^T |i\rangle\langle j|\right] =$$

$$= \sum_{i,j=1}^{d_{\text{in}}} \Lambda(|i\rangle\langle j|) \otimes \text{tr}_{\text{in}}(\rho^T |i\rangle\langle j|) = \sum_{i,j=1}^{d_{\text{in}}} \Lambda(|i\rangle\langle j|) \otimes \langle j| \rho^T |i\rangle = \sum_{i,j=1}^{d_{\text{in}}} \rho_{ij} \Lambda(|i\rangle\langle j|) = \Lambda(\rho)$$

For every positive operator $\tilde{J} \geq 0$ in $\mathcal{B}(\mathcal{H}_{\text{out}} \otimes \mathcal{H}_{\text{in}})$ we can use the prescription in (2.2.121) to define a positive linear map $\tilde{\Lambda}(\rho) := \text{tr}_{\text{in}}\left[(\mathbb{1} \otimes \rho^T)\tilde{J}\right]$. The operator $\tilde{J}$ is positive semi-definite *if and only if* $\tilde{\Lambda}$ is completely positive. This follows from the fact that $\Lambda(\rho)$ is is completely positive iff $(\Lambda \otimes \mathbb{1}_d)(|\mathbb{1}\rangle\langle\mathbb{1}|) \geq 0$. [10]

Moreover, we next proof that

$$\text{tr}\,\tilde{\Lambda}(\rho) = \text{tr}\,\rho \text{ if and only if } \text{tr}_{\text{out}}\,\tilde{J} = \mathbb{1}_{\text{in}} \tag{2.2.123}$$

or more generally

$$\text{tr}\,\tilde{\Lambda}(\rho) \leq \text{tr}\,\rho \text{ if and only if } \text{tr}_{\text{out}}\,\tilde{J} \leq \mathbb{1}_{\text{in}} \tag{2.2.124}$$

Proof: $\text{tr}\,\Lambda(\rho) = \text{tr}_{\text{out}}\,\Lambda(\rho) = \text{tr}_{\text{out}}\,\text{tr}_{\text{in}}[(\mathbb{1}_{\text{out}} \otimes \rho^T)\tilde{J}] = \text{tr}_{\text{in}}\left(\rho^T\,\text{tr}_{\text{out}}\,\tilde{J}\right) = \text{tr}(\rho^T\,\text{tr}_{\text{out}}\,\tilde{J})$. If $\text{tr}_{\text{out}}\tilde{J} \leq \mathbb{1}_{\text{in}}$, then $\text{tr}\,\Lambda(\rho) \leq \text{tr}(\rho^T\mathbb{1}_{\text{in}}) = \text{tr}\,\rho^T = \text{tr}\,\rho$. And if $\text{tr}\,\Lambda(\rho) \leq \text{tr}\,\rho = \text{tr}\,\rho^T$ for all ρ then $0 \leq \text{tr}(\rho^T - \rho^T\,\text{tr}_{\text{out}}\,\tilde{J}) = \text{tr}[\rho^T(\mathbb{1}_{\text{in}} - \text{tr}_{\text{out}}\,\tilde{J})]$ for all ρ^T, and it follows that $\mathbb{1}_{\text{in}} - \text{tr}_{\text{out}}\,\tilde{J} \geq 0$.

Hence we have shown that there is a one-to-one correspondence between bipartite states on $\mathcal{H}_{\text{out}} \otimes \mathcal{H}_{\text{in}}$ that satisfy $\text{tr}_{\text{out}}\,\tilde{J} = \mathbb{1}_{\text{in}}$ and CPTP maps $\Lambda : \mathcal{B}(\mathcal{H}_{\text{in}}) \to \mathcal{B}(\mathcal{H}_{\text{out}})$.

2.2.3 Recap: Kraus representation

Using what we learned in the previous section we can derive another widely used characterization of the the CPTP maps known as the Kraus representation. Since the Choi process matrix J is a positive semidefinite and be written it in terms of its eigenvectors as

$$J = \sum_{i=1}^{d_{\text{in}}} |K_i\rangle\langle K_i| = \sum_{i=1}^{d_{\text{in}}} (K_i \otimes \mathbb{1})\,|\mathbb{1}\rangle\langle\mathbb{1}|\,(K_i^\dagger \otimes \mathbb{1}) \tag{2.2.125}$$

where note that the positive eigenvalues of been absorbed in the projectors and that there might be some zero eigenvalues (rank of J smaller than d_{in}) effectively reducing the number of required Kraus operators. The case of unitary evolution corresponds to $\text{rank}(J) = 1$.

[10] The necessity of the condition follows from the definition of complete positivity. For sufficiency note that we can write any pure state in an extended Hilbert space $|C\rangle \in \mathcal{H} \otimes \mathbb{C}^n$ for arbitray n as $|C\rangle = \mathbb{1}_d \otimes C^T\,|\mathbb{1}\rangle$ from where $(\Lambda \otimes \mathbb{1}_d)(|C\rangle\langle C|) = \mathbb{1} \otimes C^T(\Lambda \otimes \mathbb{1})(|\mathbb{1}\rangle\langle\mathbb{1}|)\mathbb{1} \otimes C^* \geq 0$, which follows by hypothesis, $(\Lambda \otimes \mathbb{1})(|\mathbb{1}\rangle\langle\mathbb{1}|) \geq 0$, together with the fact that $B \geq 0$ implies $ABA^\dagger \geq 0$. The extension to general mixed state input ρ follows by convexity (positivity of ρ is implied by that of each element of a pure-state ensemble decomposition) or by using a purification using an enlarged ancillary system.

These operators $\{K\}_{i=1}^{N}$ are the so-called *Kraus Operators* and using the relation (2.2.121) we can arrive at the *Kraus Decomposition* representation of the channel:

$$\Lambda(\rho) = \sum_{i=1}^{d_{\text{in}}} K_i \rho K_i^\dagger \tag{2.2.126}$$

In addition, a map will be trace preserving if and only if: $\sum_{i=1}^{d_{\text{in}}} K_i^\dagger K_i = \mathbb{1}_{\text{in}}$.

$$\text{tr}(\Lambda(\rho)) = \text{tr}\left(\sum_{i=1}^{d_{\text{in}}} K_i \rho K_i^\dagger\right) = \text{tr}\left(\underbrace{\sum_{i=1}^{d_{\text{in}}} K_i^\dagger K_i}_{\mathbb{1}} \rho\right) = \text{tr}(\rho) \tag{2.2.127}$$

And similarly, the map is trace decreasing iff $\sum_{i=0}^{d_{\text{in}}} K_i^\dagger K_i \leq \mathbb{1}_{\text{in}}$. We also note that each term in the sum (2.2.127), $K_i \rho K_i^\dagger$ is completely positive[11] and therefore $\Lambda(\rho)$ is completely positive.

This also follows from the fact that the partial trace formulas (2.2.112) and (2.2.113) and the trace-preserving condition (2.2.123)

$$\mathbb{1}_{\text{in}} = \text{tr}_{\text{out}}(J) = \text{tr}_{\text{out}}\left(\sum_{i=1}^{d_{\text{in}}} |K_i\rangle\langle K_i|\right) = \sum_{i=1}^{d_{\text{in}}} (K_i^\dagger K_i)^* \quad \Longrightarrow \quad \sum_{i=0}^{d_{\text{in}}} K_i^\dagger K_i = \mathbb{1}_{\text{in}} \tag{2.2.128}$$

Kraus representation is not unique: Note that by construction, the spanning set $\{|K_i\rangle\}$ gives the minimal number of Kraus operators. However from GHJW theorem (2.2.118) it follows that

$$J = \sum_{i=1}^{d_{\text{in}}} |K_i\rangle\langle K_i| = \sum_{k=1}^{m} |M_k\rangle\langle M_k| \quad \Leftrightarrow \quad |M_k\rangle = \sum_i V_{ki} |K_i\rangle \text{ for some isometry } V.$$

which implies that two Kraus representations are equivalent (i.e. lead to the same quantum channel) iff $M_k = \sum_i V_{ki} K_i$ for some isometry V.

2.2.4 Teleporting through a channel, aka tele-stretching

Recap: Teleportation

We start by shortly presenting the teleportation protocol. Alice posses a state $|\psi\rangle_A \in \mathcal{H}_A = \text{span}\{|i\rangle\}_{i=1}^{d_A}$ that she wishes to teleport to Bob who is in a remote location. The state might be completely unknown to Alice. In addition we note that what we present here can be trivially extended (by linearity of the protocol) to the case where Alice's system is mixed or even entangled with a second system (as e.g. in entanglement swapping). Aside from this system Alice shares a maximally entangled state with Bob,

$$|\Phi^+\rangle_{AB} = \frac{1}{\sqrt{d_A}} \sum_{i=1}^{d_A} |ii\rangle_{AB} = \frac{1}{\sqrt{d_A}} |\mathbb{1}\rangle_{AB} \tag{2.2.129}$$

Alice performs a collective measurement on her systems AA' projecting on maximally entangled basis

$$\text{for } i \in [0, d_A^2 - 1] \quad |\Phi_i\rangle := \frac{1}{d_A} |U_i\rangle = \frac{1}{d_A} U_i \otimes \mathbb{1}_{A'} |\mathbb{1}\rangle_{AA'} = \frac{1}{d_A} \mathbb{1} \otimes U_i^T |\mathbb{1}\rangle_{AA'} \tag{2.2.130}$$

with $\langle U_i | U_j \rangle = \text{tr}(U_i^\dagger U_j) = d_A \delta_{ij}$. For qubit systems one could take for example the paradigmatic Bell states corresponding to $U_i \in \{\mathbb{1}, \sigma_x, -i\sigma_y, \sigma_z\}$. For convenience and wlog we take $U_0 = \mathbb{1}$. We refer to such collective measurement as (generalized) Bell measurement.

Given the input state $|\Psi\rangle_{AA'B} = |\psi\rangle_A |\Phi^+\rangle_{A'B} = \sum_k c_k |k\rangle_A |\Phi^+\rangle_{A'B}$ it is immediate to check if the Bell measurement returns outcome $i = 0$, the conditional (unnormalized) state is:

[11] ${}_{SA}\langle\Psi| K_i \otimes \mathbb{1} \rho_{SA} K_i^\dagger \otimes \mathbb{1} |\Psi\rangle_{SA} = \langle\Psi'| \rho_{SA} |\Psi'\rangle_{SA} \geq 0$ for all $\rho_{SA} \geq 0$ and for all $|\Psi\rangle_{SA}$

$$
\begin{aligned}
P_0^{AA'} \otimes \mathbb{1} \left|\psi\right\rangle_A \left|\Phi^+\right\rangle_{A'B} &= \frac{1}{d_A} \left|\Phi_0\right\rangle_{AA'} \sum_{i=1}^{d_A} {}_{AA'}\langle ii| \left(\sum_j c_j \left|j\right\rangle_A \sum_k \left|kk\right\rangle_{A'B}\right) \\
&= \frac{1}{d_A} \left|\Phi_0\right\rangle_{AA'} \sum_{i=1}^{d_A} \left(\sum_i c_i \left|i\right\rangle_B\right) = \frac{1}{d_A^2} \left|\Phi_0\right\rangle_{AA'} \left|\psi\right\rangle_B \qquad (2.2.131)
\end{aligned}
$$

That is, the systems of Alice collapse to the Bell state and Bob system is found in the same state as Alice's input state, $\left|\psi\right\rangle_B$.

Using this result we can easily compute the (unnormalized) conditional state of Bob's system for the other outcomes i of the Bell measurement

$$
\begin{aligned}
\left|\tilde{\phi}_i\right\rangle_B :=& {}_{AA'}\langle\Phi_i| \left|\Psi\right\rangle_{AA'B} &= \frac{1}{d_A} {}_{AA'}\langle\mathbb{1}| \mathbb{1}_A \otimes U_i^* \otimes \mathbb{1}_B \left|\psi\right\rangle_A \left|\mathbb{1}\right\rangle_{A'B} \\
&= \frac{1}{d_A} {}_{AA'}\langle\mathbb{1}| \mathbb{1}_A \otimes \mathbb{1}_{A'} \otimes U_i^\dagger \left|\psi\right\rangle_A \left|\mathbb{1}\right\rangle_{A'B} \\
&= \frac{1}{d_A} U_i^\dagger \left|\psi\right\rangle_B \qquad (2.2.132)
\end{aligned}
$$

where in the first equality we used the last equality of (2.2.130), in the second equality $(V\otimes\mathbb{1})\left|\mathbb{1}\right\rangle = \left|V\right\rangle = (\mathbb{1}\otimes V^T)\left|\mathbb{1}\right\rangle$ and in the third equality we have used (2.2.131), or, more succinctly, ${}_{AA'}\langle\mathbb{1}| \left|\psi\right\rangle_A \left|\mathbb{1}\right\rangle_{A'B} = \left|\psi\right\rangle_B$. The norm of the conditional state gives us the probability of the outcome: $p_i = 1/d_A^2$ for all the d_A^2 possible outcomes, which crucially does not depend on the input state $\left|\psi\right\rangle$ (otherwise one could learn information about the state without disturbing it!).

Unlike the case $i = 0$ we note that conditional state is not Alice's state, however one can simply recover it by performing a unitary transformation: $U_i \left|\phi_i\right\rangle_B = U_i U_i^\dagger \left|\psi\right\rangle_B = \left|\psi\right\rangle_B$. Therefore, Alice can teleport her state to Bob by:

0. Prerequisite: Share a maximally entangled state with Bob.
1. Alice performs Bell measurement on input+ her share of Bell state obtaining outcome i
2. Alice uses a classical channel to send the outcome obtained to Bob ($2\log_2 d_A$ bits of information)
3. Upon receiving the message from Alice, Bob performs the "correcting" unitary operation $U_i^\dagger$

Observations:

- If Bob does not perform the "correcting" unitary operation his state will be equal to the marginal of the initial Bell state, i.e. the maximally mixed state $\rho_B = \frac{1}{d_A}\mathbb{1}$. That is the teleportation channel acts as completely depolarizing channel:

$$\rho_B = \mathcal{T}(\rho) = \frac{1}{d_A^2} \sum_i U_i \rho U_i^\dagger = \frac{1}{d_B}\mathbb{1}$$

- For the very same reason, successful teleportations does not happen instantaneously, as Alice's message can travel at most at the speed of light.
- The teleportation protocol can be successfully implemented if Alice perform a collective POVM of the form $\{|\alpha_i U_i\rangle\langle\alpha_i U_i|\}$ fulfilling the completeness relation $\sum_i |\alpha_i U_i\rangle\langle\alpha_i U_i| = \mathbb{1}$ (or a continuous version of it). This guarantees that each POVM element is proportional to a projector on a maximally entangled state, and that the outcome probability is independent of the input state $\left|\psi\right\rangle_A$.
- Similarly the teleportation can be successfully implemented by using an arbitrary maximally entangled shared $\left|\Phi\right\rangle = \mathbb{1}\otimes V\left|\mathbb{1}\right\rangle_{A'B} = \left|V^T\right\rangle_{A'B}$ resource between Alice and Bob. Bob just need to adapt the "correcting" operation: $V^\dagger U_i^\dagger$ (see (2.2.133) below)

- In order to send $n = \log_2 d_A$ qubits the protocol requires $2n$ bits of classical communication + n e-bits (2 qubit Bell states). This can be contrasted to the n-qubits + n-ebits required to send $2n$ bits of classical communication via dense-coding.

- Quite importantly, the quantum resource shared between Alice and Bob can be established way before the input is provided. For instance, if Alice and Bob can only have access to a quantum channel during a limited amount of time, they can use it to share entangled states, which can later be used to send quantum information in the absence of the quantum channel. Moreover, if the channel between Alice and Bob is noisy, they can use entanglement distillation protocols to obtain a a high-fidelity maximally entangled stated, which can then be used to reliably transmit sensible QI.

- Teleportation acts as an effective ideal channel between Alice and Bob, i.e. it can be used for example by Alice to send her share of an entangled state to Bob. This is called *entanglement swapping*, and together with entanglement distillation constitutes the essence of *quantum repeaters*.

Teleporting through a channel

Here we discuss a straightforward implication of the teleportation protocols that it is widely used in the study of quantum channels and their applications.

Consider that Bob applies a channel $\Lambda(\rho) = \sum_k K_k \rho K_k^\dagger$ on Bob's share of the maximally entangled state, the initial shared resource will be

$$\sigma_{A'B} = \frac{1}{d_A} \sum_k \mathbb{1} \otimes K_i \left|\mathbb{1}\right\rangle\!\left\langle\mathbb{1}\right| \mathbb{1} \otimes K_i^\dagger = \sum_k \left|K_i^T\right\rangle\!\left\langle K_i^T\right|$$

By linearity, one readily finds that after Alice's Bell measurement outcomes $|\Phi_i\rangle$ the state of Bob will be found in

$$\begin{aligned}
\rho_B &= \sum_k ({}_{AA'}\langle \mathbb{1}| \, \mathbb{1}_A \otimes U_i^T \otimes \mathbb{1}_B \left|\psi\right\rangle_A \left|K_k^T\right\rangle_{A'B}) \langle \text{idem}| \\
&= \sum_k ({}_{AA'}\langle \mathbb{1}| \, \mathbb{1}_A \otimes \mathbb{1}_{A'} \otimes \mathbb{1}_B \left|\psi\right\rangle_A \left|U_i^T K_k^T\right\rangle_{A'B}) \langle \text{idem}| \\
&= \sum_k ({}_{AA'}\langle \mathbb{1}| \, \mathbb{1}_A \otimes \mathbb{1}_{A'} \otimes K_k U_i \left|\psi\right\rangle_A \left|\mathbb{1}\right\rangle_{A'B}) \langle \text{idem}| \\
&= \sum_k (K_k U_i \left|\psi\right\rangle_B) \langle \text{idem}| = \Lambda(U_i \left|\psi\right\rangle\!\left\langle\psi\right| U_i^\dagger)
\end{aligned} \tag{2.2.133}$$

where in order to reduce cluttering we have introduced the notation $|\psi\rangle\langle\text{idem}| := |\psi\rangle\langle\psi|$. Notice that Bob can recover the action of the channel of the input state of Alice in two cases: i) with probability $p_0 = \frac{1}{d_A^2}$ if Alice's Bell measurement returns $i = 0$, i.e. projects on $\left|\mathbb{1}\right\rangle_{AA'}$. In that case no further action is required. ii) If the unitaries U_i can be "commuted out" of the action of the channel: $\Lambda(U_i \left|\psi\right\rangle\!\left\langle\psi\right| U_i^\dagger) = V_i \Lambda(\left|\psi\right\rangle\!\left\langle\psi\right|) V_i^\dagger$, then upon receiving message i from Alice Bob need to apply the "correcting" unitary $V_i^\dagger$. Channels for which $\Lambda(U_i \left|\psi\right\rangle\!\left\langle\psi\right| U_i^\dagger) = V_i \Lambda(\left|\psi\right\rangle\!\left\langle\psi\right|) V_i^\dagger$ for some unitaries U_i and V_i, with $\sum_i \left|U_i\right\rangle\!\left\langle U_i\right| = \mathbb{1}_{\text{in}}$ are called *teleportation covariant channels*.

So, the mathematical correspondence between channels and Choi-states also holds in a restricted sense in the physical world:

1. ($\Rightarrow$) Given a channel $\Lambda(\cdot)$ one can physically produce the Choi-state:

$$\rho'_{\text{out in}} = \Lambda \otimes \mathbb{1} \left|\Phi^+\right\rangle\!\left\langle\Phi^+\right|$$

which is equivalent

2. (Restricted $\Leftarrow$) Given a Choi-state $\rho'_{\text{out in}}$, fulfilling $\rho'_{\text{out in}} \geq 0$ and $\text{tr}_{\text{out}}\, \rho'_{\text{outin}} = \frac{1}{d_{\text{in}}} \mathbb{1}_{\text{in}}$, one can use the state $\sigma_{A'B} = \rho'_{\text{in out}}$ (flipped order of in and out labels) to probabilistically recover the action of the state. If the channel is a *teleportation covariant channel*, than one can deterministically recover the state, i.e.

$$\Lambda(\rho_{\text{in}}) = \mathcal{T}(\rho_{\text{in}}, \sigma_{A'B}). \tag{2.2.134}$$

where $\mathcal{T}(\rho, \sigma_{A'B})$ denotes the teleportation protocol using $\sigma_{A'B}$ as a resource.

This will be exploited later to study the concatenation of quantum operations. The take home message is that if one has the Choi state, $\sigma_{A'B}$, of a channel $\Lambda : A \mapsto B$, if one projects $\rho_A \otimes \sigma_{A'B}$ on the (unnormalized) Bell state $|\mathbb{1}\rangle_{A'A} \langle \mathbb{1}|$ on gets:

$$\text{tr}_{AA'}[\rho_A \otimes \sigma_{A'B}(P_{AA'} \otimes \mathbb{1}_B)] = \Lambda(\rho) \in \mathcal{B}(\mathcal{H}_B) \tag{2.2.135}$$

which is equivalent to (2.2.121) which we reproduce here:

$$d_A \,\text{tr}_A \left[(\mathbb{1}_B \otimes \rho_A^T)\sigma_{BA}\right] = \text{tr}_A \left[(\mathbb{1}_B \otimes \rho_A^T) J_\Lambda\right] = \Lambda(\rho) \in \mathcal{B}(\mathcal{H}_B) \tag{2.2.136}$$

where the effect of the Bell projector and tracing out A' effectively induces the partial trace in (2.2.136).

Exercise 2.2.1. *Show that if a channel Λ fulfills the covariance property $\Lambda(U\rho U) = V\Lambda(\rho)V^\dagger$ then the corresponding Choi operator J_Λ must have the symmetry:*

$$V \otimes U^T J_\Lambda V^\dagger \otimes U^* = J_\Lambda \tag{2.2.137}$$

Hint: use (2.2.111).

Exercise 2.2.2. *Show that the depolarizing channel $\Lambda(\rho) = p\rho + (1-p)\frac{1}{2}\mathbb{1}$ is teleportation covariant. Extend the proof to general Pauli channels:*

$$\Lambda_{\text{Pauli}}(\rho) = \sum_{i=0}^{4} p_i \sigma_i \rho \sigma_i \tag{2.2.138}$$

where $\{\sigma_0, \sigma_1, \sigma_2, \sigma_3\} = \{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$, and $p_i \geq 0$ are normalized $\sum_i p_i = 1$.

2.2.5 Ultimate Quantum Limit

In this section study the optimal discrimination of quantum channels (in the sense of the minimum discrimination error). However the results easily extend to other bayesian inference problems, so long as a linear cost function is used.

Ultimate limit for teleportation-covariant channels

In the previous section we have shown that for teleportation covariant channels there is a physical correspondence between channels and bipartite quantum states. That is, given a physical channel Λ one can create a bipartite quantum state σ_Λ (Choi state) and from that state one can use teleportation to recover the action of the channel. For teleportation covariant channels states and channels are interchangeable resources. This means that if we are faced with a task involving a teleportation covariant family of channels [12] $\{\Lambda_\theta\}_\theta$: $\Lambda_\theta(U_i |\psi\rangle\langle\psi| U_i^\dagger) = V_i \Lambda_\theta(|\psi\rangle\langle\psi|) V_i^\dagger$, then this task can be implemented with exactly the same performance by using the corresponding family of Choi states $\{\sigma_\theta\}_\theta$.

For instance, if the task is to discriminate between a family of channels when n uses of the channel (black boxes) are available than the optimal probability of error is the same than that of discriminating

[12] Note that commuted correcting unitaries, V_i, have to be the same for all channels in the family.

n copies of the states $\sigma_\lambda^{\otimes n}$. In particular one can make use of the know results is quantum state discrimination, e.g. the Helstrom bound

$$\text{For two channels } \Lambda_{1,2} \text{ with priors } \eta_{1,2} \quad P_{\text{err}} = \frac{1}{2}(1 - ||\eta_1 \sigma_{\Lambda_1}^{\otimes n} - \eta_2 \sigma_{\Lambda_2}^{\otimes n}||_1) \tag{2.2.139}$$

or the asymptotic error rate, provided by the quantum Chernoff bound

$$Q_{\text{Ch}}(\Lambda_1, \Lambda_2) = -\lim_{n\to\infty} \frac{1}{n} \log P_{\text{err}} = \min_s \operatorname{tr} \sigma_{\Lambda_1}^s \sigma_{\Lambda_2}^{1-s} \tag{2.2.140}$$

Note that these are the optimal bounds under the most general strategies that can be implemented with n uses of the channels, including sequential and adaptive uses of them.

Ultimate limit as an SDP

Here we tackle the problem of channel discrimination for general families of channels, not necessarily teleportation covariant ones. Given a set of N channels $\{\Lambda_k\}$ with Choi process matrices $\{J_k\}$ (for simplicity we take equal priors), an input state ρ and a POVM at the output $\{E_k\}$ the probability of success of identifying correctly one instance of the channel can be written as:

$$P_s = \frac{1}{N}\sum_{k=1}^N \operatorname{tr}(E_k \Lambda_k^E(\rho)) = \frac{1}{N}\sum_{k=1}^N \operatorname{tr}\left((E_k \otimes \rho^T) J_k^E\right) = \frac{1}{N}\sum_{k=1}^N \operatorname{tr}\left(F_k J_k^E\right) \quad \text{with } F_k = E_k \otimes \rho^T \tag{2.2.141}$$

In this expression we have, crucially, included the possibility that an entangled probe is send through the channel as in figure 2.2.1, i.e. $\rho \in \mathcal{B}(\mathcal{H}_{\text{in}_1} \otimes \mathcal{H}_{\text{in}_2})$, $E_k \in \mathcal{B}(\mathcal{H}_{\text{out}_1} \otimes \mathcal{H}_{\text{out}_2})$ and $J_k^E = J_k \otimes |\mathbb{1}\rangle_{\text{out}_2\,\text{in}_2}\langle\mathbb{1}|$ is the Choi state of the extended channel $\Lambda_k^E = \Lambda_k \otimes \mathbb{1}$. In order to find the optimal success probability one would need to optimize over all possible input states (that can be taken to be pure) and all possible POVMs. This is a bilinear optimization problem that in principle is hard to solve.

In the same way that the trace-norm provides an operational distance measure in the context of state discrimination (2.1.68)

$$P_{\text{err}} = \frac{1}{2}(1 - \frac{1}{2}||\rho - \sigma||_1) \tag{2.2.142}$$

one can similarly define the so-called DIAMOND-NORM for channels

$$\|\Lambda_0 - \Lambda_1\|_\diamond := \sup_\rho ||\Lambda_0 \otimes \mathbb{1}(\rho) - \Lambda_1 \otimes \mathbb{1}(\rho)||_1 \geq \frac{1}{d}\|J_0 - J_1\|_1 \tag{2.2.143}$$

So that the error probability of discriminating between two channels is

$$P_{\text{err}} = \frac{1}{2}(1 - \frac{1}{2}\|\Lambda_0 - \Lambda_1\|_\diamond) \tag{2.2.144}$$

A very common way to derive bounds on optimization problems is to introduce relaxations, i.e. lift constrains or equivalently expand the range of the optimization variables. In our case, to provide this relaxation let us first note that (2.2.141) can be written as

$$P_s = \frac{1}{N}\sum_{k=1}^N \operatorname{tr}\left(F_k J_k^E\right) = \frac{1}{N}\sum_{k=1}^N \operatorname{tr}\left(\tilde{F}_k J_k\right) \qquad \text{where } \tilde{F}_k = {}_{\text{out}_2\,\text{in}_2}\langle\mathbb{1}|\, F_k \,|\mathbb{1}\rangle_{\text{out}_2\,\text{in}_2} \tag{2.2.145}$$

where $\sum_k F_k = \sum_k E_k \otimes \rho^T = \mathbb{1} \otimes \rho^T$ and hence

$$\frac{1}{N}\sum_k \tilde{F}_k = {}_{\text{out}_2\,\text{in}_2}\langle\mathbb{1}|\, \mathbb{1}_{\text{out}_1} \otimes \mathbb{1}_{\text{out}_2} \otimes \rho^T \,|\mathbb{1}\rangle_{\text{out}_2\,\text{in}_2} = \mathbb{1}_{\text{out}_1} \otimes \operatorname{tr}_{\text{in}_2}(\rho^T) =: \mathbb{1}_{\text{out}_1} \otimes \sigma$$

So, our relaxed optimization problem can be written as

$$P_s^U = \max_{\{\tilde{F}_k\}} \frac{1}{N}\sum_{k=0}^{N} \mathrm{tr}(\tilde{F}_k J_k) \qquad (2.2.146)$$
$$\text{subject to } \tilde{F}_k \geq 0 \text{ and } \frac{1}{N}\sum_k \tilde{F}_k = \mathbb{1}\otimes\sigma \text{ for some state } \sigma \in \mathcal{B}(\mathcal{H}_{\text{in}})$$

This new optimization problem constitutes a standard SDP, i.e., an optimization of a linear function with positive semi-definite constraints. It can be solved with arbitrary precision with the usual SDP tools, but remember, since our optimization variables $\tilde{F}_k$ might have additional constraints, the relaxed optimization problem could return super-optimal results: $P_s \leq P_s^U$.

However, on this occasion we are lucky because, as we show next, for every $\tilde{F}_k$ satisfying $\frac{1}{N}\sum_k \tilde{F}_k = \mathbb{1}\otimes\sigma$ it is possible to construct an input state ρ, and measurement E_k for which the success probability coincides with P_s^U. That is, the upper bound is attainable by a particular strategy, which means that P_s^U in (2.2.151) is the ultimate (attainable) quantum limit.

Indeed, given a set of operators $\{\tilde{F}_k\}$ satisfying the conditions in (2.2.151) for given σ we can construct a pure input state

$$\rho = \left|\sqrt{\sigma}^T\right\rangle\!\left\langle\sqrt{\sigma}^T\right| \quad \text{where} \quad \left|\sqrt{\sigma}^T\right\rangle = (\mathbb{1}\otimes\sqrt{\sigma})\,|\mathbb{1}\rangle \in \mathcal{H}_{\text{in}_1}\otimes\mathcal{H}_{\text{in}_2} \qquad (2.2.147)$$

and a POVM with elements

$$E_k = (\mathbb{1}\otimes\sigma^{-1/2})\tilde{F}_k(\mathbb{1}\otimes\sigma^{-1/2}) \in \mathcal{B}(\mathcal{H}_{\text{out}_1}\otimes\mathcal{H}_{\text{out}_2}) \qquad (2.2.148)$$

which by construction satisfy the completeness relation[13].

Using this ansatz we find that each term in the sum(2.2.141) is

$$\mathrm{tr}\left[(E_k\otimes\rho^T)J_k^E\right] = \mathrm{tr}\left[(E_k\otimes\rho^T)J_k\otimes|\mathbb{1}\rangle\!\langle\mathbb{1}|\right] \qquad (2.2.149)$$

Now we use that $|C\rangle\!\langle C|^T = |C^*\rangle\!\langle C^*|$ and that $\sqrt{\sigma}^* = \sqrt{\sigma}^T$ to write $\rho^T = (\mathbb{1}\otimes\sqrt{\sigma}^T)\,|\mathbb{1}\rangle\!\langle\mathbb{1}|\,(\mathbb{1}\otimes\sqrt{\sigma}^T)$ and show that

$$(\mathbb{1}_{\text{out}_1}\otimes\sigma^{-1/2})\otimes(\mathbb{1}_{\text{in}_1}\otimes\sqrt{\sigma}^T)\,|\mathbb{1}\rangle_{\text{out}_2\,\text{in}_2} = \left|\sigma^{-1/2}\sqrt{\sigma}\right\rangle_{\text{out}_2\,\text{in}_2} = |\mathbb{1}\rangle_{\text{out}_2\,\text{in}_2}$$

and analogously for its adjoint. Using these relations we can substitute the definitions of (2.2.147) and (2.2.148) in (2.2.149) to find,

$$\begin{aligned}
&\mathrm{tr}[(E_k\otimes\rho^T)J_k^E] = \mathrm{tr}_{\text{out}_1\,\text{out}_2}\,\mathrm{tr}_{\text{in}_1\,\text{in}_2}[(\tilde{F}_k\otimes|\mathbb{1}\rangle\!\langle\mathbb{1}|)(J_k\otimes|\mathbb{1}\rangle\!\langle\mathbb{1}|)] = \\
&= \mathrm{tr}_{\text{out}_1\,\text{out}_2}\sum_{ij}\tilde{F}_{k\;\text{in}_1\,\text{in}_2}\langle ii|\,(J_k\otimes|\mathbb{1}\rangle\!\langle\mathbb{1}|)\,|jj\rangle_{\text{in}_1\,\text{in}_2} = \sum_{ij}\mathrm{tr}_{\text{out}_1\,\text{out}_2}[\tilde{F}_k(J_k)_{ij}\,|i\rangle\!\langle j|] = \\
&= \mathrm{tr}(\tilde{F}_k J_k)
\end{aligned}$$

This concludes the proof, since this coincides with the corresponding term in (2.2.146).

To wrap up, we have seen that the optimization over (entangled) input states ρ_{SA} and final quantum measurement E_k can be substituted by the optimization of a single object $F = \{F_k\}$, a so-called **quantum tester**:

$$P_s = \max_{\{F_k\}} \frac{1}{N}\sum_{k=0}^{N} \mathrm{tr}(F_k J_k) \qquad (2.2.150)$$
$$\text{subject to } F_k \geq 0 \text{ and } \frac{1}{N}\sum_k F_k = \mathbb{1}\otimes\sigma \text{ for some state } \sigma \in \mathcal{B}(\mathcal{H}_{\text{in}})$$

This is illustrated in figure 2.2.2.

[13] Here we have assumed that σ has full support. Otherwise, one can complement each POVM by summing an additional

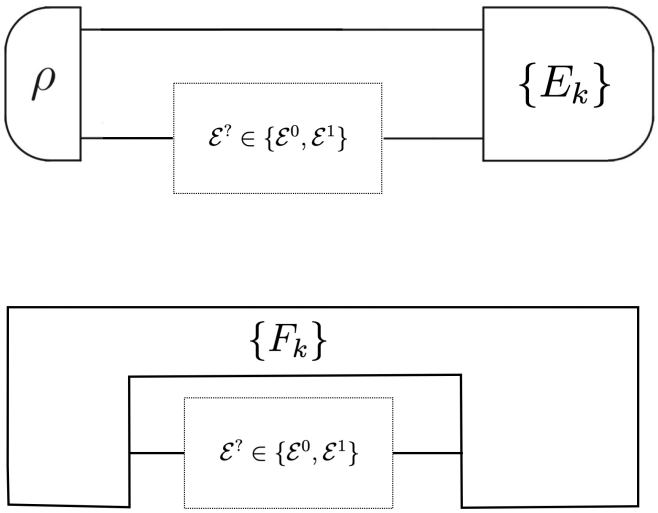


Figure 2.2.2: Most general quantum strategy test a channel $\mathcal{E}_?$. Top: physical implementation as input state preparation and measurement of output. Bottom: mathematical characterization as a quantum tester $\{F_k\}$.

Dual form

Here we derive the dual form of the above SDP (primal). Although we will do it for a particular instance, the techniques can be easily extended to all SDP's (as shown previously in section 1.7)

$$P_s = \max_{\{F_k\}} \sum_{k=0}^{N} \mathrm{tr}(F_k \pi_k J_k) \qquad (2.2.151)$$
$$\text{subject to } F_k \geq 0 \text{ and } \sum_k F_k = \mathbb{1} \otimes \sigma \text{ for some state } \sigma \in \mathcal{B}(\mathcal{H}_{\text{in}})$$

Since we are maximizing linear function over convex set there exists efficient numerical methods to navigate through the set seeking for better and better solutions. Moreover, any particular choice of feasible primal variables, F_k's and σ fulfilling the above conditions, will provide a lower-bound on $P_s \geq P_s(\{F_k\})$.

One can define a Lagrangian by adding a "penalty" term for each condition on the primal variables.

$$L(F_k, X, Y, Y_k) := \sum_{k=1}^{N} \mathrm{tr}(F_k \pi_k J_k) + \mathrm{tr}[Y_k F_k] - \mathrm{tr}[X(\sum_k F_k - \mathbb{1} \otimes \sigma)] + \mathrm{tr}[Y\sigma] - \lambda(\mathrm{tr}\,\sigma - 1) \qquad (2.2.152)$$

where the last two terms correspond to the constrains $\sigma \geq 0$ and $\mathrm{tr}\,\sigma = 1$ on the dummy variable σ, and we have introduced the dual variables X and $Y, Y_k \geq 0$ so that $P_s(\{F_k\}) \leq L$ for any feasible point $\hat{F}_k$. If we define the *Lagrange dual function* as the maximization of the Lagrangian over the unconstrained primal variables we obtain the inequalities:

$$\ell(X, Y, Y_k) := \sup_{\{F_k\}} L(F_k, X, Y, Y_k) \geq L(\hat{F}_k, X, Y, Y_k) \geq P_s(\{\hat{F}_k\}) \qquad (2.2.153)$$

Since these inequalities holds for any feasible point, it also holds for the optimal value $P_s = \max P_s(\{\hat{F}_k\})$ in (2.2.151):

$$P_s \leq \ell(X, Y, Y_k) \qquad (2.2.154)$$

This bound holds for any value of the dual variables (X and $Y, Y_k \geq 0$). However, note that it is non-trivial only in the domain of $\ell(X, Y, Y_k)$ where $\ell(X, Y, Y_k) < \infty$. To find such domain $(X, Y, Y_k) \in \mathrm{dom}(\ell)$,

and positive term A_k, such that $\sum_k E_k = \mathbb{1}$. Also in that case, $\sigma^{-1/2}$ needs to be understood as a pseudo-inverse on the relevant subspace.

i.e. the *dual feasible* variables, we can gather the terms in (2.2.152)

$$L = \sum_{k=1}^{N} \mathrm{tr}[F_k(\pi_k J_k - X + Y_k] + \mathrm{tr}[\sigma(\mathrm{tr}_{\mathrm{out}}\, X + Y - \lambda)] + \lambda \tag{2.2.155}$$

Now it is clear that maximization of L over the unconstrained primal variables becomes unbounded (∞) unless $\pi_k J_k - X + Y_k = 0$ and $\mathrm{tr}_{\mathrm{out}}\, X - Y - \lambda \mathbb{1} = 0$. These two conditions can be written as $X = \pi_k J_k + Y_k \geq \pi_k J_k$ and $\mathrm{tr}_{\mathrm{out}}\, X = \lambda \mathbb{1} - Y \leq \lambda \mathbb{1}$. Therefore we can obtain a non-trivial upper-bound for all feasible dual variables fulfilling these conditions: $P_s \leq \ell(X, Y, Y_k)$. In particular we can write the tightest bound as the following SDP:

$$\begin{aligned} P_S \leq P_S^{\mathrm{dual}} \quad &:= \quad \min \ell(X, Y, Y_k) = \min \lambda \\ &\qquad \text{s.t.} \quad X \geq \pi_k J_k \text{ and } \mathrm{tr}_{\mathrm{out}}\, X \leq \lambda \mathbb{1} \end{aligned} \tag{2.2.156}$$

which is called the dual problem corresponding to the primal problem in (2.2.151).

For any feasible points of the primal and dual variables the following inequalities hold

$$P_s(\{F_k\}) := \sum \mathrm{tr}(\pi_k J_k F_k) \leq P_s \leq \ell(X, Y, Y_k) = \lambda \tag{2.2.157}$$

In particular taking maximum of the lhs and minimum of the rhs we obtain the so-called *weak duality* $P_s \leq P_s^{\mathrm{dual}}$. Problems with the *strong duality* property are those for which the *duality gap* closes, i.e. $P_s = P_s^{\mathrm{dual}}$. We say that a SDP is *strictly feasible* if it satisfies its positive semidefiniteness requirement strictly: i.e. with positive definiteness, e.g. for our primal SDP $\exists F_k > 0\ \forall k$. If one finds some strictly feasible point, then Slater's theorem states that strong duality holds, i.e. primal and dual problem return the same optimal value.

Returning to our particular SDP problem of quantum hypothesis testing, we can write (2.2.157), defining $X =: \lambda C$

$$\begin{aligned} P_s \leq P_s^{\mathrm{dual}} \quad &:= \quad \min_{\lambda, C} \lambda \\ &\qquad \text{s.t.} \quad \lambda C \geq \pi_k J_k \text{ and } \mathrm{tr}_{\mathrm{out}}\, C \leq \mathbb{1} \end{aligned} \tag{2.2.158}$$

Since it is straightforward to give a strictly feasible point, e.g. $C = \mathbb{1} \otimes \mathbb{1}$ and $\lambda = 1$, by Slaters theorem we know that $P_s = P_s^{\mathrm{dual}}$. In addition, for any C that satisfies both inequalities we can construct a new variable $\tilde{C} = C + \rho \otimes \delta$, where $\delta := \mathbb{1} - \mathrm{tr}_{\mathrm{out}}\, C \geq 0$ and ρ is an arbitrary state, that satisfies $\lambda \tilde{C} \geq \lambda C \geq \pi_k J_k$ and $\mathrm{tr}_{\mathrm{out}}\, C = \mathbb{1}$ without affecting the value of λ, i.e. the objective value of the SDP.

Therefore the dual SDP formulation for the channel discrimination problem can be written as

$$\begin{aligned} P_s = P_s^{\mathrm{dual}} \quad &:= \quad \min_{\lambda, C} \lambda \\ &\qquad \text{s.t.} \quad \lambda C \geq \pi_k J_k \text{ and } \mathrm{tr}_{\mathrm{out}}\, C = \mathbb{1} \end{aligned} \tag{2.2.159}$$

or equivalently

$$\begin{aligned} P_s = P_s^{\mathrm{dual}} \quad &:= \quad \min_{\lambda, C} \lambda \\ &\qquad \text{s.t.} \quad \exists \text{ a quantum channel } C \text{ fulfilling } \lambda C \geq \pi_k J_k\ \forall k \end{aligned} \tag{2.2.160}$$

To close this section we note that in this (dual) form the optimal probability of success is closely related to the so called one-shot entropies — and their ϵ-smoothed versions for asymmetric scenarios as, for instance, (2.1.70). We give here the definitions, which already hint at the formal similarities. The precise relation and its consequences are discussed in [22].

Definition 2.2.1 *Let A and B be two positive operators on $(\mathcal{H})$, the* MAX ENTROPY *of A relative to B is given by*

$$D_{\max}(A\|B) := -\log \max\{w \mid wA \leqslant B\} \tag{2.2.161}$$

with the convention $\log 0 := -\infty$.

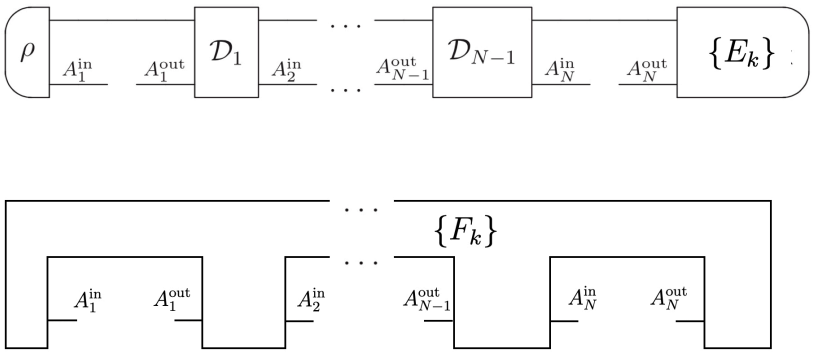


Figure 2.2.3: Most general quantum strategy to test N channel uses. Top: physical implementation consisting of: input state preparation, intermediate quantum operations and measurement of output. The presence of a quantum memory and intermediate coherent controlled actions accounts for all possible feed-forward schemes. Bottom: mathematical characterization as a generalized quantum tester $\{F_k\}$. The input and output hilbert spaces at for each of the used quantum channels (empty slot) is labelled by A_i^{in} and A_i^{out} respectively.

The max relative entropy provides one way to quantify the deviation of A from B. More generally, it is useful to consider the deviation between A and a set of operators :

Definition 2.2.2 *Let A be a positive operator on $\mathcal{H}$ and let $\mathrm{S} \subset \mathrm{Herm}(\mathcal{H})$ be a set of positive operators. The* MAX ENTROPY OF A RELATIVE TO THE SET S*, denoted as $D_{\max}(A\|\mathrm{S})$, is the quantity defined by*

$$D_{\max}(A\|\mathrm{S}) := \inf_{B\in\mathrm{S}} D_{\max}(A\|B)$$

The max relative entropy between a quantum state and a set of quantum states plays a central role also in entanglement theory and other resource theories.

2.2.6 Several uses of the channel: Quantum Combs and general quantum testers

In the previous section the task was to identify a channel based on a single use of it. However, there are many tasks for which it is feasible to make use of the channel several times. If we imagine each channel as an input-output black box it becomes apparent that there are many ways of combining them, e.g. in parallel, in series (sequentially), and even in an adaptive way were classical or quantum info can be fed forward after each use.

Luckily enough it turns out that, as we saw for the single use case, the most general testing strategy making use of N channel uses (see figure 2.2.3) can be formalized as multi-slot or generalized quantum testers: a set linear operators $\{F_k\}$ fulfilling some particular linear and positiveness. In this section we will briefly introduce this objects and at the same time introduce the notion of quantum combs (see [22] for more details).

We have already seen that there's several (equivalent) ways to arrive to the concept of channel: i) Constructive: taking unitary evolutions and making use of ancillary systems that we can trace-out (deterministic channels) or measure and get a conditional evolution (quantum instruments). ii) Formal or axiomatic: asking for the most general transformations that map quantum states into quantum states (trace-preserving completely positive maps: Λ) or those that map quantum states onto pairs $\{(p_i, \rho_i)\}$[14] of probability of outcome and post-measurement quantum state (completely positive stochastic maps $\{\Lambda_i\}$). It is gratifying to see that both approaches, the constructive and the axiomatic, lead to the same

[14] which is equivalent to map from $\rho \mapsto \{\tilde{\rho}_i\}$ where $p_i = \mathrm{tr}\,\tilde{\rho}_i$ and $\rho_i = \frac{\tilde{\rho}_i}{p_i}$

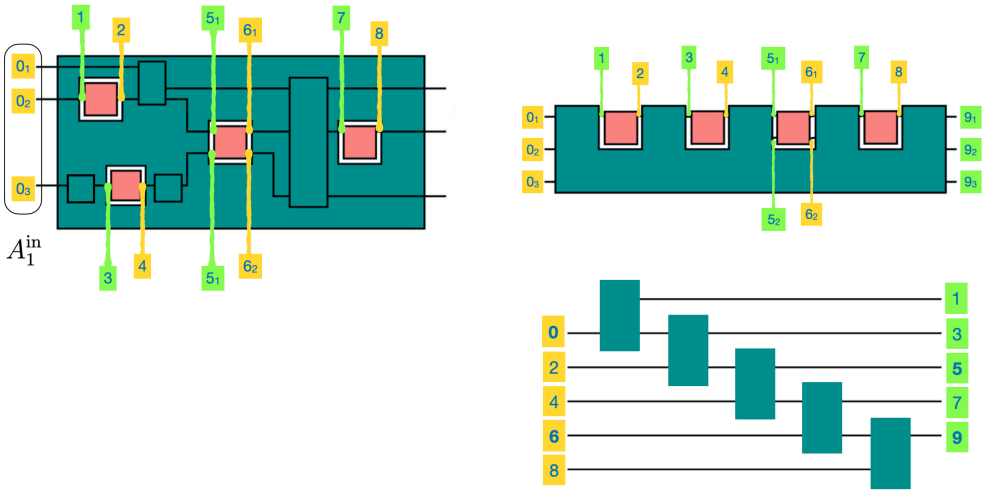


Figure 2.2.4: Quantum circuit board or sequential network (left) can be mapped to a quantum causal comb (right). Bottom picture shows the causal relation between input (yellow: even numbers) and output registers (green: odd numbers). Registered are numbered in the order they appear in the circuit. Outputs can be fed forward to next inputs, with possible intermediate actions (pink boxes) or even other quantum combs (see below). Based on original figs from Chiribella et al. [23].

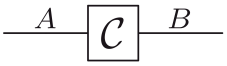
characterization, namely Choi operators $J \geq 0$ on a $B \otimes A$ with $\mathrm{tr}_B\, J = \mathbb{1}_A$, for deterministic channels and $\{J_i \geq 0\}$ fulfilling $\sum_i \mathrm{tr}_B\, J_i = \mathbb{1}_A$ for stochastic quantum channels (quantum instruments).

So, we can ask again the question: what are the most general transformations (super-channels) that map a quantum channels into quantum channels? Or following a constructive approach: what kind of transformations can one construct if we pre-process, post process quantum information going in or out of a channel. Again, the answers to both questions coincides and it turns out that such object can also be characterized by a Choi operators.

We can jump a step further and generalize the notion of super-channels to objects, called QUANTUM COMBS, that take *several* channels as input, and output a channel or actually a comb itself. We will however restricts to causal quantum combs[15] which correspond to a sequential network of quantum channels with internal memories or equivalently an arbitrary circuit board as in figure 2.2.4, i.e., a network of quantum channels with N open slots for the insertion of, yet undefined, sub-circuits (i.e. inputs of the quantum comb). By shuffling and stretching the internal wires, any circuit board can be reshaped in the form of a "comb" with an ordered sequence of slots as in figure 2.2.4. The order of the slots is the causal order induced by the flow of quantum information in the circuit board.

Let us know briefly review how to characterize the building blocks of such a comb together with the rules to compose them (constructive characterization) and finally provide a compact abstract characterization of the full comb.

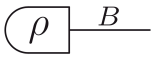
A deterministic (trace preserving) **quantum channel** $\mathcal{C}$ is represented with the diagram



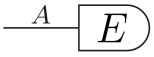
and is characterized by a Choi matrix $J_{\mathcal{C}} \geq 0$ $\mathrm{tr}_B\, J_{\mathcal{C}} = \mathbb{1}_A$. If the channel is non-deterministic (trace decreasing) then the second equality condition is replaced by the inequality $\mathrm{tr}_B\, J_{\mathcal{C}} \leq \mathbb{1}_A$

When system A is trivial ($\dim \mathcal{H}_A = 1$) the quantum operation $\mathcal{C}$ corresponds to the operation of a **state preparation** in B, represented as

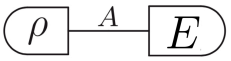
[15] More general, non-causal structures are harder to characterize and do not represent quantum strategies that one can physically implement in a context-free setting.



while when system B is trivial, the quantum operation $\mathcal{C}$ is a **measurement effect** on system A, represented as



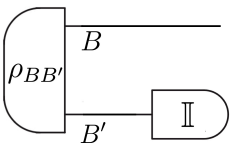
and the Choi operator operator E satisfies $0 \leq E \leq \mathbb{1}$ (POVM element)[16]. The probability of the outcome corresponding to the effect E is given by the **Born's rule:** $\mathrm{tr}(\rho E)$



In the special case where the effect is the identity operator, $E = \mathbb{1}$, the operation corresponds the action of tracing out the system: $\mathrm{tr}\, E\rho = \mathrm{tr}\, \rho$.

A quantum measurement process is described by a **quantum instruments**: a set of quantum trace decreasing operations $\{\mathcal{C}_x\}_{x\in\mathcal{X}}$ with input A and output B, subject to the condition that the sum $\sum_{x\in\mathcal{X}} \mathcal{C}_x$ is trace-preserving. That is, quantum instruments are characterized by a set $\{J_x \geq 0\}_{x\in\mathcal{X}}$ with $\mathrm{tr}_B[\sum_{x\in\mathcal{X}} J_x] = \mathbb{1}_A$, where each label x corresponds to a measurement outcome. $\mathcal{C}_x(\rho) = \tilde{\rho}_x$ is the unnormalized post-measurement state whose trace is the outcome probability. When the output system B is one-dimensional we recover the notion of generalized measurement or **POVM**, where each Choi operator $J_x = E_x \geq 0$ with $\sum_x E_x = \mathbb{1}$.

Now we will study how to combine such operations so as to construct richer structures. For this purpose it is useful to keep track of separate input or output subsystems. For instance, if we want to perform an operation on one party of a bipartite state $\rho_{BB'}$ on $\mathcal{H}_B \otimes \mathcal{H}_{B'}$ we would write



where in this case according to the rules above we have traced-out system B' thus leaving a quantum state $\rho_B = \mathrm{tr}_{B'}\, \rho_{BB'}$ in subsystem B.

Two quantum operations can be connected with each other, as long as (part of) the output of the first operation matches the (part of) input of the second. At the level of Choi operators, the connection is implemented by the operation of LINK PRODUCT.

The link product between two operators, they need to have a hilbert space in common, in the tensor product sense: let C is an operator on $\mathcal{H}_1 \otimes \mathcal{H}_2$ and D is an operator on $\mathcal{H}_2 \otimes \mathcal{H}_3$. Note that it is understood that each operator is extended to match the total hilbert space, so that for instance

$$CD = (C \otimes I_3)\,(I_1 \otimes D) \tag{2.2.162}$$

acts on $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$[17]. With this notation, the LINK PRODUCT of C and D is the operator $C * D$ defined as

$$C * D := \mathrm{Tr}_2\left[CD^{T_2}\right] \text{ on } \mathcal{H}_1 \otimes \mathcal{H}_3$$

[16] As we will shortly see the composition of two combs requires to take the transpose of the linked degrees of freedom, so, linking ρ and E leads to $\mathrm{tr}(\rho E^T)$, nevertheless for a more direct physical interpretation, in the comb diagrams we label measurement effects with the POVM elements E instead corresponding comb E^T

[17] In order to avoid confusion it is a good practice to label explicitly the subspaces where each operator acts, instead of relying on the ordering of the tensor product, e.g. $A_{12}B_{23}$

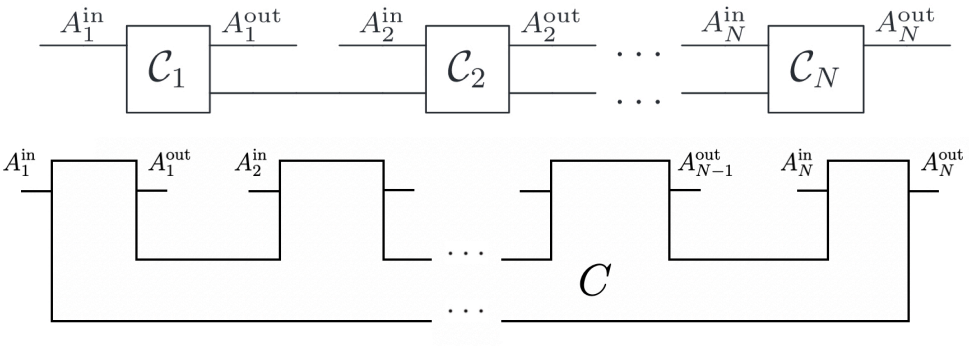


Figure 2.2.5: Quantum comb can always be understood as the concatenation of N channels that leave N open slots (top).Representation of the quantum (causal) comb C with N pairs of input-outputs. C obeys the conditions in (2.2.167) - (2.2.169) (bottom).

where D^{T_2} denotes the partial transpose of D with respect to the Hilbert space $\mathcal{H}_2$, i.e. if

$$D = \sum_{ij;kl} D_{ij;kl} \left|i,j\right\rangle_{23} \left\langle k,l\right| = \sum_{ij;kl} D_{ij;kl} \left|i\right\rangle_2 \left\langle k\right| \otimes \left|j\right\rangle_3 \left\langle l\right| \longrightarrow \tag{2.2.163}$$

$$D^{T_2} = \sum_{ij;kl} D_{ij;kl} (\left|i\right\rangle_2 \left\langle k\right|)^T \otimes \left|j\right\rangle_3 \left\langle l\right| = \sum_{ij;kl} D_{ij;kl} \left|k\right\rangle_2 \left\langle i\right| \otimes \left|j\right\rangle_3 \left\langle l\right| \tag{2.2.164}$$

The link product can be expressed as

$$C * D = \mathrm{tr}_2 \, \mathrm{tr}_{2'} \left[(C_{12} \otimes D_{2'3}) \left(\mathbb{1}_1 \otimes \left|\mathbb{1}_{22'}\right\rangle\!\left\langle \mathbb{1}_{22'}\right| \otimes \mathbb{1}_3 \right) \right]$$

This means that, up to normalization, the link product $C_{12} * D_{23}$ is the (unnormalized) state obtained when a Bell measurement, performed on the states C_{12} and $D_{2'3}$, yields the outcome $\left|\mathbb{1}\right\rangle$. By this physical reasoning or by direct comparison with (2.2.136) or (2.2.121) we arrive at the concatenation rule:

Proposition 2.2.3 (Rule for composing quantum operations) *Let $\mathcal{C}$ be a quantum operation transforming operators on $\mathcal{H}_0$ to operators on $\mathcal{H}_1 \otimes H_2$, let $\mathcal{D}$ be quantum operation transforming operators on $\mathcal{H}_2 \otimes \mathcal{H}_3$ to operators on $\mathcal{H}_4$, the the quantum operation resulting from the composition of $\mathcal{C}$ and $\mathcal{D}$, $\mathcal{F} = \mathcal{D} \circ \mathcal{C}$, is characterized by the* **link product**

$$F = C * D = \mathrm{tr}_2 \left[C D^{T_2} \right] \tag{2.2.165}$$

where $F = J_F$, $C = J_{\mathcal{C}}$ and $D = J_{\mathcal{D}}$, and are the Choi operators of $\mathcal{F}$, $\mathcal{C}$ and $\mathcal{D}$

Note that the notation for the link product does not explicitly show which are connected and traced-out Hilbert spaces. In case of ambiguity is should be stated by other means. Also note that if the two operations act on different hilbert spaces, its composition will be of the form $\mathcal{C} = \mathcal{C}_1 \otimes \mathcal{C}_2$ and the corresponding Choi operator will be $C = C_1 \otimes C_2 = C_1 C_2$, consistent with our convention (2.2.162).

Exercise 2.2.3. *Draw the diagram for teleportation protocol when the Bell measurement outcomes ϕ^+ (no correcting unitary required in that case), in terms of state preparation diagram for input (ρ) and shared entangled state (ϕ^+) and the projector on to the bell state (P_{ϕ^+}) . What is the comb that relates input and output register, i.e. without including preparation of ρ? Write the corresponding Choi operator.*

Equipped with this toolbox can connect, grow, contract combs in any desired way to obtain a new comb. This of course can include the the particular elementary combs corresponding to state preparation, measurements, tracing out etc. A general circuit board or general quantum strategy will be composed of a sequence of quantum operations $\{\mathcal{C}_i\}$ acting on $A_i^{\text{in}} \otimes M^{\text{in}} \otimes A_i^{\text{out}} \otimes M^{\text{out}}$, where M^{in} and M^{in} represent the internal registers of the circuit or internal quantum memory of the process (see figure 2.2.5). Note

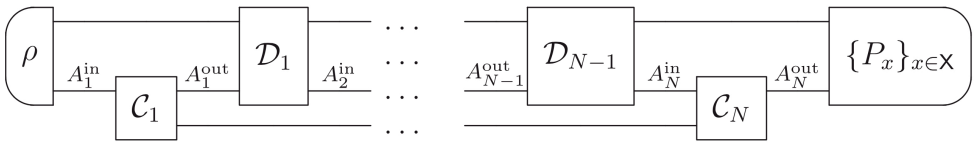


Figure 2.2.6: The quantum tester in figure 2.2.3 is wired to quantum comb in figure 2.2.5 in order to can extract information on the quantum comb.

that even if quantum operation $\mathcal{C}_i$ only acts not trivially on a reduced set of registers S_i we can extend its action over all subsystem involved (internal or external) with the identity channel. At the level of Choi matrices this extension is trivially achieved by tensoring both operations: $J_{C_i} \otimes \bigotimes_{j\notin S_i} |\mathbb{1}\rangle_{\text{in}_j\text{out}_j} \langle\mathbb{1}|$. With this convention the network has Choi operator acting on $\bigotimes_{j=1}^{N} \left(\mathcal{H}_j^{\text{out}} \otimes \mathcal{H}_j^{\text{in}}\right)$ given by

$$C = C_1 * C_2 * C_3 * \cdots * C_N \tag{2.2.166}$$

Now that we have covered how to construct general combs, the following propositions provide its general characterization.

Proposition 2.2.4 *A positive operator C is a quantum comb* if and only if *it satisfies the linear constraints*

$$\text{tr}_{A_N^{out}}\, C^{(N)} = \mathbb{1}_{A_N^{\text{in}}} \otimes C^{(N-1)} \tag{2.2.167}$$

$$\text{tr}_{A_{N-1}^{out}}\, C^{(N-1)} = \mathbb{1}_{A_{N-1}^{\text{in}}} \otimes C^{(N-2)} \tag{2.2.168}$$

$$\vdots$$

$$\text{tr}_{A_1^{out}}\, C^{(1)} = \mathbb{1}_{A_1^{\text{in}}} \tag{2.2.169}$$

where $C^{(n)}$ is a suitable operator on $\mathcal{H}_n := \bigotimes_{j=1}^{n} \left(\mathcal{H}_j^{out} \otimes \mathcal{H}_j^{\text{in}}\right)$

A generic comb of this form is shown in figure 2.2.5. Notice that every $C^{(n)}$ is itself a quantum comb. [18] The whole hierarchy of constrains is a direct consequence of the normalization condition of quantum channels. Physically, the positive operator $C^{(n)}$ represents the subnetwork transforming the first n inputs to the first n outputs. This is a consequence of temporal ordering in input-output pairs in causal combs, which means that there is no backward causation.

We finally consider non-deterministic networks containing measurement devices, which may generate random outcomes. These networks are useful for testing the quantum circuits, general physical processes, or quantum strategies of the form of figure 2.2.5 consisting of multiple time steps. Figure 2.2.3 shows a non-deterministic network designed to test, i.e. extract information from, the quantum comb in figure 2.2.5. Such non-determinsitic combs, whose only output is a set of classical measurement outcomes (and their probabilities), play the the same role for combs as POVM's do for quantum states and are called QUANTUM TESTERS.

Using the link product rule (2.2.165), we can compute the probability if getting outcome x when the quantum comb and tester quantum tester are wired together as in figure 2.2.6

$$\begin{aligned} p_x &= \rho * C_1 * D_1 * C_2 * D_2 * \cdots * D_{N-1} * C_N * P_x^T \\ &= \left(\rho * D_1 * D_2 * \cdots * D_{N-1} * P_x^T\right) * \left(C_1 * C_2 * \cdots * C_N\right) \\ &= F_x * C \\ &= \text{tr}\left[F_x C^T\right] = \text{tr}\left[F_x^T C\right] \end{aligned}$$

[18] The first condition (2.2.167) on $C^{(N)}$ can be understood from figure 2.2.4 (bottom): tracing out the last output (register 9 in figure) decouples the last channel (right most green box in fig) and what remains is equivalent to the comb (sequence of green boxes) connecting inputs 0-6 to outputs 1-7. The same happens if we continue sequentially tracing-out the last output $(7, 5, ..)$ until we arrive to the last comb in the hierarchy, $C^{(1)}$ in (2.2.169) which corresponds to the first quantum channel taking input A_1^{in} to the output that might be further processed.

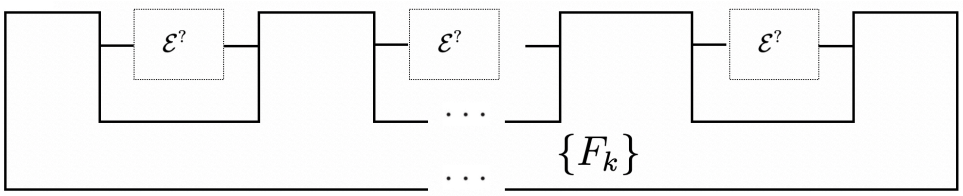


Figure 2.2.7: The quantum tester for the discrimination of channels when N uses of the channel are available.

where C is the Choi operator of the tested network, C^T is the transpose of C, and $\{F_x\}_{x\in\mathrm{X}}$ is the collection of operators defined by

$$F_x := \rho * D_1 * D_2 * \cdots * D_{N-1} * P_x^T$$

Here the transpose of P_x is needed because as mentioned in footnote 16 in comb diagrams we adopt the convention of labeling the measurement effects "boxes" with the POVM element itself P_x rather than its comb P_x^T.

We call the set of operators $F = \{F_x\}_{x\in\mathrm{X}}$ a quantum tester and equation (2.2.6) the generalized Born rule. Quantum testers admit also the following characterization:

Proposition 2.2.5 (Quantum tester) *Let* $\mathbf{F}$ *be a collection of positive operators on* $\bigotimes_{j=1}^{N}\left(\mathcal{H}_j^{\mathrm{out}}\otimes\mathcal{H}_j^{\mathrm{in}}\right)$. *$F$ is a quantum tester* if and only if

$$\sum_{x\in\mathcal{X}} F_x = \mathbb{1}_{A_N^{\mathrm{out}}}\otimes\Gamma^{(N)} \tag{2.2.170}$$

$$\mathrm{tr}_{A_N^{\mathrm{in}}}\Gamma^{(N)} = \mathbb{1}_{A_{N-1}^{\mathrm{out}}}\otimes\Gamma^{(N-1)} \tag{2.2.171}$$

$$\vdots \tag{2.2.172}$$

$$\mathrm{tr}_{A_2^{\mathrm{in}}}\Gamma^{(2)} = \mathbb{1}_{A_1^{\mathrm{out}}}\otimes\Gamma^{(1)} \tag{2.2.173}$$

$$\mathrm{tr}_{A_1^{\mathrm{in}}}\Gamma^{(1)} = 1 \tag{2.2.174}$$

where each $\Gamma^{(n)}, n = 1,\ldots,N$ *is a positive operator on* $\mathcal{H}_n^{\mathrm{in}}\otimes\left[\bigotimes_{j=1}^{n-1}\left(\mathcal{H}_j^{\mathrm{out}}\otimes\mathcal{H}_j^{\mathrm{in}}\right)\right]$

This characterization is just simple extension of the quantum tester for channels (single slot combs) found in (2.2.151).

We can extend our derivation for the optimal discrimination of quantum channels (single use) to the discrimination of quantum combs by generalized quantum testers [22]. In particular we can recover the elegant dual formulation (2.2.160).

Theorem 2.2.6 *Given a set of quantum combs* $\{C_k\}$ *appearing with prior probabilities* π_k *the optimal success probability of identifying them correctly.*

$$P_S = P_S^{\mathrm{dual}} := \min_{\lambda,C}\lambda \quad \text{s.t.} \quad \exists \text{ a quantum comb } C \text{ fulfilling } \lambda C \geq \pi_k C_k \ \forall k \tag{2.2.175}$$

Since the conditions for C to be comb, (2.2.167)-(2.2.169), are linear or positiveness constrains, the computation of P_S is an SPD. With this tool one can tackle the optimal discrimination of quantum channels after N uses: $C_x(N) = C_x^{\otimes n}$, illustrated in figure 2.2.7.

Quantum combs and quantum testers are very powerful tool whose use goes beyond quantum process discrimination. Similarly, proving that a given problem has an SDP formulation is for most practical

purposes equivalent to solving it, the reasons being: 1) There are efficient numerical codes (with implementations in python, matlab,..etc) that can provide better and better feasible solutions for the primal and dual, and thereby return provably optimal values to any desired accuracy. 2) It is often the case that inspired by the numerical results, using the symmetry of the problem or just by "brute intuition", one can give analytical proofs by constructing feasible solutions for the primal and dual problems which return the same value.

Chapter 3

Quantum Estimation
Part I: Parameter estimation

3.1 Classical Estimation

3.1.1 Introduction

We will assume that the probability mass function PMF (probability density function PDF) of a discrete (continuous) random variable, X, depends on a parameter, θ, or a set of parameters (parameter vector), $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)$. These notes are mainly devoted to the single parameter case.

Recall:

- PMF: $p_X(x_k)$ such that $p_X(x_k) := \Pr(X = x_k) = p_k$, for all values of k.
- PDF: $f_X(x)$ such that $\displaystyle\int_{-\infty}^{x} f_X(x')dx' = \Pr(X \leq x) := F_X(x)$

The function $F_X(x)$ is called the distribution function or the cumulative distribution function CDF of the random variable X. Note that $F'_X(x) = f_X(x)$. Unless it is necessary for clarity, we will often drop the subscript X and simply write $f(x)$, $F(x)$, and so on. In these notes we (mostly) focus on the continuous case, but one can check that the results we will derive hold also for discrete random variables by replacing the PDFs, $f(x;\boldsymbol{\theta})$, by the PMF, $p(x,\boldsymbol{\theta})$ and $\int \cdot\, dx \to \sum_x$.

Example 3.1.1. We say that X is normally distributed, $X \sim \mathcal{N}(\mu, \sigma^2)$, if its PDF is

$$f(x;\boldsymbol{\theta}) = f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\},$$

where $\mu = \mathbb{E}(X)$ is the *mean* and $\sigma^2 = \mathrm{var}(X) = \mathbb{E}[(X-\mu)^2] = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$ is the *variance*. We may view the mean and the variance as parameters: $\boldsymbol{\theta} = (\mu, \sigma^2)$.

Recall that $\mathbb{E}$ stands for *expectation value*:

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx.$$

Often throughout these notes, we will use boldface to denote a collection or a vector of random variables, $\boldsymbol{X} = (X_1, \ldots, X_n)$. Likewise $\boldsymbol{x}$ will be also used to denote the corresponding vector of outcomes/observation, $\boldsymbol{x} = (x_1, \ldots, x_n)$. The joint PDF will be denoted by $f_{\boldsymbol{X}}(\boldsymbol{x})$. Hence, e.g., the expectation value of $g(\boldsymbol{X})$ will be

$$\mathbb{E}[g(\boldsymbol{X})] = \int g(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) d^n x,$$

where $d^n x = dx_1 dx_2 \cdots dx_n$.

The aim of (parameter) estimation is to accurately determine the value of $\boldsymbol{\theta}$ from observations, i.e., from a set $\boldsymbol{x}$ of outcomes or realizations of the random variable X. We will refer to this set as *sample*.

The relevance of estimation for quantum information should be obvious. A quantum state ρ is just a collection of parameters (its independent entries ρ_{ab}, for instance) that describe our acknowledge about a system. To emphasize this fact, we could write $\rho_{\boldsymbol{\theta}}$ instead of just ρ. In order to have a precise mathematical description of the state, an accurate estimation of these parameters $\boldsymbol{\theta}$ is required. We can only perform measurements on the system to reveal this information. According to quantum mechanics, their outcomes are random variables, whose probability distributions are given by the Born rule, $p(\boldsymbol{x};\boldsymbol{\theta}) = \mathrm{tr}(\rho_{\boldsymbol{\theta}} E_\chi)$, where $\{E_\chi\}$ is a collection of operators defining a positive operator-valued measure (POVM) and characterizing the measurement. The classical estimation toolbox, which we are about to introduce, provides us with the means to optimally extract $\boldsymbol{\theta}$ from our measurement data. Note that the distribution $p(\boldsymbol{x};\boldsymbol{\theta})$ also depends on our choice of measurement. But what is the best

measurement we can perform on a system to estimate $\boldsymbol{\theta}$? The aim of quantum estimation is to provide means to answer this question.

In a more complex scenario, we may wish to characterize the action of a channel $\mathcal{C}_{\boldsymbol{\theta}}$. To do so, we may feed the channel with a system prepared in a fiducial or reference state, ρ_0, and perform a measurement on the output state $\rho_{\boldsymbol{\theta}} = \mathcal{C}_{\boldsymbol{\theta}}(\rho_0)$. For a fixed ρ_0 and a fixed measurement, classical estimation will provide us with the tools to obtain the most precise determination of the unknown $\boldsymbol{\theta}$.

There are several approaches to classical estimation. We will focus on two, which we will refer to as frequentist approach and Bayesian approach.

3.1.2 Frequentist approach

Within the frequentist approach the estimated parameter is assumed to be a deterministic variable with a fixed value.

Definition 3.1.1 *Given a sample of random variables (possible outcomes)* $\boldsymbol{X} = (X_1, \ldots, X_n)$*, a* statistic *$Y$ is a known function of the sample*

$$Y = f(\boldsymbol{X}).$$

When the statistic is used to estimate the value of a parameter (vector) $\boldsymbol{\theta}$ *then it is also called a* point estimate*, or* estimator *and it is usually denoted by* $\hat{\boldsymbol{\theta}}$*.*

Note that a statistic is a random variable itself.

In these notes we will always assume that X_i are independent and identically distributed (commonly abbreviated i.i.d.).

Example 3.1.2. It is well known that the sample mean (average),

$$\bar{X} := \frac{1}{n}\sum_{i=1}^{n} X_i,$$

is a "good" estimator of μ. Likewise,

$$S^2 := \frac{1}{n-1}\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2$$

is a "good" estimator of the variance. Both $\bar{X}$ and S^2 are statistics, since they just depend on $X_1, X_2, \ldots, X_n$.

To give a precise meaning to "good" above, we need to discuss some properties of the estimators.

Definition 3.1.2 *(Bias) The* bias *of an estimator $\hat{\theta}$ of a parameter θ is defined as*

$$\mathrm{Bias}(\hat{\theta}) = \mathbb{E}(\hat{\theta} - \theta).$$

If $\mathrm{Bias}(\hat{\theta}) = 0$ *then we say that the estimator is* unbiased*.*

So, if an estimator is unbiased, in average, it does give the right estimate, which, of course is a desirable property.

Example 3.1.3. The sample mean $\bar{X}$ is an unbiased estimator of the distribution mean μ, since

$$\mathbb{E}(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i) = \mathbb{E}(X) = \mu$$

by (obvious) linearity of $\mathbb{E}$. Likewise, S^2 is unbiased.

Exercise 3.1.1. *Show that S^2 is an unbiased estimator of* $\mathrm{var}(X)$.

We must show that $\mathbb{E}(S^2) = \mathrm{var}(X)$.

$$\begin{aligned}\sum_{i=1}^{n}(X_i-\bar{X})^2 &= \sum_{i=1}^{n}(X_i-\mu)^2 - \sum_{i=1}^{n}(\bar{X}-\mu)^2 - 2\sum_{i=1}^{n}(X_i-\bar{X})(\bar{X}-\mu)\\ &= \sum_{i=1}^{n}(X_i-\mu)^2 - n(\bar{X}-\mu)^2 - 2(\bar{X}-\mu)\left(\sum_{i=1}^{n}X_i - n\bar{X}\right)\\ &= \sum_{i=1}^{n}(X_i-\mu)^2 - n(\bar{X}-\mu)^2.\end{aligned}$$

We next take expectation values and recall that $\mu = \mathbb{E}(X) = \mathbb{E}(\bar{X})$:

$$(n-1)\,\mathbb{E}(S^2) = n\,\mathrm{var}(X) - n\,\mathrm{var}(\bar{X}) = (n-1)\,\mathrm{var}(X),$$

where we have used that

$$\mathrm{var}(\bar{X}) = \frac{1}{n^2}\,\mathrm{var}\left(\sum_{i=1}^{n}X_i\right) = \frac{1}{n}\,\mathrm{var}(X).$$

The estimates obtained from our samples will be always subject to errors, so we need to quantify them in a suitable way.

Definition 3.1.3 *(Mean square error) The* mean square error *of an estimator $\hat{\theta}$ is*

$$\mathrm{MSE}(\hat{\theta}) = \mathbb{E}\left[(\hat{\theta}-\theta)^2\right].$$

One can immediately check that

$$\mathrm{MSE}(\hat{\theta}) = \mathrm{var}(\hat{\theta}) + \mathrm{Bias}(\hat{\theta})^2. \tag{3.1.1}$$

Exercise 3.1.2. *Check that Eq. (3.1.1) holds.*

$$\begin{aligned}\mathbb{E}\left[(\hat{\theta}-\theta)^2\right] &= \mathbb{E}\left[\left(\hat{\theta}-\mathbb{E}(\hat{\theta})+\mathbb{E}(\hat{\theta})-\theta\right)^2\right]\\ &= \mathrm{var}(\hat{\theta}) + \left[\mathbb{E}(\hat{\theta})-\theta\right]^2 + 2\left[\mathbb{E}(\hat{\theta})-\theta\right]\mathbb{E}\left[\hat{\theta}-\mathbb{E}(\hat{\theta})\right]\\ &= \mathrm{var}(\hat{\theta}) + \left[\mathbb{E}(\hat{\theta}-\theta)\right]^2 + 2\,\mathbb{E}(\hat{\theta}-\theta)\times 0\\ &= \mathrm{var}(\hat{\theta}) + \mathrm{Bias}(\hat{\theta})^2.\end{aligned}$$

A good estimator is one that has small MSE. If it is unbiased, this is tantamount to having small variance. Notice that $\mathrm{var}(\hat{\theta})$ can be determined from the data, wheres $\mathrm{MSE}(\hat{\theta})$ cannot, since in practical applications θ is, of course, unknown.

Often the goodness of an estimator (or rather of a sequence of estimators) improves as the sample size, n, increases. The next definition captures this idea.

Definition 3.1.4 *(Consistency) A sequence of statistics* $(Y_n, n \in \mathbb{N})$ *is said to be a* consistent estimate *of a parameter* θ *if for every* $\epsilon > 0$

$$\lim_{n\to\infty} \Pr\left(|Y_n - \theta| \leq \epsilon\right) = 1.$$

Equivalently, we may write the condition as

$$\lim_{n\to\infty} Y_n = \theta \quad \text{(in probability)}.$$

The sequence $\{Y_n; n \in \mathbb{N}\}$ could, according to our notation, be denoted by $\hat{\theta}_n$, and we will often (but not always) do so, particularly if we want to emphasize that each y_n is an estimate of θ and also indicate that it is based on a sample of size n.

Exercise 3.1.3. *Show that if* Y_n *is a consistent estimator of* θ *with* $\mathbb{E}(Y_n^2) < \infty$*, then*

$$\lim_{n\to\infty} \mathbb{E}(Y_n - \theta) = 0.$$

We first note that there exists a finite C such that for all n

$$\begin{aligned}\mathbb{E}\left[(Y_n-\theta)^2\right] &\leq \mathbb{E}(Y_n^2) + \theta^2 + 2|\theta||\,\mathbb{E}(Y_n)| \leq \mathbb{E}(Y_n^2) + \theta^2 + 2|\theta|\sqrt{\mathbb{E}(Y_n^2)} \\ &= \left[|\theta| + \sqrt{\mathbb{E}(Y_n^2)}\right]^2 \leq C,\end{aligned}$$

where in the second inequality we have used that

$$\mathbb{E}(|Y_n|) \leq \sqrt{\mathbb{E}(Y_n^2)}$$

(as follows immediately from Jensen inequality). We next use Cauchy-Schwarz inequality,

$$\left[\mathbb{E}(|XY|)\right]^2 \leq \mathbb{E}\left(X^2\right)\mathbb{E}\left(Y^2\right),$$

to get that, for any $\epsilon > 0$,

$$\begin{aligned}\left[\mathbb{E}(|Y_n-\theta|)\mathbb{1}\{|Y_n-\theta| \geq \epsilon/2\}\right]^2 &\leq \mathbb{E}\left(|Y_n-\theta|^2\right)\mathbb{E}\left(\mathbb{1}\{|Y_n-\theta| \geq \epsilon/2\}\right) \\ &= \mathbb{E}\left[(Y_n-\theta)^2\right]\Pr\left(|Y_n-\theta| \geq \epsilon/2\right) \\ &\leq C\Pr\left(|Y_n-\theta| \geq \epsilon/2\right),\end{aligned}$$

where $\mathbb{1}\{\cdots\}$ is the indicator function. With this,

$$\begin{aligned}\mathbb{E}\left(|Y_n-\theta|\right) &= \mathbb{E}\left[|Y_n-\theta|\,\mathbf{1}\left\{|Y_n-\theta| < \epsilon/2\right\}\right] + \mathbb{E}\left[|Y_n-\theta|\,\mathbf{1}\left\{|Y_n-\theta| \geq \epsilon/2\right\}\right] \\ &\leq \epsilon/2 + C\Pr\left(|Y_n-\theta| \geq \epsilon/2\right). \qquad (3.1.2)\end{aligned}$$

But

$$\lim_{n\to\infty} Y_n = \theta \quad \text{(in probability)} \quad \Rightarrow \quad \lim_{n\to\infty} \Pr\left(|Y_n-\theta| \geq \epsilon/2\right) = 0.$$

This implies that there exists $N \in \mathbb{N}$ such that for any $\epsilon > 0$, $\Pr\left(|Y_n - \theta/2| \geq \epsilon\right) < \epsilon/(2C)$ provided $n > N$. Hence, from Eq. (3.1.2) we have

$$|\mathbb{E}(Y_n-\theta)| \leq \mathbb{E}\left(|Y_n-\theta|\right) < \frac{\epsilon}{2} + C\frac{\epsilon}{2C} = \epsilon,$$

which means that

$$\lim_{n\to\infty} \mathbb{E}(Y_n-\theta) = 0.$$

It is interesting to note that the claim of the exercise ceases to be true if we drop the condition $\mathbb{E}(Y_n^2) < \infty$.

Exercise 3.1.4. *Consider the sequence of random variables Y_n with probability distribution*

$$p_{Y_n}(y;\theta) = \begin{cases} \dfrac{n-1}{n} \text{ if } y = \theta \\ \dfrac{1}{n} \text{ if } y = \theta + n \\ 0 \text{ otherwise} \end{cases}$$

(a) Show that $\lim_{n\to\infty} \Pr(|Y_n - \theta| \le \epsilon) = 1$, but $\lim_{n\to\infty} \mathbb{E}(Y_n - \theta) \neq 0$.

(b) Modify the distribution slightly to show that $\lim_{n\to\infty} \Pr(|Y_n - \theta| \le \epsilon) = 1$ does not necessarily imply $\lim_{n\to\infty} \operatorname{var}(Y_n) = 0$, even if $\lim_{n\to\infty} \mathbb{E}(Y_n - \theta) = 0$ and $\mathbb{E}(Y_n^2) < \infty$.

(a) If $0 < \epsilon < 1$

$$\Pr(|Y_n - \theta| \le \epsilon) = \Pr(Y_n = \theta) = p_{Y_n}(\theta;\theta) = \frac{n-1}{n},$$

hence $\lim_{n\to\infty} \Pr(|Y_n - \theta| \le \epsilon) = 1$. However

$$\begin{aligned} \mathbb{E}(Y_n - \theta) &= (\theta - \theta)\, p_{Y_n}(\theta;\theta) + (\theta + n - \theta)\, p_{Y_n}(\theta + n;\theta) \\ &= n \cdot \frac{1}{n} = 1 \neq 0. \end{aligned}$$

(b) Consider

$$p_{Y_n}(y;\theta) = \begin{cases} \dfrac{n^2-1}{n^2} \text{ if } y = \theta \\ \dfrac{1}{n^2} \text{ if } y = \theta + n \\ 0 \text{ otherwise} \end{cases}$$

Then

$$\begin{aligned} \mathbb{E}(Y_n - \theta) &= (\theta - \theta)\, p_{Y_n}(\theta;\theta) + (\theta + n - \theta)\, p_{Y_n}(\theta + n;\theta) \\ &= n \cdot \frac{1}{n^2} \to 0. \\ \mathbb{E}(Y_n^2) &= \theta^2 \frac{n^2-1}{n^2} + (\theta + n)^2 \frac{1}{n^2} = \theta^2 + 2\frac{\theta}{n} + 1 \to \theta^2 + 1 \end{aligned}$$

Note that as a consequence of the result of Exercise 3.1.3 *any consistent estimator is asymptotically unbiased, in the sense that* $\lim_{n\to\infty} \mathbb{E}(\hat{\theta}_n) = \theta$. Hence, although the biased estimators may lead to improved precision, they may be ignored in the $n \to \infty$ limit for which the frequentist approach is really designed.

In dealing with consistency it might be useful to introduce the famous law of large numbers as follows.

Theorem 3.1.5 *(Law of Large Numbers) Suppose that $\{X_i, i \ge 0\}$ are i.i.d. with finite mean μ and variance σ^2. Then*

$$\hat{\mu}_n := \frac{1}{n} \sum_{i=1}^{n} X_i$$

is a consistent estimator of the mean. In other words for all $\epsilon > 0$

$$\Pr(|\hat{\mu}_n - \mu| > \epsilon) \to 0 \quad \text{as } n \to \infty.$$

The Law of Large Numbers follows from Chebyshev's Inequality.

Proposition 3.1.6 *(Chebyshev's Inequality) Let Y be a random variable with finite mean μ and variance σ^2. Then for any $k > 0$,*

$$\Pr\left(|Y - \mu| \geq k\sigma\right) \leq \frac{1}{k^2}.$$

This proposition in turn follows from Markov's Inequality:

Theorem 3.1.7 *(Markov's Inequality) Let $X > 0$ be a random variable, such that $\mathbb{E}(X) < \infty$ and $c > 0$ a constant. Then*

$$\Pr(X > c) \leq \frac{\mathbb{E}(X)}{c}.$$

Exercise 3.1.5. *Prove Markov's Inequality and the following slightly more general statement:*

If $X > 0$, $k \geq 1$ and $\mathbb{E}(X^k) < \infty$, then

$$\mathbb{P}(X > c) = \mathbb{P}\left(X^k > c^k\right) \leq \frac{\mathbb{E}(X^k)}{c^k}.$$

$$\Pr(X > c) = \int_c^\infty f_X(x)dx \leq \int_c^\infty \frac{x^k}{c^k} f_X(x)dx \leq \int_0^\infty \frac{x^k}{c^k} f_X(x)dx = \frac{\mathbb{E}(X^k)}{c^k}$$

Exercise 3.1.6. *Prove Chebyshev's Inequality.*

Exercise 3.1.7. *By using Chebyshev's inequality, show that if $\lim_{n\to\infty} \operatorname{var}(\hat{\theta}_n) = 0$ then asymptotic unbiasedness implies consistency.*

Exercise 3.1.8. *Prove the Law of Large Numbers.*

In addition to consistency and unbiasedness, a good estimator should have a small mean square error which for unbiased estimators is just the variance. This motivates the following

Definition 3.1.8 *(Relative Efficiency) Given two estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ of a parameter θ, the relative efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_2$, is denoted by $\operatorname{eff}(\hat{\theta}_1, \hat{\theta}_2)$ and is defined as*

$$\operatorname{eff}\left(\hat{\theta}_1, \hat{\theta}_2\right) = \frac{\operatorname{MSE}\left(\hat{\theta}_2\right)}{\operatorname{MSE}\left(\hat{\theta}_1\right)}.$$

For unbiased estimators it is equivalent to

$$\operatorname{eff}\left(\hat{\theta}_1, \hat{\theta}_2\right) = \frac{\operatorname{var}\left(\hat{\theta}_2\right)}{\operatorname{var}\left(\hat{\theta}_1\right)}.$$

Exercise 3.1.9. *Let $X_1, \ldots, X_n \sim \mathcal{U}(0, M)$ the uniform distribution on $(0, M)$. Consider the estimators*

$$\hat{\theta}_1 := \bar{X} = \frac{1}{n}\sum_{i=1}^n X_i, \quad \hat{\theta}_2 := c_n \frac{M_n}{2},$$

where

$$M_n := \max\{X_i\}_{i=1}^n.$$

and c_n is some judiciously chosen n-dependent normalization coefficient. Give c_n so that $\hat{\theta}_2$ is unbiased. Then compute $\text{eff}(\hat{\theta}_1, \hat{\theta}_2)$.

At this point, the question arises as to how to construct consistent unbiased and effective estimators. Before we attempt to answer this question we still need to give a few definitions.

Definition 3.1.9 *(Likelihood function for discrete PD) Suppose $\boldsymbol{X} = (X_1, \ldots, X_n)$ are discrete random variables whose distribution depends on a parameter(vector) $\boldsymbol{\theta}$, and have probability mass function*

$$p(\boldsymbol{x}; \boldsymbol{\theta}) := \Pr(\boldsymbol{X} = \boldsymbol{x}; \boldsymbol{\theta}),$$

where $\boldsymbol{x} = (x_1, \ldots, x_n)$ are sample observations. The likelihood of the parameter(vector) $\boldsymbol{\theta}$ given the observations $\boldsymbol{x}$ is denoted by $L(\boldsymbol{\theta} \,|\, \boldsymbol{x})$ is defined to be

$$L(\boldsymbol{\theta} \,|\, \boldsymbol{x}) := p(\boldsymbol{x}; \boldsymbol{\theta})$$

that is the joint probability mass function for the parameter(vector) $\boldsymbol{\theta}$.

Definition 3.1.10 *(Likelihood function for continuous PD) Suppose $\boldsymbol{X} = (X_1, \ldots, X_n)$ are jointly continuous random variables whose distribution depends on a parameter(vector) $\boldsymbol{\theta}$, and have PDF $f(\boldsymbol{y}; \boldsymbol{\theta})$. The likelihood of the parameter(vector) $\boldsymbol{\theta}$ given the observations $\boldsymbol{x}$ is denoted by $L(\boldsymbol{\theta} \,|\, \boldsymbol{x})$ is defined to be*

$$L(\boldsymbol{\theta} \,|\, \boldsymbol{x}) := f(\boldsymbol{x}; \boldsymbol{\theta})$$

that is the joint density for the parameter(vector) $\boldsymbol{\theta}$ evaluated at the observations.

Definition 3.1.11 *(Log-likelihood function) The log-likelihood function of the parameter (vector) $\boldsymbol{\theta}$ given the observations $\boldsymbol{x}$ is defined as*

$$l(\boldsymbol{\theta}) = l(\boldsymbol{\theta} \,|\, \boldsymbol{x}) = \log L(\boldsymbol{\theta} \,|\, \boldsymbol{x}),$$

where $L(\boldsymbol{\theta} \,|\, \boldsymbol{x})$ is the corresponding likelihood function.

If $X_1, \ldots, X_n$ are i.i.d. and $X_i \sim f(\cdot\,; \boldsymbol{\theta})$, then

$$L(\boldsymbol{\theta} \,|\, \boldsymbol{x}) = \prod_{i=1}^n f(x_i; \boldsymbol{\theta}); \quad l(\boldsymbol{\theta} \,|\, \boldsymbol{x}) = \sum_{i=1}^n \log\left[f(x_i; \boldsymbol{\theta})\right].$$

and similarly for discrete random variables.

We think of the likelihood function as a function of $\boldsymbol{\theta}$, and we treat the observations as fixed. Sometimes we will drop the observations and simply write $L(\boldsymbol{\theta})$ and $l(\boldsymbol{\theta})$.

Definition 3.1.12 *(Maximum Likelihood Estimator) Suppose that a sample $\boldsymbol{x} = (x_1, \ldots, x_n)$ has likelihood function $L(\boldsymbol{\theta}) = L(\boldsymbol{\theta} \,|\, \boldsymbol{x})$ depending on a parameter(vector) $\boldsymbol{\theta}$. Then a* maximum likelihood estimator (MLE) $\hat{\boldsymbol{\theta}}_{\text{MLE}}$ *is the value of the parameters that maximizes $L(\boldsymbol{\theta})$, if a maximum exists. In other words*

$$\hat{\boldsymbol{\theta}}_{\text{MLE}} = \arg\max_{\boldsymbol{\theta}} L(\boldsymbol{\theta} \,|\, \boldsymbol{x}) = \arg\max_{\boldsymbol{\theta}} l(\boldsymbol{\theta} \,|\, \boldsymbol{x})$$

The maximum of $L(\boldsymbol{\theta})$ may not exist, in which case the MLE cannot be constructed. The maximum, if it exists, may not be unique, in which case we will obtain several MLEs. Note that these are not the values of the parameters that are most likely, given the data. To start with, $\boldsymbol{\theta}$ in not a random variable in the frequentist approach we are discussing!

Theorem 3.1.13 *(Invariance of MLE) Suppose that $\hat{\theta}$ is the MLE for a parameter θ and let $t(\cdot)$ be a strictly monotone function of θ. Then*

$$(t(\theta))_{\text{MLE}} = t(\hat{\theta}),$$

i.e., the MLE of $t(\theta)$ is $t(\hat{\theta})$.

Exercise 3.1.10. *Consider the exponential distribution $f(x \mid \lambda) = \lambda e^{-\lambda x}$. Suppose we take a sample of size n. Show that the MLE of λ is $\hat{\lambda} = 1/\bar{X}$.*

The likelihood, is

$$L(\lambda \mid x_1, \ldots, x_n) = \prod_{i=1}^{n} \left(\lambda e^{-\lambda x_i}\right) = \lambda^n \exp\left(-\lambda \sum_{i=1}^{n} x_i\right) = \lambda^n \exp(-n\lambda \bar{x})$$

Then

$$l = n \log \lambda - n\lambda \bar{x}$$

and so

$$\frac{d}{d\lambda} l = \frac{n}{\lambda} - n\bar{x}$$

Thus L has a unique maximum at $\hat{\lambda} = 1/\bar{x}$ and this is therefore the maximum likelihood estimator of λ.

Exercise 3.1.11. *Find the maximum likelihood estimates of the parameters of the normal distribution. Note that the MLE of σ^2 is a biased estimator.*

$$l = -\frac{n}{2} \log(2\pi) - n \log \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2$$

Differentiating with respect to each parameter and setting equal to zero:

$$\frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu) = 0;$$
$$-\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^{n} (x_i - \mu)^2 = 0.$$

It follows that

$$\hat{\mu} = \frac{1}{n} \sum_{i}^{n} x_i \quad \hat{\sigma}^2 = \frac{1}{n} \sum (x_i - \bar{x})^2 .$$

We note that

$$\hat{\sigma}^2 = \frac{n-1}{n} S^2.$$

Since S^2 is unbiased, $\hat{\sigma}^2$ must be biased.

Proposition 3.1.14 *(the MLE of an i.i.d observation is consistent) Let $\{X_1, \ldots, X_n\}$ be a sequence of i.i.d. observations where*

$$X_k \overset{\text{i.i.d.}}{\sim} f(x; \theta).$$

Then the MLE of θ is consistent

Let $x_1, \ldots, x_n$ be a sample drawn from a population with PDF $f_\theta(x)$. When the sample is used to estimate the parameter θ, an obvious question arises: What is the lowest variance we can achieve?

Definition 3.1.15 *(Fisher Information) Let $X \sim f(\cdot\,; \theta)$. Then the Fisher Information is given by*

$$I_n(\theta) := n\,\mathbb{E}\left[\left(\frac{\partial l(\theta \mid x)}{\partial \theta}\right)^2\right],$$

where $l(\theta \mid x)$ is the log-likelihood.

It can be shown that if the second partial derivative exists then we also have that

$$I_n(\theta) = -n\,\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2} l(\theta\,|\,x)\right].$$

Exercise 3.1.12. *Prove this last statement.*

$$\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2} l(\theta\,|\,x)\right] = \int f(x;\theta)\partial_\theta^2 \log\left[f(x;\theta)\right] dx = \int f(x;\theta)\partial_\theta\left[\frac{\partial_\theta f(x;\theta)}{f(x;\theta)}\right] dx$$
$$= \int f(x;\theta)\left\{\frac{\partial_\theta^2 f(x;\theta)}{f(x;\theta)} - \frac{\left[\partial_\theta f(x;\theta)\right]^2}{f^2(x;\theta)}\right\} dx,$$

where we have used the obvious notation $\partial_\theta := \partial/\partial\theta$. The first term in the integral vanishes, since

$$\int f(x;\theta)\frac{\partial_\theta^2 f(x;\theta)}{f(x;\theta)} dx = \int \partial_\theta^2 f(x;\theta) dx = \partial_\theta^2 \int f(x;\theta) dx = \partial_\theta 1 = 0.$$

Finaly, the second term can be written as

$$-\int f(x;\theta)\left[\frac{\partial_\theta f(x;\theta)}{f(x;\theta)}\right]^2 dx = -\int f(x;\theta)\left[\partial_\theta l(\theta\,|\,x)\right]^2 dx = -\,\mathbb{E}\left[\left(\frac{\partial l(\theta\,|\,x)}{\partial\theta}\right)^2\right].$$

Theorem 3.1.16 *(Cramer-Rao bound) Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be i.i.d. with probability density function $f(x;\theta)$. Let $\hat{\theta}_n = g(\boldsymbol{X})$ be an unbiased estimator of θ, such that the support of $g(\boldsymbol{X})$ (the region for which the probability is not zero) does not depend on θ. Then under mild conditions we have that*

$$\mathrm{var}\left(\hat{\theta}_n\right) \geq \frac{1}{I_n(\theta)}.$$

The proof relies on the following very general theorem/definition

Theorem 3.1.17 *Let X and Y be two random variables. Then, their* correlation coefficient*, defined as*

$$\mathrm{corr}(X,Y) := \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{var}(X)\,\mathrm{var}(Y)}},$$

satisfies

$$1 \leq \mathrm{corr}(X,Y) \leq 1.$$

We recall that $\mathrm{cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y)$, $\mathrm{cov}(X,X) = \mathrm{var}(X)$. The content of this Theorem is that cov actually obeys the Cauchy-Schwarz inequality.

Exercise 3.1.13. *Prove Theorem 3.1.17.* Hint. *First check that linearity of the expectation, $\mathbb{E}$, implies $\mathrm{corr}(aX + c, bY + c') = \mathrm{corr}(X,Y)$, hence, it suffices to prove the theorem assuming $\mathbb{E}(X) = \mathbb{E}(Y) = 0$ and $\mathrm{var}(X) = \mathrm{var}(Y) = 1$. Next, consider the trivial inequality $0 \leq \mathbb{E}\left[(X - \lambda Y)^2\right]$, where $\lambda \in \mathbb{R}$.*

One can immediately check that from the very definition of $\mathbb{E}$ one has $\mathbb{E}(aX+c) = a\,\mathbb{E}(X)+c$. Then,

$$\mathrm{var}(aX + c) = \mathbb{E}\{[(aX + c - \mathbb{E}(aX + c)]^2\} = a^2\,\mathbb{E}\{[(X - \mathbb{E}(X)]^2\} = a^2\,\mathrm{var}(X),$$

and

$$\begin{aligned}\mathbb{E}[(aX + c)(bY + c')] &= \mathbb{E}[abXY + cbY + c'aX + cc'] \\ &= ab\,\mathbb{E}(XY) + cb\,\mathbb{E}(Y) + c'a\,\mathbb{E}(X) + cc' \\ &= ab[\mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y)] + [a\,\mathbb{E}(X) + c][b\,\mathbb{E}(Y) + c'],\end{aligned}$$

which proves that

$$\operatorname{cov}(aX + c, bY + c') = ab \operatorname{cov}(X, Y).$$

Hence, $\operatorname{corr}(aX + c, bY + c') = \operatorname{corr}(X, Y)$.

We see that

$$X' = \frac{X - \mathbb{E}(X)}{\sqrt{\operatorname{var}(X)}}; \quad Y' = \frac{Y - \mathbb{E}(Y)}{\sqrt{\operatorname{var}(Y)}} \quad \Rightarrow \quad \operatorname{corr}(X', Y') = \operatorname{corr}(X, Y)$$

and $\mathbb{E}(X') = \mathbb{E}(Y') = 0$ and $\operatorname{var}(X') = \operatorname{var}(Y') = 1$. This proves the first statement in the exercise.

Next, we have

$$0 \leq \mathbb{E}\left[(X - \lambda Y)^2\right] = \lambda^2 - 2\lambda\, \mathbb{E}(XY) + 1.$$

For this to hold, the polynomial in λ on the right hand side can have at most one root, which implies that the discriminant must be non-positive: $1 \geq [\mathbb{E}(X, Y)]^2 = [\operatorname{corr}(X, Y)]^2$.

Proof.
of the Cramer-Rao Bound (CRB) Consider the random variable W defined by

$$W = \partial_\theta \log f(\boldsymbol{X}; \theta) = \frac{\partial_\theta f(\boldsymbol{X}; \theta)}{f(\boldsymbol{X}; \theta)},$$

where $\boldsymbol{X} = (X_1, \ldots, X_n)$, $f(\boldsymbol{x}; \theta)$ is the joint PDF, $f(\boldsymbol{x}; \theta) = \prod_{i=1}^n f(x_i; \theta)$, and $\partial_\theta := \partial/\partial\theta$. Hence

$$\mathbb{E}(W) = \int \frac{\partial_\theta f(\boldsymbol{x}; \theta)}{f(\boldsymbol{x}; \theta)} f(\boldsymbol{x}; \theta) d^n x = \int \partial_\theta f(\boldsymbol{x}; \theta) d^n x = \frac{d}{d\theta} \int f(\boldsymbol{x}; \theta) d^n x = 0$$

under fairly general conditions that guarantee we can exchange differentiation and integration. Since $\mathbb{E}(W) = 0$, we have that $\operatorname{cov}\left(W, \hat{\theta}_n\right) = \mathbb{E}\left(W\hat{\theta}_n\right)$, thus

$$\begin{aligned} \operatorname{cov}\left(W, \hat{\theta}_n\right) &= \int g(\boldsymbol{x}) \frac{\partial_\theta f(\boldsymbol{x}; \theta)}{f(\boldsymbol{x}; \theta)} f(\boldsymbol{x}; \theta) d^n x = \int g(\boldsymbol{x})\, \partial_\theta f(\boldsymbol{x}; \theta) d^n x = \frac{d}{d\theta} \int g(\boldsymbol{x}) f(\boldsymbol{x}; \theta) d^n x \\ &= \frac{d}{d\theta} \mathbb{E}(\hat{\theta}_n) = \frac{d\theta}{d\theta} = 1. \end{aligned}$$

From Theorem 3.1.17 we have

$$1 \geq [\operatorname{corr}(W, \hat{\theta}_n)]^2 = \frac{\operatorname{cov}^2(W, \hat{\theta}_n)}{\operatorname{var}(W) \operatorname{var}(\hat{\theta}_n)},$$

which, since $\operatorname{cov}(W, \hat{\theta}_n) = 1$, implies that

$$\operatorname{var}(\hat{\theta}_n) \geq \frac{1}{\operatorname{var}(W)}.$$

Note that up to this point we have not used the i.i.d. condition, hence, the last bound holds in the general situation. Assuming now that $X_1, \ldots, X_n$ are i.i.d., we know that the joint distribution is simply $f(\boldsymbol{x}; \theta) = \prod_{i=1}^n f(x_i; \theta)$, therefore

$$W = \partial_\theta \sum_{i=1}^n \log f(X_i, \theta) = \sum_{i=1}^n \partial_\theta \log f(X_i, \theta). =: \sum_{i=1}^n W_i$$

Using again the independency condition, we have

$$\operatorname{var}(W) = \sum_{i=1}^n \operatorname{var}(W_i) = n\, \mathbb{E}\left[\partial_\theta \log f(x, \theta)\right]^2 = n\, \mathbb{E}\left[\partial_\theta l(\theta \,|\, x)\right]^2 = I_n(\theta).$$

This completes the proof. ■

In particular, the first equality in the last line of the proof states that (additivity)

$$I_1^{(1,2)}(\theta) = I_1^{(1)}(\theta) + I_1^{(2)}(\theta)$$

for independent random variables with joint distribution $f_{X_1X_2}(x_1,x_2;\theta)=f_{X_1}(x_1;\theta)f_{X_2}(x_2;\theta)$. Being non-negative [as follows from (its very) Definition 3.1.15] and additive the FI has the interpretation of an *information measure*. Its increase indicates that a higher precision is potentially achievable in parameter estimation. In particular, at a given θ_0, $I_n(\theta_0) = 0$ proves that one cannot extract any information about the parameter from a sample, whereas divergent $I_n(\theta_0) = \infty$ implies that the true value θ_0 can in principle be perfectly determined.

Definition 3.1.18 *(Efficiency) The efficiency of an unbiased estimator $\hat{\theta}_n$ of a parameter θ is defined as the ratio of the Cramer-Rao bound to the variance of $\hat{\theta}_n$, that is*

$$\mathrm{eff}\left(\hat{\theta}_n\right) = \frac{1}{I_n(\theta)\,\mathrm{var}\left(\hat{\theta}_n\right)}.$$

An estimator which has unit efficiency [the maximum value $\mathrm{eff}(\hat{\theta}_n)$ can take] is called efficient.

Exercise 3.1.14. *Let $X_1,\ldots,X_n$ be i.i.d. with PDF $f(x;\lambda) = \lambda \mathrm{e}^{-\lambda x}$. In Exercise 3.1.10 it was shown that $\hat{\lambda}_{\mathrm{MLE}} = 1/\bar{X}$. Show now that $\hat{\lambda}_{\mathrm{MLE}}$ is not unbiased, whereas $\hat{\lambda}_n = (n-1)/(\sum_{i=1}^n X_i)$ is. Show that*

$$\mathrm{eff}(\hat{\lambda}_n) = 1 - \frac{2}{n}.$$

This is less than unity, and hence, it is not efficient. However, the efficiency approaches unity as $n \to \infty$. In such cases we say that $\hat{\lambda}_n$ is an asymptotically efficient estimator.

Let us compute $\mathbb{E}(\hat{\lambda}_{\mathrm{MLE}})$. We will do it brute force:

$$\mathbb{E}(\hat{\lambda}_{\mathrm{MLE}}) = \int_{\mathbb{R}_+^n} \frac{n}{\sum_{i=1}^n x_i} \lambda^n \mathrm{e}^{-\lambda \sum_{i=1}^n x_i} d^n x.$$

Insert the identity

$$\int_0^\infty \delta\left(\textstyle\sum_{i=1}^n x_i - w\right) dw = 1,$$

where $\delta(x)$ is the Dirac delta function (distribution). We have

$$\mathbb{E}(\hat{\lambda}_{\mathrm{MLE}}) = n\lambda^n \int_0^\infty \frac{\mathrm{e}^{-\lambda w}}{w} dw \int_{\mathbb{R}_+^n} \delta\left(\textstyle\sum_{i=1}^n x_i - w\right) d^n x.$$

Scale x_i as $x_i = w y_i$, then $d^n x = w^n d^n y$, and

$$\begin{aligned}
\mathbb{E}(\hat{\lambda}_{\mathrm{MLE}}) &= n\lambda^n \int_0^\infty w^{n-1}\mathrm{e}^{-\lambda w} dw \int_{\mathbb{R}_+^n} \delta\left[w\left(\textstyle\sum_{i=1}^n y_i - 1\right)\right] d^n y \\
&= n\lambda^n \int_0^\infty w^{n-2}\mathrm{e}^{-\lambda w} dw \int_{\mathbb{R}_+^n} \delta\left(\textstyle\sum_{i=1}^n y_i - 1\right) d^n y \\
&= n(n-2)!\lambda\,\mathrm{vol}(\Delta^n),
\end{aligned}$$

where $\mathrm{vol}(\Delta^n)$ is the volume of the simplex $\Delta^n = \{(y_1,\ldots,y_n) \mid \sum_{k=1}^n y_k = 1\}$. Using the same trick with $\mathbb{E}(1) = 1$ we have

$$1 = \int_{\mathbb{R}_+^n} \lambda^n \mathrm{e}^{-\lambda \sum_{i=1}^n x_i} d^n x = \lambda^n \int_0^\infty w^{n-1}\mathrm{e}^{-\lambda w} dw \int_{\mathbb{R}_+^n} \delta\left(\textstyle\sum_{i=1}^n y_i - 1\right) d^n y = (n-1)!\,\mathrm{vol}(\Delta^n).$$

Hence $\mathrm{vol}(\Delta^n) = 1/(n-1)!$ and

$$\mathbb{E}(\hat{\lambda}_{\mathrm{MLE}}) = \frac{n(n-2)!}{(n-1)!}\lambda = \frac{n}{n-1}\lambda.$$

We see that $\hat{\lambda}_{\mathrm{MLE}}$ is *not* unbiased (though it is *asymptotically unbiased*). Thus

$$\hat{\lambda}_n = \frac{n-1}{n}\hat{\lambda}_{\mathrm{MLE}} = \frac{n-1}{\sum_{i=1}^n X_i}$$

is unbiased.

Next, let us compute the efficiency of this estimator. We first need to compute the variance, which we do applying once again the very same trick as before

$$\begin{aligned}
\mathrm{var}(\hat{\lambda}_n) &= \int_{\mathbb{R}^n_+} \left(\frac{n-1}{\sum_{i=1}^n x_i} - \lambda \right)^2 \lambda^n \mathrm{e}^{-\lambda \sum_{i=1}^n x_i} d^n x \\
&= \mathrm{vol}(\Delta^n) \lambda^n \int_0^\infty w^{n-1} \left(\frac{n-1}{w} - \lambda \right)^2 \mathrm{e}^{-\lambda w} dw \\
&= \frac{\lambda^n}{(n-1)!} \int_0^\infty \left[(n-1)^2 w^{n-3} - 2(n-1)\lambda w^{n-2} + \lambda^2 w^{n-1} \right] \mathrm{e}^{-\lambda w} dw \\
&= \frac{(n-1)^2(n-3)! - 2(n-1)(n-2)! + (n-1)!}{(n-1)!} \lambda^2 \\
&= \left(\frac{n-1}{n-2} - 1 \right) \lambda^2 \\
&= \frac{\lambda^2}{n-2}.
\end{aligned}$$

We also need the Fisher information:

$$\begin{aligned}
I_n(\lambda) &= n \, \mathbb{E} \left\{ \left[\frac{\partial}{\partial \lambda} (\log \lambda - \lambda x) \right]^2 \right\} \\
&= n \, \mathbb{E} \left[\left(\frac{1}{\lambda} - x \right)^2 \right] \\
&= n\lambda \int_0^\infty \left(\frac{1}{\lambda^2} - 2\frac{x}{\lambda} + x^2 \right) \mathrm{e}^{-\lambda x} \\
&= \frac{n}{\lambda^2}.
\end{aligned}$$

Combining these last two results we get

$$\mathrm{eff}\left(\hat{\lambda}_n\right) = \frac{1}{I_n(\theta)\,\mathrm{var}\left(\hat{\theta}_n\right)} = \frac{1}{\frac{n}{\lambda^2} \cdot \frac{\lambda^2}{n-2}} = 1 - \frac{2}{n}.$$

Alternatively, we could have computed $\mathrm{vol}(\Delta^n)$ by a change of variables. For instance, if we had to compute the integral of some function $g(y_1, y_2, \ldots, y_n)$ over the simplex Δ^n, i.e.,

$$I_g = \int_{\mathbb{R}^n_+} \delta\left(\sum_{i=1}^n y_i - 1\right) g(y_1, y_2, \ldots, y_n) d^n y,$$

we could define, e.g.,

$$\begin{aligned}
y_1 &= u_1, \\
y_2 &= (1-u_1)u_2, \\
y_3 &= (1-u_1)(1-u_2)u_3, \\
&\vdots \qquad\qquad \vdots \\
y_{n-1} &= (1-u_1)(1-u_2)\cdots u_{n-1}, \\
y_n &= (1-u_1)(1-u_2)\cdots(1-u_{n-1}).
\end{aligned}$$

(Note that the variable y_n is not independent!) So that $\sum_{i=1}^n y_i = 1$. Note that $0 \leq u_i \leq 1$, for $i = 1, 2, \ldots, n-1$. The Jacobian of the change is very easy to compute because the Jacobian matrix is

lower triangular

$$\frac{\partial(y_1,\dots,y_{n-1})}{\partial(u_1,\dots,u_{n-1})} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ * & 1-u_1 & 0 & \cdots & 0 \\ * & * & (1-u_1)(1-u_2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & \cdots & \prod_{i=1}^{n-2}(1-x_i) \end{pmatrix}.$$

Hence

$$\left|\frac{\partial(y_1,\dots,y_{n-1})}{\partial(u_1,\dots,u_{n-1})}\right| = (1-u_{n-2})(1-u_{n-3})^2(1-u_{n-4})^3\cdots(1-u_1)^{n-2},$$

and we have

$$I_g = \int_0^1 du_{n-1}\int_0^1 (1-u_{n-2})du_{n-2}\cdots\int_0^1 (1-u_1)^{n-2} g\big(u_1,(1-u_1)u_2,\dots,\textstyle\prod_{i=1}^n(1-u_i)\big)du_1.$$

In the particular case $g\equiv 1$ we obtain

$$\mathrm{vol}(\Delta^n) = I_1 = 1\cdot\frac{1}{2}\cdot\frac{1}{3}\cdots\frac{1}{n-1} = \frac{1}{(n-1)!}.$$

Proposition 3.1.19 *(Asymptotic normality) Let $\{X_1,\cdots,X_n\}$ be a sequence of i.i.d. observations where*

$$X_k \overset{\text{i.i.d.}}{\sim} f(x;\theta).$$

Let $\hat\theta$ be a MLE of θ, then

$$\sqrt{n}(\hat\theta_n-\theta) \xrightarrow{\text{d}} \mathcal{N}\left(0,\frac{1}{I_1(\theta)}\right).$$

(See Lehmann, *Elements of Large Sample Theory*, Springer, 1999 for a proof.) The meaning of convergence in distribution is given in this

Definition 3.1.20 *(Convergence in Distribution) A sequence of random variables $\{X_1,X_2,X_3,\dots\}$ converges in distribution to a random variable X, shown by $X_n \xrightarrow{\text{d}} X$, if*

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

for all x at which the CDF, $F_X(x)$, is continuous.

At this point it is also useful to recall the central limit theorem, which we quote without proof

Theorem 3.1.21 *(Central Limit Theorem). Let $X_1,X_2,\dots$ be i.i.d. random variables with $\mathbb{E}(X_i)=\mu$ and $\mathrm{var}(X_i)=\sigma^2<\infty$. Define*

$$Z_n := \frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} = \frac{\bar X-\mu}{\sigma/\sqrt{n}}.$$

Then the distribution function of Z_n converges to the distribution function of a standard normal random variable as $n\to\infty$. I.e., Z_n converges in distribution to a normally distributed random variable $X\sim\mathcal{N}(0,1)$.

Exercise 3.1.15. *Show that a MLE is asymptotically efficient.*

If $\hat\theta_n$ is a MLE, $\sqrt{n}(\hat\theta_n-\theta)\xrightarrow{\text{d}}\mathcal{N}(0,I_1^{-1}(\theta))$, which implies $\hat\theta_n\xrightarrow{\text{d}}\mathcal{N}(\theta,I_n^{-1}(\theta))$. Hence $\mathrm{var}(\hat\theta_n)\to I_n^{-1}$ and $\mathrm{eff}(\hat\theta_n)\to 1$.

We next wonder if a given estimator extracts all the information about the unknown parameter θ that is available in our samples. Assume we have observed a particular value of $\hat\theta$. In general, there would be various outcomes $\boldsymbol{x}=(x_1,\dots,x_n)$ that would lead to this particular estimate. If their distribution does not depend on θ, knowing which of them has specifically happened does not provide further information about the value of θ. This motivates the following definition.

Definition 3.1.22 *(Sufficient Statistic). Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be i.i.d. from a probability distribution with parameter θ. Then the statistic $T(\boldsymbol{X})$ is called a sufficient statistic for θ if the conditional distribution of $X_1, \ldots, X_n$ given the value of T does not depend on θ.*

There is no need to compute conditioned PDFs to check whether some statistic is sufficient thanks to the following

Theorem 3.1.23 *A statistic $T(\boldsymbol{X})$ is a sufficient statistic for θ if and only if the joint probability density of $\boldsymbol{X}$ can be factorised into two factors, one of which depends only on T and the parameters while the other is independent of the parameters:*

$$f(\boldsymbol{x}; \theta) = g(t; \theta) h(\boldsymbol{x}).$$

We do not prove this theorem here. The second factor may be written in terms of t, since it is a functions of the outcomes, but it cannot depend on θ.

Theorem 3.1.24 *Efficient estimators are sufficient.*

The converse is not true; there exist sufficient estimators/statistics that are not efficient.

Proof.
of Theorem 3.1.24 From the proof of Theorem 3.1.16 we know that if $\hat{\theta}$ is unbiased

$$\operatorname{cov}(W, \hat{\theta}) = 1.$$

If, moreover, $\hat{\theta}$ is efficient, $\operatorname{var}(\hat{\theta}) = 1/\operatorname{var}(W)$, hence

$$\begin{aligned}\mathbb{E}\left\{\left[W - \operatorname{var}(W)(\hat{\theta} - \theta)\right]^2\right\} &= \operatorname{var}(W) + \operatorname{var}(W)^2 \operatorname{var}(\hat{\theta}) - 2\operatorname{var}(W)\,\mathbb{E}[W(\hat{\theta} - \theta)] \\ &= 2\operatorname{var}(W) - 2\operatorname{var}(W)\operatorname{cov}(W, \hat{\theta}) = 0.\end{aligned}$$

Since $\mathbb{E}(X^2) = 0 \Rightarrow \Pr(X = 0) = 1$, it must be the case that

$$W = \operatorname{var}(W)(\hat{\theta} - \theta) := a(\theta)\hat{\theta} + b(\theta).$$

But, since

$$W = \frac{\partial}{\partial \theta} \log f(\boldsymbol{X}; \theta),$$

we see that

$$f(\boldsymbol{X}; \theta) = \exp\Big[A(\theta)\hat{\theta} + B(\theta) + C(\boldsymbol{X})\Big] = \exp\Big[A(\theta)\hat{\theta} + B(\theta)\Big] K(\boldsymbol{X}).$$

So, $f(\boldsymbol{x}; \theta) = \exp[A(\theta)\hat{\theta} + B(\theta)]K(\boldsymbol{x})$, and since according to Theorem 3.1.23 this is the required factorization for sufficientcy, $\hat{\theta}$ is sufficient. ■

Example 3.1.4. Suppose we use $\bar{x}$ to estimate λ, the parameter of the Poisson distribution

$$p(k; \lambda) = \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda}.$$

For a sample of size n we have

$$\begin{aligned}p(x_1, x_2, \ldots, x_n; \lambda) &= \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} = e^{-n\lambda}\lambda^{\sum_{i=1}^n x_i} \prod_{i=1}^{n} \frac{1}{x_i!} \\ &= \left(e^{-n\lambda}\lambda^{n\bar{x}}\right)\left(\prod_{i=1}^{n} \frac{1}{x_i!}\right),\end{aligned}$$

which is the required factorization according to Theorem 3.1.23. Hence $\bar{x}$ is sufficient.

Exercise 3.1.16. *Show that $(\bar{X}, S^2)$, defined in Example 3.1.2, is sufficient statistic to estimate the parameters μ and σ^2 of a normal distribution, Example 3.1.1.*

Theorem 3.1.25 *From any unbiased estimator that is not based on a sufficient statistic, an improved estimate can be obtained which is based on the sufficient statistic. It is unbiased and it has smaller variance, and is obtained by averaging with respect to the conditional distribution given the sufficient statistic.*

So, if $R(\boldsymbol{X})$ is an unbiased estimate of the parameter θ and $T(\boldsymbol{X})$ is a sufficient statistic for θ. The conditional distribution of R given T is

$$f_{R\,|\,T}(r\,|\,t) = \frac{f_{RT}(r,t;\theta)}{f_T(t;\theta)},$$

where $f_{RT}(r,t;\theta)$ is the joint probability density function of R and T, and

$$f_T(t;\theta) = \int_{-\infty}^{\infty} f(r,t;\theta)dr$$

is the marginal distribution for T. Because T is a sufficient statistic $f_{R\,|\,T}(r\,|\,t)$ does not depend on θ. The improved estimate of θ, $S(T)$, is the function of T that is obtained by averaging R with respect to its conditional distribution given T.

$$S(T) := \mathbb{E}[R\,|\,T] = \int_{-\infty}^{\infty} r f_{R\,|\,T}(r\,|\,T)dr.$$

Exercise 3.1.17. *Prove Theorem 3.1.25.*

Since R is an unbiased estimator of θ it satisfies

$$\mathbb{E}[R] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} r f_{RT}(r,t;\theta)dr\,dt = \theta.$$

Let us check that S is also unbiased:

$$\begin{aligned}\mathbb{E}(S) &= \int_{-\infty}^{\infty} s(t) f_T(t;\theta)\,dt = \int_{-\infty}^{\infty}\int_{\infty}^{\infty} r f_{R\,|\,T}(r,t) f_T(t;\theta)dr\,dt \\ &= \int_{-\infty}^{\infty}\int_{\infty}^{\infty} r f_{RT}(r,t;\theta)dr\,dt = \mathbb{E}(R) = \theta.\end{aligned}$$

It only remains to check that $\mathrm{var}(S) \le \mathrm{var}(R)$:

$$\begin{aligned}\mathrm{var}(R) &= \mathbb{E}\left[(R-\theta)^2\right] = \mathbb{E}\left\{[(S-\theta)+(R-S)]^2\right\} \\ &= \mathrm{var}(S) + \mathbb{E}\left[(R-S)^2\right] + 2\mathbb{E}[(R-S)(S-\theta)].\end{aligned} \tag{3.1.3}$$

Let us check that the last term is identically zero:

$$\begin{aligned}\mathbb{E}[(R-S)(S-\theta)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[r-s(t)][s(t)-\theta] f_{R,T}(r,t;\theta)dr\,dt \\ &= \int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty}[r-s(t)] f_{R\,|\,T}(r\,|\,t)dr\right\}[s(t)-\theta] f_T(t;\theta)dt.\end{aligned}$$

But the inner integral is zero. Since $\mathbb{E}[(R-S)^2] \ge 0$ we see from Eq. (3.1.3) that $\mathrm{var}(S) \le \mathrm{var}(R)$.

Exercise 3.1.18. *We toss a coin n times [head $= 1$, tail $= 0$]. We decide to use X_1 (the result of the first toss, ignoring the rest) as estimate of the probability p of a head, i.e., $\hat{\theta} = X_1$. Check $\hat{\theta}$ is unbiased. Compute its variance. Next, check that in n trials the proportion of heads is a sufficient statistic. Construct an improved estimate from $\hat{\theta}$ based on the proportion of heads in n trials using Theorem 3.1.25 and check explicitly that it has smaller variance.*

3.1.3 Bayesian approach

Within the Bayesian approach, the estimated parameter θ is assumed to be a random variable, Θ, that is distributed according to a prior PDF, $f(\theta)$, representing the knowledge about θ one possesses before performing the estimation. Therefore, in contrast to the frequentist philosophy, where the estimated parameter was assumed to have a fixed, well defined value, it is a particular realization of the parameter that is really estimated in a real-life experiment. As a consequence, an optimal estimator must not only be global and minimize the MSE, but also has to take into account which values of Θ are more probable according to $f(\theta)$. Hence, such an estimator must minimize the Average Mean Squared Error ($\overline{\text{MSE}}$):

$$\overline{\text{MSE}}(\hat{\theta}) = \int f(\theta) d\theta \int (\hat{\theta} - \theta)^2 f(\boldsymbol{x} \,|\, \theta) d^n x,$$

where we recall that the estimator $\hat{\theta}$ is a function of the sample $\boldsymbol{x} = (x_1, \ldots, x_n)$ and where $f(\boldsymbol{x} \,|\, \theta)$ is the PDF previously labelled as $f(\boldsymbol{x}; \theta)$ within the frequentist approach, which due to stochastic character of the parameter now represents a *conditional density*. The last definition can also be written as

$$\overline{\text{MES}}(\hat{\theta}) = \int (\hat{\theta} - \theta)^2 f(\boldsymbol{x}, \theta) d^n x \, d\theta,$$

where the joint PDF, $f(\boldsymbol{x}, \theta)$, is defined via Bayes' theorem –hence the name of the approach– in two equivalent ways:

$$f(\boldsymbol{x}, \theta) = f(\boldsymbol{x} \,|\, \theta) f(\theta) = f(\theta \,|\, \boldsymbol{x}) f(\boldsymbol{x})$$

[we abuse notation here by using the same letter f to denote all PDFs. In a more precise notation one should write $f_{\boldsymbol{X}\Theta}(\boldsymbol{x}, \theta) = f_{\boldsymbol{X}|\Theta}(\boldsymbol{x}|\theta) f_\Theta(\theta) = f_{\Theta|\boldsymbol{X}}(\theta|\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x})$, but we drop the subscripts to ease the notation] In general, the conditional PDFs satisfy $\int f(\boldsymbol{x} \,|\, \theta) d^n x = \int f(\theta \,|\, \boldsymbol{x}) d\theta = 1$ and the probability of a particular sample corresponds to the marginal $f(\boldsymbol{x}) = \int f(\boldsymbol{x}, \theta) d\theta$. We, then, can also write

$$\overline{\text{MSE}}(\hat{\theta}) = \int f(\boldsymbol{x}) d^n x \left[\int (\hat{\theta} - \theta)^2 f(\theta \,|\, \boldsymbol{x}) d\theta \right]. \tag{3.1.4}$$

The minimum of this expression is attained by minimizing the square bracket for each outcome $\boldsymbol{x}$:

$$\partial_{\hat{\theta}} \int (\hat{\theta} - \theta)^2 f(\theta \,|\, \boldsymbol{x}) d\theta = 0 \quad \Rightarrow \quad \hat{\theta} = \int \theta f(\theta \,|\, \boldsymbol{x}) d\theta = \mathbb{E}_{\Theta|\boldsymbol{X}}(\Theta). \tag{3.1.5}$$

The optimal *Minimum Mean Squared Error* (MMSE) estimator simply corresponds to the average parameter value computed with respect to the posterior PDF, $f(\theta \,|\, \boldsymbol{x})$, that in principle may always be computed using Bayes' theorem:

$$f(\theta \,|\, \boldsymbol{x}) = \frac{f(\boldsymbol{x} \,|\, \theta) f(\theta)}{\int f(\boldsymbol{x} \,|\, \theta) f(\theta) d\theta}. \tag{3.1.6}$$

Within the Bayesian framework, one should view the process of data inference as a procedure in which the effective PDF of the estimated parameter θ becomes updated. Hence, the posterior PDF $f(\theta \,|\, \boldsymbol{x})$ represents the prior $f(\theta)$ that has been reshaped and narrowed-down after learning the sample $\boldsymbol{x}$:

$$f(\theta) \quad \overset{\text{observe } \boldsymbol{x}}{\longrightarrow} \quad f(\theta \,|\, \boldsymbol{x})$$

whereas the MMSE estimator (3.1.5) just outputs the mean of such an effective distribution. Moreover, the minimal $\overline{\text{MSE}}$ (3.1.4) then reads

$$\overline{\text{MSE}}(\hat{\theta}) = \int f(\boldsymbol{x}) d^n x \left[\int f(\theta \,|\, \boldsymbol{x}) \left(\theta - \mathbb{E}_{\Theta|\boldsymbol{X}}(\Theta) \right)^2 \right] = \int f(\boldsymbol{x}) \, \text{var}_{\Theta|\boldsymbol{X}}(\Theta) d^n x,$$

so that it represents the variance of the parameter Θ computed also with respect to $f(\theta \,|\, \boldsymbol{x})$ and averaged over all the possible outcomes.

It is really important within the Bayesian approach to choose an appropriate $f(\theta)$ such that, on one hand, it adequately represents the knowledge about the parameter before the estimation, but, on the other, it does not significantly overshadow the information obtained from the data collected.

Exercise 3.1.19. *Consider the extremal case where the prior PDF is the Dirac delta distribution, $f_\delta(\theta) = \delta(\theta - \theta_0)$, which represents the case when we perfectly know the estimated parameter before*

performing the estimation. Compute the MMSE $\hat{\theta}$ and discuss the role of the observations. What is $\overline{\mathrm{MSE}}(\hat{\theta})$?

Note that so far we did not require at any stage the sampled data to be independently distributed. Such property, which previously was heavily used within the frequentist approach, is not necessary in the derivation of the optimal Bayesian estimator, which relies only on the form of the posterior PDF (3.1.6). In fact, as independently distributed data may be interpreted as if it was collected carrying out consecutive repetitions of the estimation protocol, the Bayesian results in such a case may be understood as a progressive updating of the knowledge we possess about the parameter, where at each step the posterior is calculated based only on the outcomes x_i but for the prior already updated with the results x_{i-1}.

$$f(\theta) \overset{\text{observe } x_1}{\longrightarrow} f(\theta \,|\, x_1) \overset{\text{observe } x_2}{\longrightarrow} f(\theta \,|\, x_1, x_2) \overset{\text{observe } x_3}{\longrightarrow} f(\theta \,|\, x_1, x_2, x_3) \cdots$$

Exercise 3.1.20. *Show that the interpretation of (3.1.6) for independent samples as progressive updating of the prior PDFs is correct.*

Consider the obvious relations [We use independency in the very first line: $f(x_1, \ldots, x_n \,|\, \theta) = \prod_{i=1}^{n} f(x_i \,|\, \theta) = f(x_n \,|\, \theta) f(x_1, \ldots, x_{n-1} \,|\, \theta)$]:

$$\begin{aligned} f(\theta \,|\, x_1, x_2, \ldots, x_n) &= \frac{f(x_n \,|\, \theta) f(x_1, \ldots, x_{n-1} \,|\, \theta) f(\theta)}{\int f(x_n \,|\, \theta) f(x_1, \ldots, x_{n-1} \,|\, \theta) f(\theta) d\theta} \\ &= \frac{f(x_n \,|\, \theta) f(\theta \,|\, x_1, \ldots, x_{n-1}) f(x_1, \ldots, x_n)}{\int f(x_n \,|\, \theta) f(\theta \,|\, x_1, \ldots, x_{n-1}) f(x_1, \ldots, x_n) d\theta} \\ &= \frac{f(x_n \,|\, \theta) f(\theta \,|\, x_1, \ldots, x_{n-1})}{\int f(x_n \,|\, \theta) f(\theta \,|\, x_1, \ldots, x_{n-1}) d\theta}. \end{aligned}$$

So, the updating in the last step from the prior PDF $f(\theta \,|\, x_1, \ldots, x_{n-1})$ is based entirely on the observation x_n. We can, obviously, repeat the procedure with $f(\theta \,|\, x_1, \ldots, x_{n-1})$, and so on.

The MMSE estimator plays and special role because of the following result. For any regular prior $f(\theta)$ one has

$$\overline{\mathrm{MSE}}(\hat{\theta}) \overset{n \to \infty}{=} \mathbb{E}_\Theta \left[\frac{1}{I_n(\Theta)} \right] \geq \frac{1}{\mathbb{E}_\Theta \left[I_n(\Theta) \right]}, \tag{3.1.7}$$

where the last expression follows from the Jensen inequality stating that for any convex function $f(X)$ one has $\mathbb{E}[f(X)] \geq f[\mathbb{E}(X)]$. If the Fisher information I_n is independent of θ we can, of course, drop the expectations and the bound saturates. This relation enables us to establish a connexion between the Bayesian and frequentist approaches.

Alternatively, one could prove the last inequality in (3.1.7) by invoking the Hlder inequality:

$$\int |g(x)h(x)| dx \leq \left(\int |g(x)|^p dx \right)^{1/p} \left(\int |h(x)|^q dx \right)^{1/q}, \quad \text{for all } p, q \text{ such that } \frac{1}{p} + \frac{1}{q} = 1.$$

For $g(x) = \sqrt{x f(x)}$, $h(x) = \sqrt{f(x)/x}$, $p = q = 2$, we have

$$1 = \int f(x) dx \leq \left(\int |x f(x) dx \right)^{1/2} \left(\int \frac{f(x)}{x} dx \right)^{1/2}, \quad \text{assuming } x > 0.$$

Therefore

$$1 \leq \mathbb{E}(X)\, \mathbb{E}(X^{-1}) \quad \Rightarrow \quad \mathbb{E}(X^{-1}) \geq \frac{1}{\mathbb{E}(X)}.$$

Within the Bayesian framework, nothing prevents us to consider other figures of merit, i.e., cost functions $C(\hat{\theta},\theta)$, in order to generalize the $\overline{\mathrm{MSE}}$, and define the average cost, $\mathbb{E}_\Theta[\mathcal{C}(\hat{\theta})]$:

$$\overline{\mathcal{C}}(\hat{\theta}) = \int f(\theta)d\theta \int C(\hat{\theta},\theta) f(x\,|\,\theta)dx.$$

The $\overline{\mathrm{MSE}}$ is the special case $C(\hat{\theta},\theta) = (\hat{\theta}-\theta)^2$.

Example 3.1.5. To deal with a circularly symmetric parameter, we can consider the simplest cost function introduced by Holevo:

$$C_{\mathrm{H}}(\hat{\theta},\theta) = C_{\mathrm{H}}(\hat{\theta}-\theta) = 4\sin^2\left(\frac{\hat{\theta}-\theta}{2}\right). \tag{3.1.8}$$

It is periodic (as it should if one has circular symmetry) and $C_{\mathrm{H}}(\hat{\theta},\theta) \sim (\hat{\theta}-\theta)^2$ (i.e., approaches the squared error) as $\hat{\theta}\to\theta$.

3.2 Quantum Estimation

3.2.1 Frequentist (pointwise) approach

The quantum Cramer-Rao bound

As we mentioned in the introduction, in quantum mechanics we have (Born rule) $f(x;\theta) = \mathrm{tr}\,(E_x\rho_\theta)$, where $\{E_x\}, \int dx E_x = \mathbb{1}$, are the elements of a POVM and ρ_θ is the density operator parametrized by the quantity we want to estimate. Let us introduce the

Definition 3.2.1 *[Symmetric Logarithmic Derivative (SLD)] The SLD, L_θ is the self-adjoint operator satisfying the equation*

$$\frac{L_\theta\rho_\theta + \rho_\theta L_\theta}{2} = \frac{\partial\rho_\theta}{\partial\theta} = \partial_\theta\rho_\theta.$$

Note that

$$\begin{aligned}
\partial_\theta f(x\,|\,\theta) &= \partial_\theta \,\mathrm{tr}\,[E_x\rho_\theta] = \mathrm{tr}\,[E_x\partial_\theta\rho_\theta] \\
&= \mathrm{tr}\left[E_x\left(\frac{L_\theta\rho_\theta+\rho_\theta L_\theta}{2}\right)\right] \\
&= \frac{1}{2}\,\mathrm{tr}\,[E_x L_\theta\rho_\theta] + \frac{1}{2}\,\mathrm{tr}\,[E_x\rho_\theta L_\theta] \\
&= \frac{1}{2}\,\mathrm{tr}\,[E_x L_\theta\rho_\theta] + \frac{1}{2}\left[\mathrm{tr}\,(E_x\rho_\theta L_\theta)^\dagger\right]^* \\
&= \frac{1}{2}\,\mathrm{tr}\,[E_x L_\theta\rho_\theta] + \frac{1}{2}\left[\mathrm{tr}\,L_\theta\rho_\theta E_x\right]^*,
\end{aligned}$$

where we have used the cyclic property of the trace. We can then write

$$\partial_\theta f(x\,|\,\theta) = \Re\left[\mathrm{tr}\,(\rho_\theta E_x L_\theta)\right].$$

We can use this result to express the Fisher information as

$$I_1(\theta) = \int dx \frac{\left\{\Re\left[\mathrm{tr}\,(\rho_\theta E_x L_\theta)\right]\right\}^2}{\mathrm{tr}(\rho_\theta E_x)}. \tag{3.2.9}$$

The numerator of the integrant can be bounded as

$$\left\{\Re\left[\mathrm{tr}\,(\rho_\theta E_x L_\theta)\right]\right\}^2 \le \left|\mathrm{tr}\,(\rho_\theta E_x L_\theta)\right|^2 = \left|\mathrm{tr}\left(\sqrt{\rho_\theta}\sqrt{E_x}\sqrt{E_x}L_\theta\sqrt{\rho_\theta}\right)\right|^2 \tag{3.2.10}$$

$$\le \mathrm{tr}\left(\sqrt{\rho_\theta}\sqrt{E_x}\sqrt{E_x}\sqrt{\rho_\theta}\right)\mathrm{tr}\left(\sqrt{E_x}L_\theta\sqrt{\rho_\theta}\sqrt{\rho_\theta}L_\theta\sqrt{E_x}\right) \tag{3.2.11}$$

$$= \mathrm{tr}(\rho_\theta E_x)\,\mathrm{tr}(L_\theta E_x L_\theta\rho_\theta),$$

where we have used the Schwartz inequality:

$$\left|\mathrm{tr}\left(A^\dagger B\right)\right|^2 \leq \mathrm{tr}\left(A^\dagger A\right)\mathrm{tr}\left(B^\dagger B\right). \tag{3.2.12}$$

We have also used that $\rho_\theta, E_x \geq 0$ and L_θ is self-adjoint. Substituting this bound in Eq. (3.2.9) we have

$$I_1(\theta) \leq \int dx\,\mathrm{tr}(L_\theta E_x L_\theta \rho_\theta) = \mathrm{tr}\left[L_\theta\left(\int dx E_x\right)L_\theta\rho_\theta\right] = \mathrm{tr}(L_\theta^2\rho_\theta), \tag{3.2.13}$$

which states that the Fisher information $I_1(\theta)$ of any quantum measurement is bounded by the so-called

Definition 3.2.2 *(Quantum Fisher Information QFI). The QFI is defined as*

$$H(\theta) := \mathrm{tr}(L_\theta^2\rho_\theta) = \mathrm{tr}\left[L_\theta\left(\partial_\theta\rho_\theta\right)\right].$$

The second form of the definition follows from

$$\mathrm{tr}(L_\theta^2\rho_\theta) = \frac{\mathrm{tr}(L_\theta\rho_\theta L_\theta) + \mathrm{tr}(L_\theta^2\rho_\theta)}{2} = \mathrm{tr}\{L_\theta\frac{\rho_\theta L_\theta + L_\theta\rho_\theta}{2}\} = \mathrm{tr}\left\{L_\theta\left(\partial_\theta\rho_\theta\right)\right\}.$$

Eq. (3.2.13), through Theorem 3.1.16, leads to

Theorem 3.2.3 *(Quantum Cramr-Rao bound) The variance of any estimator $\hat{\theta}_n$ of the parameter θ characterizing the family of states ρ_θ is bounded by*

$$\mathrm{var}(\hat{\theta}_n) \geq \frac{1}{nH(\theta)}.$$

This is the quantum version of the Cramr-Rao theorem and provides an ultimate bound the the sensitivity that can be achieved in parameter estimation in the quantum mechanical framework.

The quantum Fisher Information is an upper bound for the Fisher Information as it embodies the optimization of the Fisher Information over any possible measurement. Optimal quantum measurements for the estimation of θ thus correspond to POVM with Fisher information equal to the quantum Fisher information, i.e., those saturating both inequalities Eq. (3.2.10) and Eq. (3.2.11). The first one is saturated when $\mathrm{tr}\left[\rho_\theta E_x L_\theta\right]$ is a real number. The second one is based on the Schwartz inequality, Eq. (3.2.12), which is saturated when matrices A and B are proportional. Hence, we must have

$$\sqrt{E_x}\sqrt{\rho_\theta} = c_x\sqrt{E_x}L_\theta\sqrt{\rho_\theta}$$

for all x. This condition can always be met by choosing the operators E_x to be assembled from one-dimensional projectors onto a complete set of orthonormal eigenstates of L_θ.

Exercise 3.2.1. *Consider the set of qubit pure states that lie in the equator of the Bloch sphere ($\theta = \pi/2$). They are parametrized by the azimuthal angle ϕ. Compute the SLD, L_ϕ, for this one-parameter family. Compute the QFI, $H(\phi)$. Show that by measuring n copies of these equatorial states on the orthogonal bases $\{|+\rangle, |-\rangle\}$ and using the MLE estimator to process the classical data obtained from the measurement the QCR bound is saturated asymptotically.*

You should attempt to solve this exercise by yourself, but because of its relevance to a discussion about the Heisenberg limit below, we provide a solution here.

SOLUTION.

Let us compute the SLD brute force (we will learn about other methods below). Define the unit vectors $\hat{\boldsymbol{r}} = (\cos\phi, \sin\phi, 0)$ and $\hat{\boldsymbol{\phi}} = (-\sin\phi, \cos\phi, 0) = \partial_\phi\hat{\boldsymbol{r}}$, we have

$$\rho_\phi = \frac{\mathbb{1} + \hat{\boldsymbol{r}}\cdot\boldsymbol{\sigma}}{2} \quad \Rightarrow \quad \partial_\phi\rho_\phi = \frac{\hat{\boldsymbol{\phi}}\cdot\boldsymbol{\sigma}}{2}.$$

Write

$$L_\phi = a\mathbb{1} + \boldsymbol{b}\cdot\boldsymbol{\sigma}.$$

Then, using the definition of the SLD, we must have

$$\begin{aligned} 2\hat{\boldsymbol{\phi}}\cdot\boldsymbol{\sigma} &= (a\mathbb{1}+\boldsymbol{b}\cdot\boldsymbol{\sigma})(\mathbb{1}+\hat{\boldsymbol{r}}\cdot\boldsymbol{\sigma}) + (\mathbb{1}+\hat{\boldsymbol{r}}\cdot\boldsymbol{\sigma})(a\mathbb{1}+\boldsymbol{b}\cdot\boldsymbol{\sigma}) \\ &= 2a\mathbb{1} + 2(\boldsymbol{b}+a\hat{\boldsymbol{r}})\cdot\boldsymbol{\sigma} + \sum_{ij}(b_i\hat{r}_j + b_j\hat{r}_i)\sigma_i\sigma_j \\ &= 2(a+\boldsymbol{b}\cdot\hat{\boldsymbol{r}})\mathbb{1} + 2(\boldsymbol{b}+a\hat{\boldsymbol{r}})\cdot\boldsymbol{\sigma}. \end{aligned}$$

It follows that $a+\boldsymbol{b}\cdot\hat{\boldsymbol{r}}=0$ and $\boldsymbol{b}+a\hat{\boldsymbol{r}}=\hat{\boldsymbol{\phi}}$. Since $\hat{\boldsymbol{r}}\cdot\hat{\boldsymbol{\phi}}=0$, the second condition implies the first one and $\boldsymbol{b}=\hat{\boldsymbol{\phi}}-a\hat{\boldsymbol{r}}$ for any a. Hence, *the SLD is not uniquely defined*:

$$L_\phi = a\mathbb{1} + (\hat{\boldsymbol{\phi}}-a\hat{\boldsymbol{r}})\cdot\boldsymbol{\sigma} \quad \text{for any } a\in\mathbb{R}.$$

From the definition of QFI, we readily see that

$$H(\phi) = \operatorname{tr}\left\{\left[a\mathbb{1}+(\hat{\boldsymbol{\phi}}-a\hat{\boldsymbol{r}})\cdot\boldsymbol{\sigma}\right]\frac{\hat{\boldsymbol{\phi}}\cdot\boldsymbol{\sigma}}{2}\right\} = (\hat{\boldsymbol{\phi}}-a\hat{\boldsymbol{r}})\cdot\hat{\boldsymbol{\phi}} = 1.$$

The PMF of the measurement $\{|+\rangle,|-\rangle\}$ ($\phi=0,\pi$) is

$$\begin{aligned} p(+;\phi) &= \operatorname{tr}\left(|+\rangle\langle+|\rho_\phi\right) = |\langle+|\psi_\phi\rangle|^2 = \left|\frac{1+\mathrm{e}^{i\phi}}{2}\right|^2 \\ &= \cos^2\left(\frac{\phi}{2}\right), \\ p(-;\phi) &= \sin^2\left(\frac{\phi}{2}\right). \end{aligned}$$

So, we see that the outcome of each individual measurement is a Bernoulli random variable. By assigning 1 to outcome + and 0 to outcome −, the PMF of such variable is

$$X\sim p(x;\phi) = \cos^{2x}\left(\frac{\phi}{2}\right)\sin^{2(1-x)}\left(\frac{\phi}{2}\right).$$

The (log-)likelihood function is

$$L(\phi\,|\,\pm) = \frac{1\pm\cos\phi}{2}; \qquad l(\phi\,|\,\pm) = \log\left(\frac{1\pm\cos\phi}{2}\right).$$

Hence

$$\partial_\phi l(\phi\,|\,\pm) = \frac{\mp\sin\phi}{1\pm\cos\phi},$$

and

$$I_1(\phi) = \mathbb{E}\left[\left(\frac{\mp\sin\phi}{1\pm\cos\phi}\right)^2\right] = \frac{\sin^2\phi}{(1+\cos\phi)^2}\frac{1+\cos\phi}{2} + \frac{\sin^2\phi}{(1-\cos\phi)^2}\frac{1-\cos\phi}{2} = 1.$$

From which $I_n(\phi)=n$, and we know that the MLE of this measurement will saturate the bound asymptotically.

Measuring each individual copy of the given n states on the $\{|+\rangle,|-\rangle\}$ basis we have

$$p(x_1,\dots x_n;\phi) = \left[\cos^2\left(\frac{\phi}{2}\right)\right]^{n\bar{x}}\left[\sin^2\left(\frac{\phi}{2}\right)\right]^{n(1-\bar{x})}.$$

Checking the right hand side of this expression we clearly see that $\bar{X}$ is a sufficient statistic for ϕ. We also see that $\bar{X}$ is binomially distributed:

$$\bar{X}\sim \operatorname{Bin}\left[n,\cos^2\left(\frac{\phi}{2}\right)\right] \quad\Leftrightarrow\quad p_{\bar{X}}(\bar{x};\phi) = \binom{n}{n\bar{x}}\left[\cos^2\left(\frac{\phi}{2}\right)\right]^{n\bar{x}}\left[\sin^2\left(\frac{\phi}{2}\right)\right]^{n(1-\bar{x})}.$$

Hence, $L(\phi \mid \bar{x}) = p_{\bar{X}}(\bar{x}; \phi)$. A straightforward derivation leads to the expression of the MLE

$$\hat{\phi} = \arccos(1 - 2\bar{x}).$$

This completes the solution of the exercise. ■

Going back to our general discussion, one can find a closed form for the SLD in terms of the spectral representation of ρ_θ,

$$\rho_\theta = \sum_a \lambda_a |\psi_a\rangle\langle\psi_a|.$$

It is given by

$$L_\theta = \sum_{a,b} \frac{2\langle\psi_a|\partial_\theta\rho_\theta|\psi_b\rangle}{\lambda_a + \lambda_b} |\psi_a\rangle\langle\psi_b|, \tag{3.2.14}$$

where the sum extends to all a and b such that $\lambda_a + \lambda_b \neq 0$.

Exercise 3.2.2. *Prove Eq. (3.2.14).*

The pure state model

A much simpler expression can be written for pure states. A straightforward calculation gives

$$L_\theta = 2\partial_\theta\rho_\theta = 2\left(|\partial_\theta\psi_\theta\rangle\langle\psi_\theta| + |\psi_\theta\rangle\langle\partial_\theta\psi_\theta|\right). \tag{3.2.15}$$

From this result one can easily obtain the QFI of this so called pure state model:

$$H(\theta) = 4\left(\langle\partial_\theta\psi_\theta|\psi_\theta\rangle^2 + \langle\partial_\theta\psi_\theta|\partial_\theta\psi_\theta\rangle\right). \tag{3.2.16}$$

Exercise 3.2.3. *Prove that for pure states one can choose the SLD to be given by Eq. (3.2.15). By using this result, prove Eq. (3.2.16).*

Pure states satisfy $\rho_\theta^2 = \rho_\theta$, hence

$$\partial_\theta\rho_\theta = \partial_\theta\left(\rho_\theta^2\right) = (\partial_\theta\rho_\theta)\,\rho_\theta + \rho_\theta\,(\partial_\theta\rho_\theta).$$

By comparing with the definition of L_θ,

$$\partial\rho_\theta = \frac{L_\theta\rho_\theta + \rho_\theta L_\theta}{2},$$

we see that we can choose $L_\theta/2 = \partial_\theta\rho_\theta$, and

$$L_\theta = 2\partial_\theta\rho_\theta = 2\partial_\theta\left(|\psi_\theta\rangle\langle\psi_\theta|\right) = 2\left(|\partial_\theta\psi_\theta\rangle\langle\psi_\theta| + |\psi_\theta\rangle\langle\partial_\theta\psi_\theta|\right),$$

where, assuming $\{|\alpha\rangle\}$ is a fixed (θ-independent) basis of the Hilbert space,

$$|\partial_\theta\psi_\theta\rangle = \partial_\theta\left(\sum_\alpha \psi_\theta^\alpha|\alpha\rangle\right) = \sum_\alpha (\partial_\theta\psi_\theta^\alpha)\,|\alpha\rangle$$

and

$$\langle\partial_\theta\psi_\theta| = \partial_\theta\left(\sum_\alpha (\psi_\theta^\alpha)^*\langle\alpha|\right) = \sum_\alpha (\partial_\theta\psi_\theta^\alpha)^*\,\langle\alpha|.$$

The QFI is easily derived noticing that for pure states one can write

$$H(\theta) = \mathrm{tr}\left[L_\theta\,(\partial_\theta\rho_\theta)\right] = \frac{1}{2}\,\mathrm{tr}\left(L_\theta^2\right).$$

Then,

$$\begin{aligned} H(\theta) &= 2\,\mathrm{tr}\left[\left(|\partial_\theta\psi_\theta\rangle\langle\psi_\theta| + |\psi_\theta\rangle\langle\partial_\theta\psi_\theta|\right)\left(|\partial_\theta\psi_\theta\rangle\langle\psi_\theta| + |\psi_\theta\rangle\langle\partial_\theta\psi_\theta|\right)\right] \\ &= 2\left[\langle\psi_\theta|\partial_\theta\psi_\theta\rangle^2 + \langle\partial_\theta\psi_\theta|\psi_\theta\rangle^2 + 2\langle\partial_\theta\psi_\theta|\partial_\theta\psi_\theta\rangle\right] \\ &= 4\left[\langle\partial_\theta\psi_\theta|\psi_\theta\rangle^2 + \langle\partial_\theta\psi_\theta|\partial_\theta\psi_\theta\rangle\right], \end{aligned}$$

where we have used that

$$\langle\psi_\theta|\partial_\theta\psi_\theta\rangle = -\langle\partial_\theta\psi_\theta|\psi_\theta\rangle,$$

which follows from taking derivative of $\langle\psi_\theta|\psi_\theta\rangle = 1$.

Let us consider the case where the parameter of interest, θ, is the amplitude of a unitary perturbation imprinted to a given initial pure state $|\psi_0\rangle$. The family of quantum states we are dealing with may be expressed as

$$|\psi_\theta\rangle = U_\theta|\psi_0\rangle,$$

where $U_\theta = \exp\{-i\theta H\}$ is a unitary operator and H is the corresponding Hermitian generator (we may think of it as the "Hamiltonian" of the system). This example is of particular interest in metrology.

From Eq. (3.2.16) the QFI can be easily computed to be

$$H(\theta) = 4(\Delta H)^2_{\psi_\theta} = 4(\Delta H)^2_{\psi_0}, \tag{3.2.17}$$

where $(\Delta H)_\psi$ is the standard deviation (uncertainly) of the hermitian operator H in the state $|\psi\rangle$, defined through

$$(\Delta H)^2_\psi = \langle\psi|H^2|\psi\rangle - \langle\psi|H|\psi\rangle^2$$

(it is just the variance in the quantum mechanical sense). The QCR bound is then

$$\mathrm{var}(\hat{\theta}) \geq \frac{1}{4n(\Delta H)^2_{\psi_0}},$$

and we note that it is *independent of* θ, hence providing a global bound.

Exercise 3.2.4. *Derive Eq. (3.2.17).*

$$|\partial_\theta\psi_\theta\rangle = \partial_\theta U_\theta|\psi_0\rangle = -iHU_\theta|\psi_0\rangle = -iH|\psi_\theta\rangle.$$

$$\begin{aligned} \langle\partial_\theta\psi_\theta|\psi_\theta\rangle &= i\langle\psi_\theta|H|\psi_\theta\rangle = i\langle\psi_0|H|\psi_0\rangle, \\ \langle\partial_\theta\psi_\theta|\partial_\theta\psi_\theta\rangle &= \langle\psi_\theta|H^2|\psi_\theta\rangle = \langle\psi_0|H^2|\psi_0\rangle. \end{aligned}$$

Substituting these expressions in Eq. (3.2.16) we obtain the desired result

$$H(\theta) = 4\left[\left(i\langle\psi_\theta|H|\psi_\theta\rangle\right)^2 + \langle\psi_\theta|H^2|\psi_\theta\rangle\right] = 4\,(\Delta H)^2_{\psi_\theta} = 4\,(\Delta H)^2_{\psi_0}\,.$$

The lower bound we have obtained is a function of the reference state. We should now find what is the best state $|\psi_0\rangle$ to estimate θ. We will prove that

Claim 3.2.4 *The maximum value that $(\Delta H)_\psi$ can achieve is half of the so called* spread of H*, namely,*

$$(\Delta H)_\psi \leq \frac{|\lambda_{\max} - \lambda_{\min}|}{2}.$$

This value is attainable with the choice

$$|\psi_0\rangle = \frac{|\lambda_{\min}\rangle + |\lambda_{\max}\rangle}{\sqrt{2}}, \tag{3.2.18}$$

where $|\lambda_{\min}\rangle$ ($|\lambda_{\max}\rangle$) is the eigenstate of the minimum (maximum) eigenvalue, $\lambda_{\min}$, ($\lambda_{\max}$) of H.

Then, the QCR bound for this family of states reads

$$\mathrm{var}(\hat{\theta}) \geq \frac{1}{n(\lambda_{\max} - \lambda_{\min})^2}. \tag{3.2.19}$$

PROOF.
of the claim Let the spectral decomposition of H be given by

$$H = \sum_a \lambda_a |\lambda_a\rangle\langle\lambda_a|.$$

A generic state $|\psi\rangle$ can be written in this eigenbasis as

$$|\psi\rangle = \sum_a \psi_a |\lambda_a\rangle.$$

Then,

$$(\Delta H)^2_\psi = \sum_{a=0} \lambda_a^2 p_a - \left(\sum_a \lambda_a p_a\right)^2, \tag{3.2.20}$$

where $p_a := |\psi_a|^2 \geq 0$ are (of course!) probabilities, $\sum_a p_a = 1$. Instead H let us consider the operator

$$\tilde{H} = \frac{H - \lambda_{\min}\mathbb{1}}{\lambda_{\max} - \lambda_{\min}},$$

whose minim eigenvalue, $\tilde{\lambda}_{\min}$, is zero and maximum eigenvalue, $\tilde{\lambda}_{\max}$, is one. We obviously have

$$(\Delta H)^2_\psi = (\lambda_{\max} - \lambda_{\min})^2 \left(\Delta \tilde{H}\right)^2_\psi. \tag{3.2.21}$$

Let us now show that $(\Delta \tilde{H})^2_\psi = 1/4$. Eq. (3.2.20) is completely general, so it also holds for $\tilde{H}$, but $\tilde{\lambda}_a^2 \leq \tilde{\lambda}_a$. Hence, if we define $u := \sum_a \tilde{\lambda}_a p_a \geq 0$, we have

$$\left(\Delta \tilde{H}\right)^2_\psi \leq u - u^2, \quad \text{for all } u \geq 0. \tag{3.2.22}$$

The maximum value of the right hand side is $1/4$ (for $u = 1/2$). Substituting in (3.2.21) we obtained the desired bound:

$$(\Delta H)^2_\psi = \left(\frac{\lambda_{\max} - \lambda_{\min}}{2}\right)^2.$$

The bound (3.2.22) is attained iff $p_a = 0$ for all $\tilde{\lambda}_a$ with the exception of $\tilde{\lambda}_b = 1 = \tilde{\lambda}_{\max}$ and $\tilde{\lambda}_c = 0 = \tilde{\lambda}_{\min}$ (since it follows from the inequality $\tilde{\lambda}_a^2 \leq \tilde{\lambda}_a$). Thus, any state of the form

$$|\psi\rangle = \frac{|\tilde{\lambda}_{\min}\rangle + \mathrm{e}^{iw}|\tilde{\lambda}_{\max}\rangle}{\sqrt{2}} = \frac{|\lambda_{\min}\rangle + \mathrm{e}^{iw}|\lambda_{\max}\rangle}{\sqrt{2}}.$$

attains the maximum. In particular we can choose $w = 0$. ■

The Heisenberg limit

Now we are going to show something remarkable. Let us go back to Exercise 3.2.1. The state ρ_ϕ is pure, so $\rho_\phi = |\psi_\phi\rangle\langle\psi_\phi|$, and

$$|\psi_\phi\rangle = \frac{|0\rangle + \mathrm{e}^{-i\phi}|1\rangle}{\sqrt{2}} = U_\phi|\psi_0\rangle,$$

where

$$U_\phi = \mathrm{e}^{-i\phi H} = |0\rangle\langle 0| + \mathrm{e}^{-i\phi}|1\rangle\langle 1|, \quad |\psi_0\rangle = |+\rangle, \quad H = |1\rangle\langle 1| \quad \Rightarrow \quad \begin{cases} \lambda_{\min} = 0 \\ \lambda_{\max} = 1 \end{cases}.$$

Using the QCR bound derived in Eq. (3.2.19) we recover the result of the exercise, namely

$$\mathrm{var}(\hat{\phi}) \geq \frac{1}{n(1-0)^2} = \frac{1}{n}.$$

and this bound is attainable. If we repeat the same experiment (measuring n copies to estimate ϕ) N times the variance would, of course, be bounded as

$$\mathrm{var}(\hat{\phi}) \geq \frac{1}{Nn}. \tag{3.2.23}$$

Suppose, however, that we proceed in a different way. Instead of preparing a product state, $|\psi_0\rangle^{\otimes n}$, and measuring each copy separately, we view the n copies as a whole system, S, with "Hamiltonian" $H_S = \sum_{k=1}^{n} H_k$, where $H_k = |1\rangle_k\langle 1| \otimes \mathbb{1}_{S-\{k\}}$ is the "Hamiltonian" of the k-th qubit, and choose the fiducial state as in (3.2.18). In this case

$$|\psi_0\rangle = \frac{|\lambda_{\min}\rangle + |\lambda_{\max}\rangle}{\sqrt{2}} = \frac{|0\rangle^{\otimes n} + |1\rangle^{\otimes n}}{\sqrt{2}} \quad \Rightarrow \quad \begin{cases} \lambda_{\min} &= 0 \\ \lambda_{\max} &= n \end{cases}$$

(note it is a highly entangled –hence very fragile– state!). Then, according to (3.2.19), if we repeat the experiment N times we will get a much enhanced sensitivity, with a variance scaling quadratically with the inverse of the size of the system

$$\mathrm{var}(\hat{\theta}) \geq \frac{1}{N(n-0)^2} = \frac{1}{Nn^2}.$$

We refer to this behavior, $\mathrm{var}(\hat{\theta}) \sim 1/n^2$, as the *Heisenberg limit*, in contrast to the *standard quantum (or shot-noise) limit*, where $\mathrm{var}(\hat{\theta}) \sim 1/n$ [as in Eq. (3.2.23)]. In quantum metrology based on interferometry, the Heisenberg limit can be achieved using squeezed states (instead of the "classical" coherent states that have shot-noise limited sensitivity). This falls beyond the scope of this course and will not be discussed here.

3.2.2 Bayesian approach

To give a flavor of what the Bayesian approach is about, let us look at the pure state model of Section 3.2.1 from this point of view. We will use the cost function (3.1.8) introduced in Example 3.1.5. The averaged cost function is

$$\overline{\mathcal{C}}_{\mathrm{H}}(\hat{\phi}) = \int_0^{2\pi} \frac{d\phi}{2\pi} \int 4\sin^2\left(\frac{\hat{\phi}_{\boldsymbol{x}} - \phi}{2}\right) \mathrm{tr}\Big(E_{\boldsymbol{x}} U_\phi |\psi_0\rangle\langle\psi_0| U_\phi^\dagger\Big)\, d^m x,$$

where $|\psi_0\rangle \in (\mathbb{C}^2)^{\otimes n}$ and we emphasize that the estimate $\hat{\phi}$ depends on the outcomes $\boldsymbol{x}$ by writing $\hat{\phi}_{\boldsymbol{x}}$. We assume, without loss of generality, that the outcomes, $\boldsymbol{x}$, of the measurement are continuous random variables (vectors of some dimension m; the "volume element" $d^m x$ is normalized such that $\int d^m x = 1$). This can be witten as

$$\begin{aligned}
\overline{\mathcal{C}}_{\mathrm{H}}(\hat{\phi}) &= 2\int_0^{2\pi} \frac{d\phi}{2\pi} \int \Big(1 - \cos(\hat{\phi}_{\boldsymbol{x}} - \phi)\Big)\, \langle\psi_0| U_\phi^\dagger E_{\boldsymbol{x}} U_\phi |\psi_0\rangle d^m x \\
&= 2 - \Re\left\{\int_0^{2\pi} \frac{d\phi}{2\pi} \int \mathrm{e}^{i\hat{\phi}_{\boldsymbol{x}}} \mathrm{e}^{-i\phi} \langle\psi_0| U_\phi^\dagger E_{\boldsymbol{x}} U_\phi |\psi_0\rangle d^m x\right\}.
\end{aligned}$$

The unitary matrix U_ϕ acts on $(\mathbb{C}^2)^{\otimes n}$, so it can be written as

$$U_\phi = \sum_{k=0}^{n} \mathrm{e}^{-ik\phi} |k\rangle\langle k|, \tag{3.2.24}$$

where each $|k\rangle$ spans the (one-dimentional) irreducible representations of the unitary group $U(1)$. Explicitly they are

$$|k\rangle = \binom{n}{k}^{-1/2} \mathrm{e}^{i\varphi_k} \Big(|0\rangle \otimes \cdots \otimes |0\rangle \otimes \underbrace{|1\rangle \otimes \cdots \otimes |1\rangle}_{k} + \text{permutations}\Big). \tag{3.2.25}$$

The normalization coefficient comes about because there are $\binom{n}{k}$ different orthogonal terms on the right hand side and the phases φ_k are arbitrary; we can choose as we wish. We now note that

$$\int_0^{2\pi} \frac{d\phi}{2\pi} \mathrm{e}^{-i\phi} U_\phi^\dagger \otimes U_\phi = \sum_{k=0}^{n} |k+1\rangle\langle k+1| \otimes |k\rangle\langle k|. \tag{3.2.26}$$

This is so because $\int_0^{2\pi} d\phi/(2\pi) \exp(is\phi) = 0$ for any $s \in \mathbb{Z}$, $s \neq 0$. Using this property, we readily see that

$$\overline{\mathcal{C}}_{\mathrm{H}}(\hat{\phi}) = 2 - 2\Re\left\{ \sum_{k=0}^{n} \int \mathrm{e}^{i\hat{\phi}_{\boldsymbol{x}}} c_{k+1} c_k (E_{\boldsymbol{x}})_{k+1,k}\, d^m x \right\}. \tag{3.2.27}$$

Here we have introduced the definitions $(E_{\boldsymbol{x}})_{k+1,k} := \langle k+1|E_{\boldsymbol{x}}|k\rangle$ and $c_k := \langle k|\psi_0\rangle$, where the arbitrary phases φ_k in (3.2.25) have been chosen so that $c_k \geq 0$. We know that these phases have no physical relevance, so this choice cannot affect our result. Now, note the following chain of inequalities:

$$\begin{aligned}
\overline{\mathcal{C}}_{\mathrm{H}}(\hat{\phi}) &\geq 2 - 2\left| \sum_{k=0}^{n} \int \mathrm{e}^{i\hat{\phi}_{\boldsymbol{x}}} c_{k+1} c_k (E_{\boldsymbol{x}})_{k+1,k}\, d^m x \right| \\
&\geq 2 - 2\sum_{k=0}^{n} \int c_{k+1} c_k \left| (E_{\boldsymbol{x}})_{k+1,k} \right| d^m x \\
&\geq 2 - 2\sum_{k=0}^{n} c_{k+1} c_k \int \sqrt{(E_{\boldsymbol{x}})_{k,k}} \sqrt{(E_{\boldsymbol{x}})_{k+1,k+1}}\, d^m x \\
&\geq 2 - 2\sum_{k=0}^{n} c_{k+1} c_k \sqrt{\left(\int (E_{\boldsymbol{x}})_{k,k}\, d^m x \right) \left(\int (E_{\boldsymbol{x}})_{k+1,k+1}\, d^m x \right)} \\
&= 2 - 2\sum_{k=0}^{n} c_{k+1} c_k .
\end{aligned} \tag{3.2.28}$$

In the first line we have used that $|\Re(z)| \leq |z|$ for any $z \in \mathbb{C}$. The triangle inequality led to the second line. The positivity condition $E_{\boldsymbol{x}} \geq 0$ implies $|(E_{\boldsymbol{x}})_{k+1,k}|^2 \leq (E_{\boldsymbol{x}})_{k,k}(E_{\boldsymbol{x}})_{k+1,k+1}$, $(E_{\boldsymbol{x}})_{k,k} \geq 0$, $(E_{\boldsymbol{x}})_{k+1,k+1} \geq 0$, which enabled us to write the third inequality. Schwarz inequality led to the forth line, and finally, the POVM condition $\int E_{\boldsymbol{x}}\, d^m x = \mathbb{1}$ enabled us to get rid of the POVM operators in the last line and got an absolute bound. Attainability is shown by providing an explicit measurement that saturates the bound, e.g.,

$$E_{\hat{\phi}} = U_{\hat{\phi}} |\Phi_0\rangle\langle\Phi_0| U_{\hat{\phi}}^\dagger,$$

where

$$|\Phi_0\rangle = \sum_{k=0}^{n} |k\rangle$$

(note $|\Phi_0\rangle$ is not normalized).

Exercise 3.2.5. *Check that the set of operators $\{E_{\hat{\phi}} \mid \hat{\phi} \in [0, 2\pi)\}$ defines a (continuous) POVM. Show that it saturates the bound in Eq. (3.2.28).*

The operators $E_{\hat{\phi}}$ are manifestly positive, so we only need to check that they add up to the identity operator:

$$\int \frac{d\hat{\phi}}{2\pi} E_{\hat{\phi}} = \int \frac{d\hat{\phi}}{2\pi} U_{\hat{\phi}} |\Phi_0\rangle\langle\Phi_0| U_{\hat{\phi}}^\dagger.$$

From Eq. (3.2.24) we have

$$\int_0^{2\pi} \frac{d\hat{\phi}}{2\pi} U_{\hat{\phi}} \otimes U_{\hat{\phi}}^\dagger = \sum_{k=0}^{n} \sum_{l=0}^{n} |k\rangle\langle k| \otimes |l\rangle\langle l| \int_0^{2\pi} \mathrm{e}^{-i(k-l)\hat{\phi}} \frac{d\hat{\phi}}{2\pi} = \sum_{k=0}^{n} |k\rangle\langle k| \otimes |k\rangle\langle k|,$$

hence

$$\int \frac{d\hat{\phi}}{2\pi} E_{\hat{\phi}} = \int \frac{d\hat{\phi}}{2\pi} U_{\hat{\phi}} |\Phi_0\rangle\langle\Phi_0| U_{\hat{\phi}}^\dagger = \sum_{k=0}^{n} |\langle k|\Phi_0\rangle|^2 |k\rangle\langle k|.$$

But

$$\langle k|\Phi_0\rangle = \langle k|\left(\sum_{l=0}^{n}|l\rangle\right) = \langle k|k\rangle = 1.$$

Substituting this in the previous equation we get

$$\int \frac{d\hat{\phi}}{2\pi} E_{\hat{\phi}} = \sum_{k=0}^{n} |k\rangle\langle k| = \mathbb{1}.$$

So the set $\{E_{\hat{\phi}} \mid \hat{\phi} \in [0, 2\pi)\}$ defines a proper POVM. Let us now check that it saturates de bound. For our POVM Eq. (3.2.27) can be written as

$$\overline{\mathcal{C}}_{\mathrm{H}}(\hat{\phi}) = 2 - 2\Re\left\{\sum_{k=0}^{n} c_{k+1}c_k \int_0^{2\pi} \frac{d\hat{\phi}}{2\pi}\, \mathrm{e}^{i\hat{\phi}}(E_{\hat{\phi}})_{k+1,k}\right\}.$$

Let us compute the integral:

$$\begin{aligned}\int_0^{2\pi} \frac{d\hat{\phi}}{2\pi}\mathrm{e}^{i\hat{\phi}}(E_{\hat{\phi}})_{k+1,k} &= \langle k+1|\left(\int_0^{2\pi} \frac{d\hat{\phi}}{2\pi}\mathrm{e}^{i\hat{\phi}}U_{\hat{\phi}}|\Phi_0\rangle\langle\Phi_0|U_{\hat{\phi}}^{\dagger}\right)|k\rangle \\ &= \sum_l \langle k+1|\left(\sum_{l=0}^{n}|l+1\rangle\langle l+1|\Phi_0\rangle\langle\Phi_0|l\rangle\langle l|\right)|k\rangle \\ &= \langle k+1|\Phi_0\rangle\langle\Phi_0|k\rangle = 1\end{aligned}$$

[we have used the hermitian conjugate of Eq. (3.2.26) in the second line]. Hence

$$\overline{\mathcal{C}}_{\mathrm{H}}(\hat{\phi}) = 2 - 2\Re\left\{\sum_{k=0}^{n} c_{k+1}c_k\right\} = 2 - 2\sum_{k=0}^{n} c_{k+1}c_k.$$

Our last task is to minimize the cost function $\overline{\mathcal{C}}_{\mathrm{H}}(\hat{\phi})$. Since c_k are the components of the reference state $|\psi_0\rangle$ in the basis $\{|k\rangle\}_{k=0}^{n}$ and they are real (because of our choice of phases) we have $\sum_{k=0}^{n} c_k^2 = 1$. So,

$$\overline{\mathcal{C}}_{\mathrm{H}}(\hat{\phi}) \geq \begin{pmatrix} c_0 & c_1 & \dots & c_n \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 2 & -1 & \ddots & 0 & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix}. \tag{3.2.29}$$

The minimum of this quadratic form is given by the minimum eigenvalue of the symmetric matrix on the right hand side.

Exercise 3.2.6. *Let M be a $n \times n$ symmetric matrix with entries M_{ij} and $\{c_k\}_{k=1}^{n}$ real coefficients so that $\sum_{k=1}^{n} c_k^2 = 1$. Show the the minimum of $C = \sum_{i,j=1}^{n} c_j M_{jk} c_k$ is the minimum eigenvalue of M.*

Diagonalizing the matrix in (3.2.29) is an easy task that surely enough you have accomplished in your classical mechanics courses when dealing with coupled oscillators, as it shows up when a chain of $n+1$ equal masses are connected with springs of equal strength (the loaded string). The eigenvalues are there computed to be[1]

$$\lambda_j = 4\sin^2\left[\frac{j\pi}{2(n+2)}\right], \quad j = 1, 2, \dots, n+1.$$

By borrowing this result we obtain that the minimum cost is

$$\overline{\mathcal{C}}_{\mathrm{H}}^{\min}(\hat{\phi}) = 4\sin^2\left[\frac{\pi}{2(n+2)}\right].$$

[1] See, e.g., *Classical Dynamics of particles and systems* (fifth edition), Stephen T. Thornton and Jerry B. Marion, Brooks/Cole (2004)

Exercise 3.2.7. *Find the optimal reference state $|\psi_0\rangle$. Namely, compute the coefficients c_k for which the minimum cost is attained.*

Quite remarkably this average cost is exact for all n! One can easily verify that

$$\overline{\mathcal{C}}_{\mathrm{H}}^{\min}(\hat{\phi}) \sim \frac{\pi^2}{n^2} \quad \text{as } n \to \infty.$$

Since in average $\hat{\phi}$ is very close to ϕ asymptotically, we have

$$\overline{\mathrm{MSE}}(\hat{\theta}) \approx \overline{\mathcal{C}}_{\mathrm{H}}^{\min}(\hat{\phi}) \sim \frac{\pi^2}{n^2} \quad \text{as } n \to \infty.$$

This tell as that the protocol consisting in preparing a system of n qubits in the state $|\psi_0\rangle$ and performing the measurement defined by $\{E_{\hat{\phi}} \mid 0 \leq \hat{\phi} < 2\pi\}$, whose outcome is the estimate, $\hat{\phi}$, has Heisenberg-limited sensitivity. The factor π^2 is the price we pay for using a global estimator.

Bibliography

[1] *Probability and Statistics II*. Notes by G. Deligiannidis.

[2] *Quantum estimation for quantum technology*, M. G. A. Paris, arXiv:0804.2981v3

[3] *Precision bounds in noisy quantum metrology*, J. Kolodynski, arXiv:1409.0535v2

[4] *Nonlinear quantum metrology*, S. Boixo, PhD thesis, University of New Mexico (2008).

Chapter 4

Quantum Estimation
Part II: Quantum state tomography

4.1 Introduction

Quantum state tomography refers to the experimental task of estimating the unknown $d \times d$ density matrix of a quantum system, ρ. It is thus a multiparameter estimation task, where nothing is assumed about ρ except that it corresponds to a physical state, i.e., it is a valid density matrix. To be able to carry out such task, we are assumed to have the ability to prepare and measure N copies of the system in the same state, $\rho^{\otimes N}$.

A closely related task is *quantum process tomography*, where one must determine the Choi matrix J_Λ associated to a quantum process $\Lambda : \rho \mapsto \Lambda(\rho)$. Here one has N uses of the process Λ, and has the liberty of preparing any input or "probe" state ρ for each use, including a d^N-dimensional multipartite entangled state where Λ acts locally on each party.

In this lecture we will focus on quantum state tomography. This task is acknowledged to be vital for implementing any quantum technology and, at the same time, extremely challenging to carry out in practice. The main reason for its difficulty is rooted in the so-called *curse of dimensionality*: the Hilbert space of a multipartite system grows exponentially in the local dimension, and with it so does the number of parameters to estimate in an unknown quantum state of such system. Not only the number of state preparations and measurements required to fully characterize a, say, 53-qubit state (the size of Google's Sycamore quantum processor that achieved quantum computational advantage for the first time [24]) up to a reasonably small error, but also the computational resources to compute an estimate of it, are enormous. As we will see later, this apparent roadblock has spurred considerable interest in alternative methods that are less demanding than full tomography.

4.1.1 The unknown quantum state

Before commencing, let us (superficially) explore a deeper, more philosophical side of tomography that has to do with the notion of an *unknown* quantum state. We will roughly follow the discussion on the topic in Ref. [25]. Tomography, like any statistical inference task, can be undertaken using a frequentist or a Bayesian approach. In this section we discuss the implications of the formulation of the tomography problem under a Bayesian perspective, since an apparent contradiction arises at a conceptual level, before writing any equations.

Following an ontological interpretation of quantum states, a quantum state to be characterized via tomography is unknown in the sense that we, as experimentalists, do not have a classical description for it (a density matrix written in a piece of paper), but the state itself surely exists independently of whether we observe it or not. This view is in line with how frequentists define a probability: the asymptotic limit to which measurable frequencies tend. Like a probability, a quantum state is of ontic nature, it is a property of nature that exists independently of us. An unknown quantum state is therefore a well defined concept.

Under an epistemological approach, however, a quantum state is a state of belief of some agent that quantifies their degree of knowledge about the state of the system. This interpretation of a quantum state is very much in line with the Bayesian interpretation of probability. What does it mean, then, for a quantum state to be unknown? If it is unknown to us experimentalists, who holds this state of belief that we are trying to access? The usual Bayesian conceptual assumption is that there is uncertainty on the *preparation procedure* of a given quantum state from our point of view as experimentalists-measurers, but the experimentalist that carried out the preparation does hold the knowledge about what state they are preparing. This is a well motivated framework in several scenarios involving various agents, e.g., in quantum communication. But tomography, as an attempt to characterize a completely *unknown* quantum state that not even the builder of the source producing it knows, does not quite fit this assumption.

Luckily for Bayesian statisticians, tomography can be rigorously formulated without the need of involving an external agent holding the state of belief that we try to unravel, nor does it need the term "unknown state" at all. The only assumption that the experimentalist needs to make is that the states produced by the source are indeed indistinguishable from each other, and nothing else. That is, if $\rho^{(N)}$

comprises their overall state of knowledge of the N copies prior to any measurement, they will assign the same state $\rho^{(N)}$ to any permutation of the copies[1]. The key point now is the *quantum de Finetti representation theorem*, which states that, if $\rho^{(N)}$ corresponds to an exchangeable sequence of states—in the above sense—, it can be expressed as

$$\rho^{(N)} = \int P(\rho)\rho^{\otimes N} d\rho \,, \tag{4.1.1}$$

where $P(\rho)$ is a prior probability distribution for ρ. Plainly put into words, the experimentalist can regard their prior state of knowledge $\rho^{(N)}$ *as if* it were a probabilistic mixture of tensor product states $\rho^{\otimes N}$, where all they can tell about the unknown state is encapsulated in the probability distribution $P(\rho)$.

Upon obtaining information from measuring some of the copies, $P(\rho)$ is shaped accordingly by means of Bayes' rule, until, when enough copies are measured, the state of knowledge of the experimentalist for the remaining copies resembles a product state. Furthermore, this updating process guarantees that two independent agents would come to agreement based on the same measurement outcomes, regardless the prior probability with which each one starts. Say the initial overall state of $N + M$ copies produced by the source is

$$\rho^{(N+M)} = \int P(\rho)\rho^{\otimes (N+M)} d\rho \,. \tag{4.1.2}$$

The first N copies are measured, and the obtained information is represented by the (multidimensional) random variable k. It can be shown that the remaining M copies are left in the post-measurement state

$$\rho_k^{(M)} = \int P(\rho|k)\rho^{\otimes M} d\rho \,, \tag{4.1.3}$$

where $P(\rho|k)$ is calculated through Bayes' rule. When N is large enough, the probability $P(\rho|k)$ gets highly peaked on a certain state ρ_k determined by the measurement outcomes, independently of the prior probability $P(\rho)$. If two agents start with different priors $P_i(\rho)$, $i = 1, 2$, both will adjust their state of knowledge after the measurement to the same product state $\rho_k^{\otimes M}$, as $\int P_i(\rho|k)\rho^{\otimes M} d\rho \to \rho_k^{\otimes M}$ for N sufficiently large.

The Bayesian interpretation of the tomography process shifts in this way the focus from accessing the "true" state of the system to agents agreeing to a common state of knowledge in the light of evidence. Bayesian theory does not describe how the physical world behaves, but rather how us, observers of that world, should act if we want to make rational assessments about it.

4.2 Measurement schemes and data acquisition

The task of quantum state tomography begins by preparing N copies of ρ, which we then measure to gather information and infer what is ρ. We first need to choose a *measurement scheme*, that is, a particular strategy to measure each copy of ρ individually that is able to provide enough information to reconstruct ρ. Associated to a measurement scheme, there is a tomographic map

$$\mathcal{T} : \rho \mapsto \boldsymbol{p} \tag{4.2.4}$$

that takes states into vectors of probabilities $\boldsymbol{p}$. The map $\mathcal{T}$ is linear in ρ, and needs to be *tomographically complete*, i.e.,

$$\rho \neq \rho' \Rightarrow \mathcal{T}(\rho) \neq \mathcal{T}(\rho') \,. \tag{4.2.5}$$

Consequently, $\mathcal{T}$ is a linear and injective map and as such has a left-inverse $\mathcal{T}^{-1}$ with $\mathcal{T}^{-1}[\mathcal{T}(\rho)] = \rho$. Hence, given $\boldsymbol{p}$, we can unambiguously obtain ρ via $\rho = \mathcal{T}^{-1}(\boldsymbol{p})$.

The notion of tomographic completeness of measurement schemes is very much related to that of information completeness of POVMs. Let $\mathcal{H}_d$ be a d-dimensional Hilbert space and $\mathcal{L}(\mathcal{H}_d)$ the set of Hermitian operators on that space. Let $\{E_i\}_{i=1}^m$ be a POVM on $\mathcal{H}_d$, i.e., a set of positive semidefinite

[1] Actually, the requirement that $\rho^{(N)}$ be derivable from $\rho^{(N+1)}$ for any N is also necessary.

operators $E_i \in \mathcal{L}(\mathcal{H}_d)$ so that $\sum_{i=1}^m E_i = \mathbb{1}$. This POVM is said to be *informationally complete* (IC) if the set of operators E_i span $\mathcal{L}(\mathcal{H}_d)$. Equivalently, IC means that for every pair of distinct states $\rho \neq \rho'$, there exists $i \in \{1, \dots, m\}$ such that $\mathrm{tr}(E_i\rho) \neq \mathrm{tr}(E_i\rho')$. We will see later how this idea plays out under usual measurement schemes for state tomography.

For now, let us assume that ρ is d-dimensional and we measure each copy of ρ with an IC POVM $\{E_i\}_{i=1}^m$. When we measure ρ, we obtain outcome i with probability determined by the Born rule as

$$p_i := \Pr(i|\rho) = \mathrm{tr}(E_i\rho)\,. \tag{4.2.6}$$

and p_i for $i = 1, \dots, m$ comprise the vector of probabilities $\boldsymbol{p}$. In practice, we cannot access the true probabilities p_i. If we could, inverting the system of equations given by Eq. (4.2.6) would yield the true state ρ unambiguously. What we can do is to estimate p_i by the *empirical frequency*

$$f_i = \frac{n_i}{N}\,, \tag{4.2.7}$$

where n_i is the number of times outcome i was observed, and N is the total number of copies of ρ measured. Indeed, by the law of large numbers,

$$\langle \boldsymbol{f} \rangle = \boldsymbol{p}\,, \quad \text{and} \quad \boldsymbol{f} \to \boldsymbol{p} \quad \text{as} \quad N \to \infty\,, \tag{4.2.8}$$

where the expectation value is taken over the N repetitions.

4.3 Reconstruction of the density operator: point estimators

Since $\boldsymbol{f}$ is all the information we can hope to get from our experiment, it makes sense to take it as an approximation of $\boldsymbol{p} = \mathcal{T}(\rho)$ and select a state ρ for which $\boldsymbol{f}$ and $\boldsymbol{p}$ are most similar. We could naively formalize the following optimization problem (a least-squares fit) to find an *estimate* $\hat{\rho}$ of the true state:

$$\hat{\rho} = \underset{\rho \in \mathcal{L}(\mathcal{H}_d)}{\mathrm{argmin}} ||\boldsymbol{f} - \mathcal{T}(\rho)||_2^2\,, \tag{4.3.9}$$

where ρ is understood as a valid density matrix, i.e., $\rho \geq 0$ and $\mathrm{tr}\,\rho = 1$. Unfortunately, the resulting estimate from this optimization is *biased*, that is, $\langle \hat{\rho} \rangle \neq \rho$. A method to produce an estimate $\hat{\rho}$ from $\boldsymbol{f}$ is called an *estimator* (formally, a map $\boldsymbol{f} \to \hat{\rho}$). Unsurprisingly, an estimator is *unbiased* if

$$\langle \hat{\rho} \rangle_\rho = \rho\,, \tag{4.3.10}$$

that is, if the estimate $\hat{\rho}$ fluctuates around the actual state ρ from which the data $\boldsymbol{f}$ was obtained. Of course, we would like our estimators to be unbiased, however they do not exist for density matrices!

Unbiased estimators that yield valid density matrices do not exist. PROOF.
Assume that we measure a pure state $\rho = |\psi\rangle\langle\psi|$. We have

$$\langle \hat{\rho} \rangle_\rho = \sum_{\boldsymbol{f}} \Pr(\boldsymbol{f}|\rho)\hat{\rho}(\boldsymbol{f})\,, \tag{4.3.11}$$

where $\Pr(\boldsymbol{f}|\rho)$ is the probability of the observed frequencies to occur given that the state was ρ, and $\hat{\rho}(\boldsymbol{f})$ is the estimated state for $\boldsymbol{f}$ according to an estimator. Since ρ is pure, it cannot be expressed as a convex combination of other states. Hence $\langle \hat{\rho} \rangle = \rho$ can only hold if $\hat{\rho}(\boldsymbol{f}) = \rho$ for all $\boldsymbol{f} \in S$, where $S := \{\boldsymbol{f} | \Pr(\boldsymbol{f}|\rho) > 0\}$ is the set of possible data (note that such estimator would already be pathological).

Now consider a different non-orthogonal state $\rho' = |\psi'\rangle\langle\psi'|$ ($\langle\psi|\psi'\rangle \neq 0$) with a dataset S'. Because the states are not perfectly distinguishable in any finite experiment, there exists a nonempty set of data $S_{\mathrm{int}} = S \cap S' = \{\boldsymbol{f} | \Pr(\boldsymbol{f}|\rho) \cdot \Pr(\boldsymbol{f}|\rho') > 0\}$ which can occur for both states. As we have argued, for $\boldsymbol{f} \in S_{\mathrm{int}}$ we have $\hat{\rho}(\boldsymbol{f}) = \rho$, but since $\boldsymbol{f}$ can also occur under state ρ' we also have $\hat{\rho}(\boldsymbol{f}) = \rho'$ if $\boldsymbol{f} \in S'/S_{\mathrm{int}}$. Hence, Eq. (4.3.11) for state ρ' would be a mixture of at least two pure states, violating the unbiasedness condition $\langle \hat{\rho} \rangle_{\rho'} = |\psi'\rangle\langle\psi'|$. The only way to avoid this situation would be for the set S_{int} to be empty, which can only happen if $|\psi\rangle$ and $|\psi'\rangle$ are orthogonal, but this contradicts our assumption that $\langle\psi|\psi'\rangle \neq 0$. ■

What we can do is to consider a relaxation of the optimization problem Eq. (4.3.9) that allows for estimates that are not necessarily positive-semidefinite and thus are not valid density matrices. We then have

$$\hat{\rho} = \underset{X \in \mathcal{L}(\mathcal{H}_d)}{\operatorname{argmin}} ||\boldsymbol{f} - \mathcal{T}(X)||_2^2 = \mathcal{T}^{-1}(\boldsymbol{f})\,, \tag{4.3.12}$$

where the optimization is now over all Hermitian operators X of unit trace[2]. Typically, the resulting $\hat{\rho}$ will display negative eigenvalues. We can argue that this occurs mainly for two reasons: first, in an experiment we always have to deal with finite statistics, hence $\boldsymbol{f}$ may not correspond to a valid probability vector for the given measurement scheme $\mathcal{T}$; and second, we assume that all measured ρ's are identical and uncorrelated (the *i.i.d.* assumption), but this hardly occurs in practice. In principle, however, with an increasing number of samples these negative eigenvalues will become smaller.

In addition, any point estimator coming out of an experiment is useless, unless we also provide error bars. In the case of tomography, these are multidimensional error bars in the state space, and receive the name of *confidence regions*, or *credible regions* in the case of using a Bayesian approach (see Section 4.5). In a successful tomography experiment, the confidence region will contain with high probability also operators that qualify as proper density operators.

We now review in detail two of the most used point estimators in tomography: *least squares* and *maximum likelihood*.

4.3.1 The least squares estimator

The *least squares (LS) estimator* is the solution to Eq. (4.3.12). Let us look for more explicit forms of $\mathcal{T}^{-1}(\boldsymbol{f})$. A way to invert $\mathcal{T}$ is to consider a vectorization of the Born rule, that is, a way to write it as an inner product. For a POVM with elements $\{E_i\}_{i=1}^m$, the map $\mathcal{T}$ is defined component-wise as $[\mathcal{T}(X)]_i = \operatorname{tr} E_i X$. We can write in matrix form

$$[\mathcal{T}(\rho)]_i = \operatorname{tr} E_i \rho = \sum_{jk} [E_i]_{jk} \rho_{kj} = \tilde{E}_i \cdot \tilde{\rho} = p_i\,, \tag{4.3.13}$$

with the vectorizations

$$\tilde{E}_i := ([E_i]_{11}, \ldots, [E_i]_{1d}, [E_i]_{21}, \ldots, [E_i]_{2d}, \ldots, [E_i]_{d1}, \ldots, [E_i]_{dd})\,, \tag{4.3.14}$$

$$\tilde{\rho} := (\rho_{11}, \ldots, \rho_{d1}, \rho_{12}, \ldots, \rho_{d2}, \ldots, \rho_{1d}, \ldots, \rho_{dd})^T\,, \tag{4.3.15}$$

and the subindices indicate elements according to some basis (e.g., the computational basis). Defining $\tilde{E}$ as the $m \times d^2$ matrix whose rows are given by $\tilde{E}_i$, $i = 1, \ldots, m$, the map operation $\mathcal{T}(\rho) = \boldsymbol{p}$ can be expressed as the single matrix equation

$$\tilde{E} \cdot \tilde{\rho} = \boldsymbol{p}\,. \tag{4.3.16}$$

We can readily invert this linear system of equations. We have

$$\tilde{E}^T \tilde{E} \cdot \tilde{\rho} = \tilde{E}^T \cdot \boldsymbol{p} \implies \tilde{\rho} = (\tilde{E}^T \tilde{E})^{-1} \cdot \tilde{E}^T \cdot \boldsymbol{p} \tag{4.3.17}$$

(note that we could not directly invert $\tilde{E}$ without multiplying from the left by $\tilde{E}^T$ since it is not a square matrix). Thus, substituting $\boldsymbol{p}$ by the empirical frequencies $\boldsymbol{f}$ in Eq. (4.3.17) and undoing the vectorization of $\tilde{\rho}$ we have an explicit way of computing $\mathcal{T}^{-1}(\boldsymbol{f})$.

Exercise 4.3.1. *Show that $\tilde{E}^T \cdot \boldsymbol{p} = \sum_{i=1}^m p_i E_i$ in operator form by undoing the vectorization.*

Exercise 4.3.2. *Another way to vectorize the Born rule is to use the generalized Bloch representation. Work out explicitly this vectorization for a Hilbert space of dimension 2^n so that we can write the Born rule as in Eq. (4.3.16).*

[2] The inverse map $\mathcal{T}^{-1}$ is at this point technically ill-defined if it acts on vectors $\boldsymbol{f}$ which do not correspond to probabilities arising from $\mathcal{T}$, but we will define it properly in the next section.

Expressed in terms of maps, the least squares estimator reads as

$$\hat{\rho}_{\mathrm{LS}} = (\mathcal{T}^\dagger \mathcal{T})^{-1} \mathcal{T}^\dagger(\boldsymbol{f}) \,, \tag{4.3.18}$$

where $\mathcal{T}^\dagger : \mathbb{R}^m \to \mathcal{L}(\mathcal{H}_d)$ is the dual map of $\mathcal{T}$ [3]. Note that since we have an IC POVM $\{E_i\}_{i=1}^m$ associated to $\mathcal{T}$, the operators E_i span the whole space $\mathcal{L}(\mathcal{H}_d)$ and thus form an overcomplete operator basis for that space, thus we can express any ρ as

$$\rho = \sum_{i=1}^{m} \mathrm{tr}(E_i \rho) E_i \,. \tag{4.3.19}$$

Since $\mathcal{T}(\rho) = \boldsymbol{p}$ and the components of $\boldsymbol{p}$ are precisely the coefficients in the above equation, we see that the dual map acting on an arbitrary vector $\boldsymbol{x} \in \mathbb{R}^m$ gives an operator in $\mathcal{L}(\mathcal{H}_d)$ as above, i.e.,

$$\mathcal{T}^\dagger(\boldsymbol{x}) = \sum_{i=1}^{m} x_i E_i \tag{4.3.20}$$

We readily see that $\mathcal{T}$ also obeys

$$(\mathcal{T}^\dagger \mathcal{T})(X) = \sum_{i=1}^{m} \mathrm{tr}(E_i X) E_i \,. \tag{4.3.21}$$

From the last equation we see that the inverse $(\mathcal{T}^\dagger \mathcal{T})^{-1}$ will only exist if the operators E_i span the whole space. Otherwise, $(\mathcal{T}^\dagger \mathcal{T})(X)$ would have a nontrivial kernel and thus would be non-invertible. The particular form of this inverse depends on the measurement setting used (see Section 4.3.2).

The *projected* least squares estimator

A way of obtaining a physical (although biased) estimate is to project $\hat{\rho}_{\mathrm{LS}}$ onto the set of quantum states with respect to the Hilbert-Schmidt distance (or 2-norm), that is,

$$\hat{\rho}_{\mathrm{PLS}} = \underset{\sigma \in \mathcal{L}(\mathcal{H}_d)}{\mathrm{argmin}} \, \|\hat{\rho}_{\mathrm{LS}} - \sigma\|_2 \,, \tag{4.3.22}$$

where σ are valid density matrices, i.e., $\mathrm{tr}\,\sigma = 1$, and $\sigma \geq 0$. The PLS estimator behaves better than the one in Eq. (4.3.9). In particular, rigorous confidence regions have been established for the former (see Section 4.5), while a similar result for the latter is lacking. As discussed above, providing a confidence region is absolutely necessary to report valid experimental results, and mitigates in some way the uncertainty that the bias of the estimator generates.

4.3.2 LS for specific measurement schemes

Prominent measurement schemes for state tomography are local Pauli measurements, Pauli observables, uniform POVMs, and SIC-POVMs. In this section we will explicitly evaluate Eq. (4.3.18) for some of these measurements. Details of the derivations can be found in Appendix 4.A.

Local Pauli measurements

Let us assume that ρ is a multipartite state of n qubits. The experimentally easiest and most common way to measure multiqubit states is to apply Pauli measurements to each qubit. There are 3^n different arrangements of Pauli measurements that we can apply on ρ, that we label by $\boldsymbol{s} \in \{X, Y, Z\}^{\times n}$. Each possible value of $\boldsymbol{s}$ is called a *measurement setting*, and it has 2^n possible outcomes enumerated by $\boldsymbol{x} \in \{+1, -1\}^{\times n}$. To carry out a tomography experiment successfully, we need to get data from all

[3] This form is a more developed expression of the left-inverse map $\mathcal{T}^{-1}$ in terms of $\mathcal{T}^\dagger$, that explicitly takes into account that the state space and the space of probabilities generally have different dimensions.

measurement settings $\boldsymbol{s}$. This becomes intuitively clear by noting that any n-qubit state can be generally expressed as

$$\rho = \sum_{i_1,\ldots,i_n} \alpha_{i_1,\ldots,i_n} \sigma_{i_1} \otimes \cdots \otimes \sigma_{i_n} \,, \tag{4.3.23}$$

where $\alpha_{i_1,\ldots,i_n}$ are real coefficients, $i_k \in \{0, X, Y, Z\}$ for qubit k, and σ_{i_k} are the Pauli matrices with $\sigma_0 = \mathbb{1}$. In other words, the operators $\sigma_{i_1} \otimes \cdots \otimes \sigma_{i_n}$ form an operator basis for $\mathcal{L}(\mathcal{H}_{2^n})$, hence we need to get information on every coefficient $\alpha_{i_1,\ldots,i_n}$ to fully determine ρ.

Exercise 4.3.3. *Suppose our measurement scheme consists in using only one measurement setting, where we apply a 2-outcome σ_Z measurement on each qubit, for each preparation of ρ. Show by counterexample that the arising map $\mathcal{T}$ is not tomographically complete.*

The outcomes of measurement setting $\boldsymbol{s}$ are associated to projections onto the vectors

$$|(\boldsymbol{s}, \boldsymbol{x})\rangle = \bigotimes_{k=1}^{n} |(s_k, x_k)\rangle \,, \tag{4.3.24}$$

where $|(s, x)\rangle$ is the eigenstate of the Pauli operator σ_s with eigenvalue x. These measurements yield on the state ρ the 6^n probabilities

$$p_{\boldsymbol{s},\boldsymbol{x}} = \langle(\boldsymbol{s}, \boldsymbol{x})| \rho |(\boldsymbol{s}, \boldsymbol{x})\rangle \,, \tag{4.3.25}$$

with $\sum_{\boldsymbol{x}} p_{\boldsymbol{s},\boldsymbol{x}} = 1$ for each setting $\boldsymbol{s}$. We write $\boldsymbol{p}$ for the collection of all probabilities.

For local Pauli measurements, the union of all projections onto the 6^n basis vectors Eq. (4.3.24) conform a map $\mathcal{T}$ that obeys

$$(\mathcal{T}^\dagger \mathcal{T})(X) = 3^n \mathcal{D}_{1/3}^{\otimes n}(X) \,, \tag{4.3.26}$$

where $\mathcal{D}_{1/3}(X) = \frac{1}{3}X + (1 - \frac{1}{3})\frac{\mathrm{tr}\, X}{2}\mathbb{1}$ is a single-qubit depolarizing channel with loss parameter $p = 1/3$. Eq. (4.3.18) takes the explicit form

$$\hat{\rho}_{\mathrm{LS}} = \frac{1}{3^n} \sum_{\boldsymbol{s},\boldsymbol{x}} f_{\boldsymbol{s},\boldsymbol{x}} (\mathcal{D}_{1/3}^{\otimes n})^{-1} \left(|(\boldsymbol{s}, \boldsymbol{x})\rangle\langle(\boldsymbol{s}, \boldsymbol{x})| \right) \tag{4.3.27}$$

$$= \frac{1}{3^n} \sum_{\boldsymbol{s},\boldsymbol{x}} f_{\boldsymbol{s},\boldsymbol{x}} \bigotimes_{k=1}^{n} (3\, |(s_k, x_k)\rangle\langle(s_k, x_k)| - \mathbb{1}) \,. \tag{4.3.28}$$

Pauli observables

Each operator in the decomposition of the n-qubit state ρ in Eq. (4.3.23) can be associated with a Pauli observable W_i, where $i = 1, \ldots, d^2$ labels all possible n-fold tensor products of Pauli matrices. The observables W_i form an operator basis of $\mathcal{L}(\mathcal{H}_{2^n})$, and thus

$$\rho = \frac{1}{d} \sum_{i=1}^{d^2} \mathrm{tr}(W_i \rho) W_i \tag{4.3.29}$$

By associating a 2-outcome POVM to each W_i with elements $P_i^{\pm} = \frac{1}{2}(\mathbb{1} \pm W_i)$, we can approximate the expectation value $\mathrm{tr}(W_i\rho)$ by the empirical mean $f_i^+ - f_i^-$, where $f_i^{\pm}$ are the observed frequencies of outcomes associated to $P_i^{\pm}$. We assume that each frequency $f_i^{\pm}$ is calculated over $N/(d^2 - 1)$ samples, i.e., we share the total number of samples equally among all nontrivial Pauli observables (one of them, say W_1, is simply the identity matrix, so we don't have to measure it). The least squares estimator is simply

$$\hat{\rho}_{\mathrm{LS}} = \frac{1}{d} \sum_{i=1}^{d^2} (f_i^+ - f_i^-) W_i \,. \tag{4.3.30}$$

In this case it is very easy to check that $\hat{\rho}_{\mathrm{LS}}$ is indeed unbiased:

$$\langle \hat{\rho}_{\mathrm{LS}} \rangle = \frac{1}{d} \sum_{i=1}^{d^2} \langle f_i^+ - f_i^- \rangle W_i = \frac{1}{d} \sum_{i=1}^{d^2} \mathrm{tr}(W_i \rho) W_i = \rho \tag{4.3.31}$$

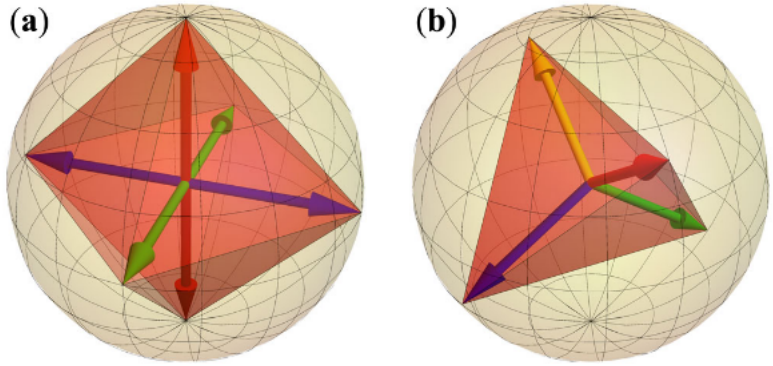


Figure 4.3.1: Bloch sphere representation of two IC POVMs on a qubit: (a) a set of three mutually unbiased bases along directions X, Y, and Z with six outcomes; and (b) a SIC-POVM with four outcomes pointing at the vertices of a regular tetrahedron. Source: Ref. [26].

Note that, from the implementation point of view, this measurement scheme is identical to the one in the previous section. What changes is the data processing, by means of an alternative definition of the empirical frequencies. The resulting estimator is thus different.

Symmetric informationally complete POVMs

Symmetric informationally complete (SIC) POVMs are a special kind of IC POVMs. Usual SIC-POVMs are made of d^2 rank-1 elements $\frac{1}{d}|v_i\rangle\langle v_i|$, with $|v_i\rangle$ normalized, that for $i \neq j$ fulfill

$$|\langle v_i|v_j\rangle|^2 = \frac{1}{d+1}\,. \tag{4.3.32}$$

It is an open conjecture whether rank-1 SIC-POVMs exist in all dimensions. For $d = 2$, the vectors $|v_i\rangle$ form a regular tetrahedron in the Bloch sphere (see Fig. 4.3.1). A very important property of SIC-POVMs is that they form *2-designs*. For our purposes, it will suffice to recall here that a set of rank-1 projectors $\{|v_i\rangle\langle v_i|\}_{i=1}^m$ form a 2-design if

$$\frac{1}{m}\sum_{i=1}^{m}|v_i\rangle\langle v_i|^{\otimes 2} = \frac{2}{d(d+1)}P_{\mathrm{Sym}^{(2)}}\,, \tag{4.3.33}$$

where $P_{\mathrm{Sym}^{(2)}}$ is the projector onto the fully symmetric subspace of $\mathcal{H}_d^{\otimes 2}$.

Considering a measurement scheme for tomography that measures every copy of ρ with a SIC-POVM $\{\frac{1}{d}|v_i\rangle\langle v_i|\}_{i=1}^{d^2}$, we can use this useful property (see Appendix 4.A for more details) to show that the associated map $\mathcal{T}$ fulfills

$$(\mathcal{T}^\dagger\mathcal{T})(X) = \sum_{i=1}^{d^2}\langle v_i|X|v_i\rangle\,|v_i\rangle\langle v_i| = \frac{d^2}{d+1}[X + (\mathrm{tr}\,X)\mathbb{1}]\,, \tag{4.3.34}$$

which can be readily inverted, yielding the least squares estimator

$$\hat{\rho}_{\mathrm{LS}} = (d+1)\sum_{i=1}^{d^2} f_i\,|v_i\rangle\langle v_i| - \mathbb{1}\,. \tag{4.3.35}$$

Note that in the case that ρ is an n-qubit state, measuring each qubit with a tetrahedron (a SIC-POVM for $d = 2$) is also a valid tomographic measurement scheme, that is, the resulting scheme is also tomographically complete. In such case, however, one should modify appropriately the above estimator.

4.3.3 The maximum likelihood estimator

Maximum likelihood (ML) estimation was proposed in quantum state tomography as a 'solution' to the problem of the least squares estimator being typically unphysical [27, 28]. Instead of finding which Hermitian matrix is determined by the observed frequencies (inverting the Born rule), ML asks *what quantum state is the one that is most likely to have generated the data?* The probability of a state ρ generating the frequency vector $\boldsymbol{f}$ is given by the likelihood function

$$\mathcal{L}(\rho) = \Pr(\boldsymbol{f}|\rho)\,. \tag{4.3.36}$$

The particular form of this probability will obviously depend on $\mathcal{T}$. If we measure each ρ with a POVM $\{E_i\}_{i=1}^m$, the probabilities of observing each outcome i follow a multinomial distribution, and we have

$$\mathcal{L}(\rho) = \frac{N!}{\prod_i n_i!} \prod_j \operatorname{tr}(E_j\rho)^{n_j}\,, \tag{4.3.37}$$

where recall that n_j is the total count of observations of outcome j (the multinomial factor on the above expression is not important for the discussion and we will omit it). Thus, we define the maximum likelihood estimator as the maximization over density matrices

$$\hat{\rho}_{\mathrm{ML}} = \operatorname*{argmax}_{\rho\in\mathcal{L}(\mathcal{H}_d)} \mathcal{L}(\rho)\,. \tag{4.3.38}$$

The ML estimator has a number of desirable properties:

- *The likelihood function has a unique local maximum.* First, note that $\mathcal{L}(\rho)$ is non-negative, hence we can maximize $\log\mathcal{L}(\rho)$ instead. And second, $\log\mathcal{L}(\rho)$ is concave. The proof goes as follows: realize that
$$\log\mathcal{L}(\rho) = \sum_i n_i \log\operatorname{tr}(E_i\rho)\,; \tag{4.3.39}$$
the trace is a non-negative linear function of ρ, the log of a linear function is concave, and the sum of concave functions is concave. A maximization of a concave function over a convex set has a unique maximizer.

- *The ML and LS estimators coincide if* $\hat{\rho}_{\mathrm{LS}} \geq 0$. Recall that $\hat{\rho}_{\mathrm{LS}}$ returns a Hermitian matrix that exactly reproduces the observed frequencies $\boldsymbol{f}$, as it is just the inverse map $\mathcal{T}^{-1}(\boldsymbol{f})$. From Eq. (4.3.37), we write
$$\log\mathcal{L}(\rho) = \sum_j n_j \log\operatorname{tr}(E_j\rho) = \sum_j n_j \log p_j = N\sum_j f_j \log p_j \tag{4.3.40}$$
$$= N\sum_j \left[f_j\log f_j - (f_j\log f_j - f_j\log p_j)\right] \tag{4.3.41}$$
$$= -N\left[H(\boldsymbol{f}) + D(\boldsymbol{f}||\boldsymbol{p})\right]\,, \tag{4.3.42}$$
where H is the Shannon entropy, and D is the relative entropy. Note that $H(\boldsymbol{f})$ does not depend on $\boldsymbol{p}$ and thus it is not relevant for the maximization. On the other hand, $D(\boldsymbol{f}||\boldsymbol{p})$ is always non-negative and has a unique zero when $\boldsymbol{p} = \boldsymbol{f}$, implying that $\log\mathcal{L}(\rho)$ is uniquely maximized at this point. Thus, if $\hat{\rho}_{\mathrm{LS}} \geq 0$, then $\boldsymbol{f}$ is an exactly realizable probability vector, one can assert that $\boldsymbol{p} = \boldsymbol{f}$, and $\hat{\rho}_{\mathrm{ML}} = \hat{\rho}_{\mathrm{LS}}$.

However, if $\hat{\rho}_{\mathrm{LS}} \not\geq 0$, $\hat{\rho}_{\mathrm{ML}}$ does not coincide with $\hat{\rho}_{\mathrm{LS}}$ and is always a rank-deficient state. To understand this, consider the polytope in Hilbert space described by facets set by the equalities $\operatorname{tr} E_i\rho = 0$. The LS estimator predicts non-negative probabilities for each observable outcome, i.e., $\operatorname{tr} E_i\rho \geq 0$, hence $\hat{\rho}_{\mathrm{LS}}$ is always a point inside this polytope. Since ρ has a unique maximizer at $\rho = \rho_{\mathrm{LS}}$, the maximization in Eq. (4.3.38) over a closed subset of the polytope (the state space) that does not contain ρ_{LS} always yields a point in the boundary of the subset, that is, a rank-deficient state $\hat{\rho}_{\mathrm{ML}}$ (see Fig. 4.3.2).

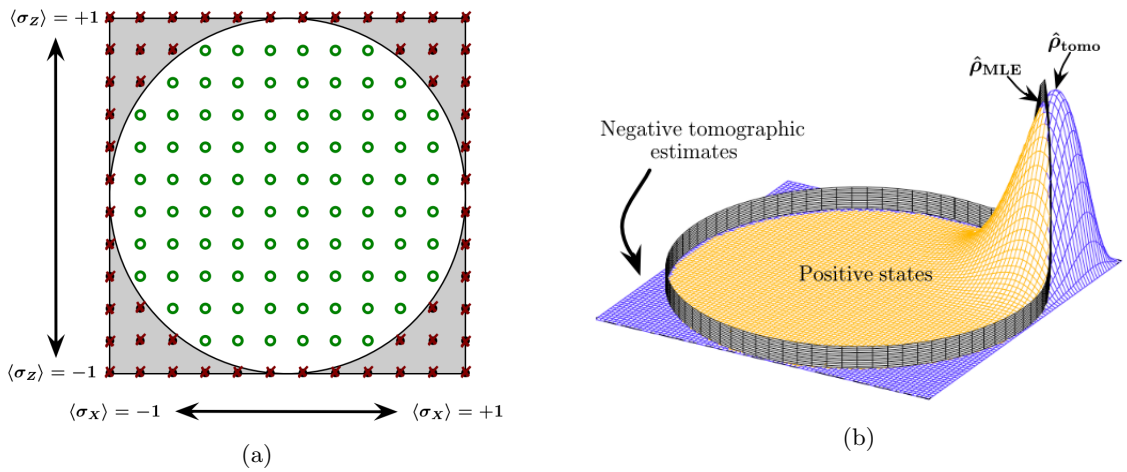


Figure 4.3.2: (a) For a single qubit, a section of the polytope defined by $\mathrm{tr}\, E_{s,x}\rho \geq 0$, where $s = X, Z$, $x = \pm 1$, and $\mathrm{tr}\, E_{Y,x}\rho = 0$. The points correspond to possible estimates from $N = 11$ measurements of σ_Z and σ_X each. The circle is the inscribed Bloch equator, separating physical from unphysical estimates. (b) Illustration of how the constrained maximization of ML makes $\hat{\rho}_{\mathrm{ML}}$ lie in the boundary of the state space, if $\hat{\rho}_{\mathrm{LS}}$ is unphysical. Source: Ref. [29].

Example 4.3.1. Suppose that we measure σ_X, σ_Y, and σ_Z over a single qubit, applying each measurement just once. In each case, we observe the outcome $+1$. The LS estimator

$$\hat{\rho}_{\mathrm{LS}} = \begin{pmatrix} 1 & \frac{1+i}{2} \\ \frac{1-i}{2} & 0 \end{pmatrix} \tag{4.3.43}$$

satisties $\langle \sigma_X \rangle = \langle \sigma_Y \rangle = \langle \sigma_Z \rangle = 1$ and is Hermitian, but it has a negative eigenvalue $\lambda = 1 - \sqrt{3}/2 \approx -0.366$. It therefore lies outside the state space. It is also clear that no quantum state exists for which these three spin measurements are perfectly predictable.

The problem with having a point estimator with zero eigenvalues is that these are unjustifiable in any experiment, no matter the amount of data collected, since any amount of uncertainty around a probability $\hat{p} = 0$ would mean accepting negative probabilities. Thus claiming $\hat{p} = 0$ for a measurement outcome necessarily means two things: that we are *absolutely certain* that such is the true probability for that outcome given the observed data, and that we are *absolutely certain* that this outcome will not come up in any future measurement.[4]

So, we have seen that an unconstrained version of the ML estimator, without the condition $\rho \geq 0$, coincides with the LS estimator. Conceptually, adding the semidefinite positive constraint means that we seek an estimator capable of making predictions on future measurements (such predictions being dictated by the Born rule), not just explaining well the observed data. Apart from the issue of estimates with zero eigenvalues, there is the additional issue that, as a result of imposing this constraint, $\hat{\rho}_{\mathrm{ML}}$ is *biased*. An illustration of this effect can be seen in Fig. 4.3.3.

4.4 Bayesian mean estimation

An alternative approach to point estimators is Bayesian mean estimation (BME) [29]. This approach is fundamentally different (even at a philosophical level) from the frequentist approaches that report point estimators that we have seen, and it comes with a number of advantages over those: (a) the estimates produced are *always full rank, physical density matrices*; (b) the method gives a natural way of computing *error bars*; and (c) BME is optimal in the sense that it minimizes any *average operational divergence*, e.g., the squared Hilbert-Schmidt distance $\mathrm{tr}[(\rho - \hat{\rho})^2]$ or the quantum relative entropy $\mathrm{tr}[\rho(\log \rho - \log \hat{\rho})]$.

[4] Would you say that the probability of tossing a coin and getting heads is zero just because you never got tails in 100 tosses? How many times do you have to toss to confidently make such claim?

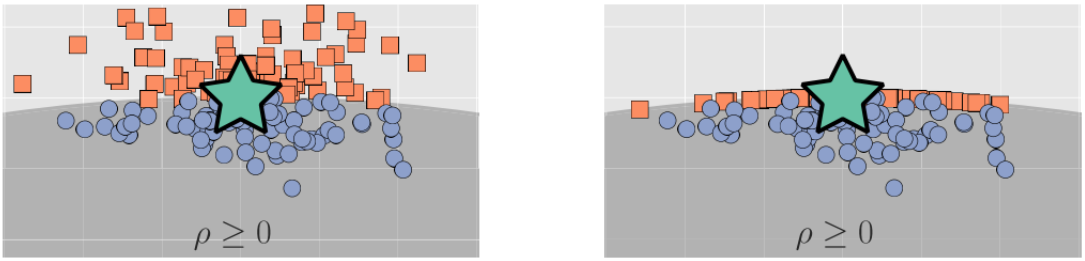


Figure 4.3.3: Left: ML estimates without imposing $\rho \geq 0$ over a pure qutrit state (star), where the orange squares are outside the state space and the blue circles inside; the estimates are (asymptotically) Gaussian-distributed. Right: the same estimates enforcing $\rho \geq 0$; all orange squares lie on the boundary, the distribution of the estimates is no longer Gaussian, and its mean will be clearly biased towards the inside of the state space. Source: Ref. [30].

The method works as follows:

1. Generate a likelihood function from the observed data $\mathcal{L}(\rho) = \Pr(\boldsymbol{f}|\rho)$.
2. Choose a prior distribution over states, $\pi_0(\rho)d\rho$.
3. Multiply them and normalize the result to obtain a posterior distribution $\pi_f(\rho)d\rho \propto \mathcal{L}(\rho)\pi_0(\rho)d\rho$.
4. The BME estimator is the mean of this posterior:

$$\hat{\rho}_{\text{BME}} = \int \rho\, \pi_f(\rho) d\rho\,. \tag{4.4.44}$$

A couple of noteworthy comments. First, one should choose the prior $\pi_0(\rho)d\rho$ as uninformatively as possible, for one should make minimum assumptions on the state in the setting of quantum state tomography, which generally assumes ρ to be *completely unknown*. The prior is supposed to capture our state of knowledge before performing the experiment. The very fact of having to select a prior makes $\hat{\rho}_{\text{BME}}$, yet again, a *biased* estimator, no matter how uninformative the prior is. A way to directly observe this effect is to realize that, if the prior is uniform, the peak of the posterior distribution $\pi_f(\rho)d\rho$ is given by the most likely state, i.e., $\hat{\rho}_{\text{ML}} = \text{argmax}_\rho \mathcal{L}(\rho)$, which we already know is a biased estimator. And as important as selecting the prior, we need to choose a measure over quantum states, that is, a specific form of $d\rho$. Whereas the Haar measure is unique in the sense that it is the only unitarily invariant measure under the group $\text{SU}(d)$, this holds for pure states, and there is no canonical way of extending this measure to mixed states.[5]

Then, $\hat{\rho}_{\text{BME}}$ implies computing an integral that is generally not amenable to analytical methods. Typically, one resorts to using a Metropolis-Hastings algorithm, a clever Monte Carlo integration method that samples from a given probability distribution more efficiently than a simple rejection-sampling algorithm. We will not enter here into details about how this algorithm works (a good, practical reference is [31]).

Let us now comment further on the advantages of BME briefly mentioned above:

(a) *The estimator $\hat{\rho}_{\text{BME}}$ is always a full rank state.* This holds true for any reasonable and robust prior. Without entering into what constitutes the technical definition of "reasonable", just observe that, intuitively, any such prior (defined over the whole state space) should give some nonzero weight to a full-rank state, say the maximally mixed state $\mathbb{1}_d/d$. If this is the case, no amount of data will ever rule out this state (its likelihood will not vanish), since any measurement outcome has nonzero probability of occurring upon measuring it. Thus, the maximally mixed state will be part

[5] All this choice is the root of the main criticisms that frequentism holds against Bayesian methods. At the same time, Bayesian statistics is, arguably, the more natural way of thinking about probability.

of the integration domain in Eq. (4.4.44), and consequently $\hat{\rho}_{\rm BME}$ will be a convex combination of $\mathbb{1}_d/d$ and other states, resulting in a full rank estimate $\hat{\rho}_{\rm BME}$.

A robust prior is one for which there is no finite-length measurement record that annihilates it. Equivalently, a prior that assigns zero probability to a particular event is fragile, in the sense that it will set π_f to zero upon observing such event, making the Bayesian approach fail completely.

Example 4.4.1. To further illustrate the fact that BME reports more sensible estimates than ML, consider the example of estimating the bias b of a coin, with b the probability of obtaining 'heads' and $1-b$ that of obtaining 'tails'. After N tosses, we obtain 'heads' n times and 'tails' $N-n$ times. The likelihood function is

$$\mathcal{L}(b) = b^n (1-b)^{N-n} ,$$

and the ML estimate is

$$\hat{b}_{\rm ML} = \frac{n}{N} .$$

A problem appears if $n = 0$ (N), since $\hat{b}_{\rm ML} = 0$ (1) regardless of the value of N, i.e., we assess that either one outcome or the other never occurs.

A more reasonable estimate is provided by BME. Choosing a uniform prior for b in the $[0,1]$ interval, we have

$$\hat{b}_{\rm BME} = \frac{n+1}{N+2} .$$

Since $0 \leq n \leq N$, this estimate never assigns a zero probability to any of the outcomes of a tossing. With no data, $\hat{b}_{\rm BME} = 1/2$. After N trials, if 'heads' is never observed, its probability is still estimated to be $\hat{b}_{\rm BME} = 1/(N+2) \approx 1/N$, which is a reasonable guess.

(b) *Natural error bars.* BME provides a very natural way of computing the variance of any observable X measured on $\hat{\rho}_{\rm BME}$. Denote by $\Delta\rho$ the superoperator

$$\Delta\rho = \int \rho \otimes \rho \, \pi_f(\rho) d\rho - \hat{\rho}_{\rm BME} \otimes \hat{\rho}_{\rm BME} . \tag{4.4.45}$$

This superoperator, acting on two copies of X, produces the variance $\Delta\langle X\rangle^2$ of X, where $\langle\cdot\rangle$ means the expectation value over $\hat{\rho}_{\rm BME}$. Indeed,

$$\begin{aligned}
\Delta\langle X\rangle^2 &= {\rm Tr}(X \otimes X \Delta\rho) && (4.4.46)\\
&= {\rm Tr}\left[X \otimes X \int \rho \otimes \rho \, \pi_f(\rho) {\rm d}\rho\right] - {\rm Tr}\left[X \otimes X \hat{\rho}_{\rm BME} \otimes \hat{\rho}_{\rm BME}\right] && (4.4.47)\\
&= \int {\rm Tr}[X\rho] \cdot {\rm Tr}[X\rho] \pi_f(\rho) {\rm d}\rho - {\rm Tr}\left[X \hat{\rho}_{\rm BME}\right]^2 && (4.4.48)\\
&= \int \langle X\rangle_\rho^2 \pi_f(\rho) {\rm d}\rho - \left[\int \langle X\rangle_\rho \pi_f(\rho) {\rm d}\rho\right]^2 , && (4.4.49)
\end{aligned}$$

which resembles the variance of the distribution $\pi_f(\langle X\rangle)$.[6]

Besides the point estimator $\hat{\rho}_{\rm BME}$, it is fairly straightforward to define *credible regions* $\mathcal{R}$ around it consisting in regions in the state space for which the likelihood $\mathcal{L}(\rho)$ is above a certain threshold. These regions have a clear meaning: one can compute a function of them called *credibility*, which is interpreted as the probability that $\rho \in \mathcal{R}$ if $\boldsymbol{f}$ was obtained, i.e., the posterior probability of $\mathcal{R}$. For details on how to compute credible regions, see [32].

[6] To see this even more clearly, define the univariate distribution $\pi_f(\langle X\rangle) d\langle X\rangle \equiv \int_\sigma \pi_f(\rho) d\rho$, where σ includes the other d^2-2 dimensions besides the one spanned by $\langle X\rangle$.

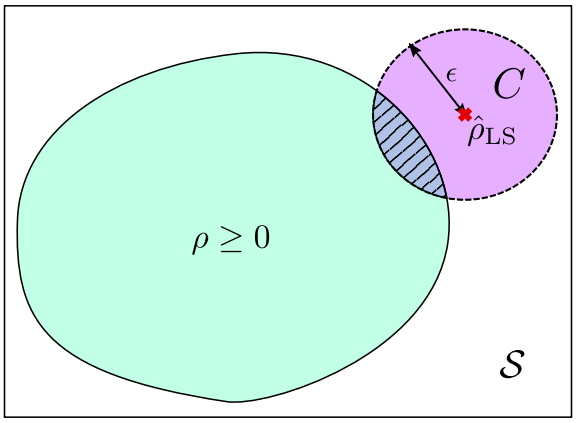


Figure 4.5.1: Illustration of the state space (green) as a subset of the space of Hermitian matrices $\mathcal{S}$, and its intersection with a confidence region C (purple) for the unphysical estimate $\hat{\rho}_{\mathrm{LS}}$ of size ϵ and confidence level $1-\delta$.

4.5 Confidence regions

A *confidence region* for quantum state tomography is a function C from the measurement data to the space of Hermitian operators such that it determines a region in this space that contains the true state with high probability. Mathematically, we express this requirement for C as

$$\Pr[\rho \in C(\boldsymbol{f})] \geq 1-\delta\,, \quad \text{or equivalently} \quad \Pr[\rho \notin C(\boldsymbol{f})] \leq \delta\,, \tag{4.5.50}$$

where the state ρ and the data $\boldsymbol{f}$ are connected via $\boldsymbol{f} = \mathcal{T}(\rho)$. The parameter δ is typically small for a good confidence region when N is reasonably large, and $1-\delta$ is the *confidence* with which we can claim that $\rho \in C$.

We write $\mathcal{X}$ for a linear space in which all possible data can be represented and $\mathcal{S}$ for the linear space space spanned by all Hermitian matrices in dimension d. With these notions, a confidence region is conveniently defined by means of a norm, either in $\mathcal{S}$ or in $\mathcal{X}$. In the first case one uses a point estimator $\mathcal{R}\colon \mathcal{X} \to \mathcal{S}$ and a norm $\|\cdot\|_{\mathcal{S}}$ in $\mathcal{S}$ to define for $\boldsymbol{f} \in \mathcal{X}$

$$C_{\mathrm{inv}}(\boldsymbol{f}) = \{\rho \in \mathcal{S}, \rho \geq 0, \mathrm{tr}(\rho) = 1 |\, \|\rho - \mathcal{R}(\boldsymbol{f})\|_{\mathcal{S}} \leq \epsilon\}\,. \tag{4.5.51}$$

In the second case one assumes a prediction mapping $\mathcal{T}\colon \mathcal{S} \to \mathcal{X}$ and a norm $\|\cdot\|_{\mathcal{X}}$. Then we let for $x \in \mathcal{X}$

$$C_{\mathrm{fwd}}(\boldsymbol{f}) = \{\rho \in \mathcal{S}, \rho \geq 0, \mathrm{tr}(\rho) = 1 |\, \|\mathcal{T}(\rho) - \boldsymbol{f}\|_{\mathcal{X}} \leq \tau\}\,. \tag{4.5.52}$$

As a technical detail, $\mathcal{R}$ might be only defined on the possible data, not on all of $\mathcal{X}$ and similarly, $\mathcal{T}$ might be only defined on the set of states. Figure 4.5.1 illustrates a confidence region of the first type around the LS estimator $\hat{\rho}_{\mathrm{LS}}$. In this lecture we will only focus on this type of confidence region.

Note that the least overlap C has with the state space, the stronger is the statement of Eq. (4.5.50), since we are only interested in the *physical* states contained in C (naturally, we are confident that the state prepared in the laboratory is physical). Also note that the size of the region ϵ and the confidence level $1-\delta$ are interrelated parameters. Indeed, intuitively, for the same amount of data, lowering the confidence allows to define a smaller region, while high confidence levels require regions to be larger. The particular form of this relation strongly depends on the method chosen to construct C.

We will review a few different ways to construct confidence regions as well as credible regions, the analogous concept in a Bayesian approach.

4.5.1 Confidence region for the LS estimator

We first show an analytical definition of C for the LS estimator, based on random matrix theory and concentration inequalities.

The first step is to realize that $\hat{\rho}_{\text{LS}}$ can be regarded as a sum of N independent random matrices. To see this, consider $N = 1$ and a SIC-POVM. The corresponding LS estimator is given by Eq. (4.3.35). Then $\hat{\rho}_{\text{LS}}$ is an instance of the random matrix $X = (d+1)\,|v_i\rangle\langle v_i| - \mathbb{1}$, where the randomness is in the index $i \in [1, \ldots, d^2]$, which occurs with probability $\frac{1}{d}\,\langle v_i|\,\rho\,|v_i\rangle$. After N trials, we find

$$\hat{\rho}_{\text{LS}} = \frac{1}{N}\sum_{j=1}^{N} X_j\,,$$

where the matrices X_j are statistically independent. Such sums of random matrices concentrate sharply around their expectation value $\langle\hat{\rho}_{\text{LS}}\rangle = \rho$. Matrix concentration inequalities, in particular the *matrix Bernstein inequality*, help us quantify this convergence.

Theorem 4.5.1 (Matrix Bernstein inequality) *Consider a sequence of N independent Hermitian matrices $A_1, \ldots, A_N \in \mathcal{H}_d$. Assume that each A_i satisfies*

$$\langle A_i\rangle = 0\,, \quad \text{and} \quad \|A_i\|_\infty \leq R\,.$$

Then, for any $t > 0$,

$$\Pr\left[\|\sum_{i=1}^{N}(A_i - \langle A_i\rangle)\|_\infty \geq t\right] \leq \begin{cases} d\,\mathrm{e}^{-\frac{3t^2}{8\sigma^2}} & t \leq \frac{\sigma^2}{R} \\ d\,\mathrm{e}^{-\frac{3t}{8R}} & t \geq \frac{\sigma^2}{R} \end{cases}$$

where $\sigma^2 = \|\sum_{i=1}^{N}\langle A_i^2\rangle\|_\infty$.

As already argued, in our case we have the random matrices X_i, all with the same mean $\langle X\rangle = \rho$, so we can already identify in the above theorem the distance $\|\hat{\rho}_{\text{LS}} - \rho\|_\infty$ in operator norm.[7] Using this result, one can see for the case of SIC-POVMs that

$$\Pr\left[\|\hat{\rho}_{\text{LS}} - \rho\|_\infty \geq \epsilon\right] \leq d\,\mathrm{e}^{-\frac{3N\epsilon^2}{16d}}\,. \tag{4.5.53}$$

We include a proof of this result in Appendix 4.B, extracted from Ref. [33]. For other types of POVMs, what changes in the above equation is the factor 16. From Eq. (4.5.53) we can confirm some intuitions that we advanced above: for fixed N, larger values of ϵ mean a larger confidence region, which implies a smaller right hand side, corresponding to greater confidence (and viceversa). If we want to achieve a desired confidence, we just equal the right hand side to δ and invert the relation to obtain the corresponding size of the region $\epsilon(\delta)$.

One can bound the operator norm with the trace distance or 1-norm, a widely used and operationally motivated distance measure between quantum states, and adapt Eq. (4.5.53) to a confidence region in trace distance. Doing so however can only be at the expense of enlarging the region for fixed N and confidence level (see [33] for more details, including a confidence region in trace distance for the *projected* LS estimator).

4.5.2 Confidence polytope for the LS estimator

In this section we show an alternative, algorithmic method to compute a confidence region in the form of a polytope in state space, described in [34].

Assume as usual that we measure N copies of ρ with a POVM $\{E_i\}_{i=1}^m$ to obtain a frequency vector $\boldsymbol{f}$. The approach here is to build a confidence region C as the intersection of *confidence half-spaces* C_i, one for each POVM element E_i. The half-spaces C_i depend on the observed frequencies f_i, and are defined as

$$C_i(f_i) = \{\rho \in \mathcal{S} : \operatorname{tr}(E_i\rho) \leq f_i + \gamma_N(f_i, \delta_i)\}\,, \tag{4.5.54}$$

[7] The operator or infinity norm $\|A\|_\infty$ corresponds to the largest eigenvalue of A, and it is a proper distance measure.

where $\gamma_N(f_i, \delta_i)$ is the positive root of $D(f_i||f_i + \gamma_N) = -\frac{1}{N}\log\delta_i$, $D(x||y) = x\log(\frac{x}{y}) + (1-x)\log(\frac{1-x}{1-y})$ is the Kullback-Leibler divergence, and $\sum_i \delta_i = \delta$ with $1-\delta$ the desired confidence level. Then

$$C(\boldsymbol{f}) := \bigcap_i C_i(f_i) \tag{4.5.55}$$

is a $1-\delta$ confidence region. See Appendix 4.C for a proof.

Each half-space is a so-called *Clopper-Pearson confidence interval*, which is a well established method for computing confidence intervals for the binomial distribution. Eq. (4.5.55) defines a region around the estimator $\hat{\rho}_{\text{LS}}$ by means of the buffer terms γ_N (if we set all γ_N to zero the region only contains one point, the linearly inverted state). The resulting region can be easily computed using the Bloch form of ρ and E_i:

$$\rho = \frac{1}{d}\left(\mathbb{1} + \sqrt{\frac{d(d-1)}{2}}\boldsymbol{r}\cdot\boldsymbol{\lambda}\right), \tag{4.5.56}$$

$$E_i = \frac{1}{a_i}\left(\mathbb{1} + \sqrt{\frac{d(d-1)}{2}}\boldsymbol{\eta}_i\cdot\boldsymbol{\lambda}\right), \tag{4.5.57}$$

where we defined by $\boldsymbol{r}$ ($\boldsymbol{\eta}_i$) the Bloch vector associated to ρ (E_i), $\boldsymbol{\lambda} = \{\hat{\lambda}_j\}_{j=1}^{d^2-1}$ is a vector of the generators of $\mathrm{SU}(d)$ with the orthogonality relation $\mathrm{tr}\,\hat{\lambda}_i\hat{\lambda}_j = 2\delta_{ij}$, and the constants a_i satisfy $\sum_i 1/m_i = 1$. Then, the half-spaces $C_i(f_i)$ can be expressed as inequalities

$$1 + (d-1)\boldsymbol{r}\cdot\boldsymbol{\eta}_i \leq m_i\left[f_i + \gamma_N(f_i, \delta_i)\right]. \tag{4.5.58}$$

In the parameter space of Bloch vectors $\boldsymbol{r}$, the confidence region is a polytope in $\mathbb{R}^{d^2-1}$. If we use a SIC-POVM, the polytope is a simplex in this space. For Pauli measurements on a qubit, the resulting polytope is a rectangular cuboid, a facet for each of the six measurement outcomes.

Exercise 4.5.1. *Assume we do tomography of a qubit state with a local Pauli basis measurement scheme. Will the shape of the confidence polytope* (4.5.55) *depend on the true state? If so, what are the conditions on it so that we obtain a regular cube?*

One of the advantages of this construction is that we can refine the region in a specified direction by performing more measurements along it, since the more measurements we do in said direction the smaller the buffer γ_N is.

A final observation: the comparison between this confidence region and the one of Section 4.5.1 is not easy. Which one is smaller for the same N and δ? For starters, one is a polytope and the other is a "ball". In addition, the size of the latter is reported in terms of ∞-norm or trace distance and is data independent, whereas that is not the case for the polytope. A way to achieve a rigorous comparison is to quantify the power of confidence regions in terms of operational distinguishability tasks, e.g., how many copies do we need to statistically distinguish (i.e., non-intersecting regions) two different states with each method? Following these ideas, a comparative analysis of different confidence regions for state tomography was recently carried out in Ref. [35].

4.5.3 Single-shot measurement schemes

So far we have discussed measurement schemes where we measure each copy of ρ individually. However, quantum mechanics allows for more general, *collective* measurements. Typically, collective measurements yield better performance than local ones in quantum information processing tasks, and tomography is no different. Without entering into too much details, we will report here the main tools used and the results (for more details, see [36, 37]).

An important difference of collective measurements with respect to the local schemes that we have seen is that these are *single-shot* measurements, that is, we measure $\rho^{\otimes N}$ to obtain a single instance of a

stochastic action. Thus, there are no frequencies, just a single outcome. We can still model a collective POVM by a set of operators $\{M_\sigma\}$ acting on $\mathcal{H}_d^{\otimes N}$, where the outcome σ is directly our estimate for ρ.

When optimizing over all possible collective measurements, one has to exploit symmetries. The first obvious symmetry is that the POVM can be chosen to be invariant under the symmetric group of permutations S_N. The second one is that our POVM should not perform differently if we replace ρ by $U\rho U^\dagger$ (we do not assume any distribution over the states), hence

$$M_{U\sigma U^\dagger} = (U^\dagger)^{\otimes N} M_\sigma U^{\otimes N}\,. \tag{4.5.59}$$

This second property of our POVM is referred to as group covariance (the group here is $\mathrm{SU}(d)^{\otimes N}$). These symmetries impose a lot of structure in the POVM to be optimized without compromising its performance, and the particular structure that stems from the interplay between the groups $\mathrm{SU}(d)^{\otimes N}$ and S_N is given by the so-called *Schur-Weyl duality*.

For fixed error bounds in trace distance, ϵ, or fidelity, τ, the question is how many copies do we need to measure to guarantee said bounds with high probability using the optimal measurement.[8] In the first case, the minimum number of copies needed to achieve $\|\hat{\rho}-\rho\|_1 \leq \epsilon$ is $N = O(d^2/\epsilon^2)$. If the error bound is in fidelity $F(\rho,\sigma) = \mathrm{tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}$,[9] then $F(\hat{\rho},\rho) \geq 1-\tau$ can be guaranteed using $N = O(d/\tau)$ copies. We cannot hope to do better than this.

4.6 Alternative schemes for quantum state tomography

One could say that the state of a qubit, say $\alpha|0\rangle + \beta|1\rangle$, contains infinite information, for the possible values of the complex parameters α, β are infinite. Of course, this is unjustified in the physical reality since we will never be able to specify α and β with infinite precision. But, as we have seen, even when we are only hoping for approximations, if d increases exponentially with the number of qubits in our system, the necessary number of state preparations and measurements to perform tomography also grows exponentially. We may argue that the root of the problem lies in the fact that the task that tomography pretends to solve is too ambitious.

Indeed, tomography aims at getting an estimate of a density matrix of a completely unknown quantum state. In other words, it aims at being able to predict the answer to *any question* that we may wish to ask in the future to our system, without assuming *anything* about it (with the exception of its dimension in most cases). Thus, there are two ways to come up with estimation schemes that are more efficient than full tomography: by focusing on highly-structured classes such as low-rank states (compressed sensing), matrix product states (MPS tomography) or neural-network states (NN tomography); and by being contempt with predicting answers to only a pre-specified and finite set of questions. We shortly review some results along these two approaches.

Compressed sensing

Intuitively, if a density matrix has small rank, then it should be easier to estimate from experimental data, since in this case only a few eigenvectors need to be learned. This is the idea of *compressed sensing* [38, 39], a method designed to work with Pauli observables. By forcing the estimator to yield low-rank density matrices, in the event of ρ actually having low rank, compressed sensing achieves estimates with small error using much fewer copies. More precisely, we construct the estimator by the optimization problem

$$\hat{\rho}_{\mathrm{CS}} = \underset{X\in\mathcal{L}(\mathcal{H}_d)}{\mathrm{argmin}} \|X\|_1 \quad \text{s.t.} \quad \|\mathcal{T}^\dagger(\mathcal{T}(X) - \boldsymbol{f})\|_\infty \leq \lambda\,, \tag{4.6.61}$$

[8] Here "high probability" means that the probability will tend to one as N increases. This is less informative that the probabilities reported in the previous sections. If one wants confidence level $1-\delta$, one needs to increase N by $\log(1/\delta)$ copies.

[9] The fidelity and the trace distance are related via

$$1 - F(\cdot) \leq \frac{1}{2}\|\cdot\|_1 \leq \sqrt{1-F^2(\cdot)}\,. \tag{4.5.60}$$

where $\mathcal{T}$ is the map associated with Pauli observables (see Section 4.3.2), and λ is some positive constant. Then, we can guarantee with high probability that

$$\|\hat{\rho}_{\mathrm{CS}} - \rho\|_1 \leq C_0 r\lambda + C_1 \|\rho_c\|_1 \tag{4.6.62}$$

using $N = O(r^2 d^2 \log d)$ copies, where C_0, C_1 are positive constants. We write for any state $\rho = \rho_r + \rho_c$, where ρ_r is the best rank-r approximation of ρ (largest r eigenvalues and eigenvectors), and the complement ρ_c is what appears above. So, as long as ρ is indeed well approximated by ρ_r, this is a good error bound.

Shadow tomography & classical shadows

It is natural to think that we will only ever be interested in asking certain questions about a quantum state, not the infinitely many possible ones. *Shadow tomography* [40] formalizes the problem of finding a good estimate with respect to predictions on the outcomes of a finite set of future measurements.

The problem is as follows: given a d-dimensional unknown state ρ and known two-outcome measurements $\{E_1, \ldots, E_M\}$, each of which accepts ρ with probability $\operatorname{tr} E_i\rho$ and rejects with probability $1 - \operatorname{tr} E_i\rho$, output numbers $b_1, \ldots, b_M \in [0,1]$ such that $|b_i - \operatorname{tr} E_i\rho| \leq \epsilon$ for all i, with success probability at least $1-\delta$. Do this by measuring $\rho^{\otimes N}$, where $N(d, M, \epsilon, \delta)$ is as small as possible. As shown in [40], this problem is solvable using $N = \tilde{O}(\frac{\log 1/\delta}{\epsilon^4} \log^4 M \log d)$ copies, which sharply contrasts with the (optimal) sample complexity of $O(d^2/\epsilon^2)$ that we saw in Section 4.5.3 (the notation $\tilde{O}$ means that polylogarithmic factors are ignored). Unfortunately, implementing the protocol of [40] is still very demanding: a collective, entangled measurement of all available copies is required, which needs exponentially long quantum circuits and keeping all copies coherently in a quantum memory.

Closely related to the mindset of shadow tomography is the framework of *classical shadows* [41]. It can be used to very efficiently predict values of linear functions of the state, of which the above shadow tomography protocol is an example (also some non-linear ones). Many physically relevant properties of a quantum state correspond indeed to expected values $o_i(\rho) = \operatorname{tr} O_i\rho$ for some observables O_i, e.g., the fidelity w.r.t. a target pure state, the value of an entanglement witness, or the probability distribution of a measurement. Crucially, the sample complexity of classical shadows does not directly depend on the dimension of the system, but on the *number of target functions* M and the *structure of the observables* O_i through a so-called *shadow norm* $\|O_i\|^2_{\mathrm{shadow}}$.

The procedure is as follows: select at random a unitary U from some specified ensemble $\mathcal{U}$, apply it to a multi-qubit state ρ, and measure in the computational basis. Upon obtaining an outcome $b \in \{0,1\}^n$, store the operator $U^\dagger |b\rangle\langle b| U$ in a classical memory.[10] We can think of the average operator, both over unitaries and over measurement outcomes, as a channel $\rho \mapsto \mathcal{M}(\rho)$ where

$$\mathcal{M}(\rho) = \mathbb{E}\left(U^\dagger |b\rangle\langle b| U\right) = \mathbb{E}_{U\sim\mathcal{U}} \sum_{b\in\{0,1\}^n} \langle b| U\rho U^\dagger |b\rangle \left(U^\dagger |b\rangle\langle b| U\right) . \tag{4.6.63}$$

As we have seen, tomographic completeness implies that $\mathcal{M}$ has a unique inverse $\mathcal{M}^{-1}$, so we can construct the *classical shadow*

$$\hat{\rho} = \mathcal{M}^{-1}\left(U^\dagger |b\rangle\langle b| U\right) , \tag{4.6.64}$$

This is nothing else but the least squares estimator (4.3.18) in the single-shot limit. As such, it is unbiased, i.e. by linearity it fulfils $\rho = \mathbb{E}(\hat{\rho})$, but it need not be positive semidefinite. We can then make a prediction for a linear function o_i just replacing ρ by $\hat{\rho}$, that is,

$$\hat{o}_i = \operatorname{tr} O_i\hat{\rho} , \quad \mathbb{E}(\hat{o}_i) = o_i .$$

The prediction $\hat{o}_i$ has the right expectation value, and its variance obeys

$$\operatorname{Var}[\hat{o}] = \mathbb{E}\left[(\hat{o} - \mathbb{E}(\hat{o}))^2\right] \leq \|O - \tfrac{\operatorname{tr} O}{2^n}\mathbb{1}\|^2_{\mathrm{shadow}} ,$$

[10] Crucially, for certain relevant ensembles of unitaries, it is proven that this operator admits an efficient classical description.

where the norm $\|\cdot\|_{\text{shadow}}$ only depends on the observable O and the ensemble of unitaries $\mathcal{U}$ (see Appendix 4.D for details). Convergence of $\hat{o}_i$ to the desired target o_i can be boosted by concentration arguments (Chernoff-Hoeffding theorem) when taking an average over N (random and independent) classical shadows $\hat{o}_i(N,1) := \frac{1}{N}\sum_{j=1}^{N} \operatorname{tr}(O_i\hat{\rho}_j)$, which yields a sample complexity of $N \sim \operatorname{Var}[\hat{o}_i]/(\delta\epsilon^2)$ to achieve accuracy ϵ with success probability $1-\delta$. We can exponentially improve the δ scaling by, instead of doing a mean of estimators, we do a *median of means*:

$$\hat{o}_i(N,K) = \text{median}\left\{\hat{o}_i^{(1)}(N,1),\ldots,\hat{o}_i^{(K)}(N,1)\right\} \quad \text{where} \quad \hat{o}_i^{(k)} = \frac{1}{N}\sum_{j=N(k-1)+1}^{Nk} \operatorname{tr}\left(O_i\hat{\rho}_j\right)$$

for $1 \le k \le K$. We now require NK samples, but this method is much more robust in the sense that $|\hat{o}(N,K) - o| > \epsilon$ iff more than half of the empirical means individually deviate more than ϵ, which probability decreases exponentially with K. Finally, we have the following theorem:

Theorem 4.6.1 (Classical shadows, Thm. 1 in [41]) *Fix a measurement primitive $\mathcal{U}$, a collection $O_1,\ldots,O_M$ of $2^n\times 2^n$ Hermitian matrices and accuracy parameters $\epsilon,\delta\in[0,1]$. Set*

$$K = 2\log(2M/\delta) \quad \text{and} \quad N = \frac{34}{\epsilon^2}\max_{1\le i\le M}\left\|O_i - \tfrac{\operatorname{tr}O_i}{2^n}\mathbb{1}\right\|_{\text{shadow}}^2$$

Then, a collection of NK independent classical shadows allow for accurately predicting all features via median of means prediction:

$$|\hat{o}_i(N,K) - \operatorname{tr}\left(O_i\rho\right)| \le \epsilon \quad \text{for all } 1\le i\le M$$

with probability at least $1-\delta$.

In other words, by the above theorem we have that, in order to accurately predict M linear target functions $\operatorname{tr}(O_i\rho)$, we require a number of samples of the order

$$N = \mathcal{O}\left(\frac{\log M}{\epsilon^2}\max_{1\le i\le M}\|O_i - \frac{\operatorname{tr}O_i}{2^n}\mathbb{1}\|_{\text{shadow}}\right). \tag{4.6.65}$$

The shadow norm can be explicitly evaluated for particular ensembles of unitaries $\mathcal{U}$. In Ref. [41] we have two examples: the global n-qubit Clifford group $\mathrm{Cl}(2^n)$, and random single-qubit Pauli measurements:

- *Random Clifford measurements.* The n-qubit Clifford group $\mathrm{Cl}(2^n)$ is comprised by all quantum circuits of n qubits composed of CNOT, Hadamard and phase gates.[11] The group $\mathrm{Cl}(2^n)$ also comprises a *unitary 3-design*, i.e., sampling uniformly Clifford circuits reproduces the first 3 moments of the full unitary group under the Haar measure. Because of this, it can be shown that the following simple expressions for the classical shadows and the shadow norm hold [cf. Eq. (4.3.35)]:

$$\hat{\rho} = (2^n+1)U^\dagger|b\rangle\langle b|U - \mathbb{1}\,, \quad \|O - \tfrac{\operatorname{tr}O}{2^n}\mathbb{1}\|_{\text{shadow}}^2 \le 3\operatorname{tr}(O^2)\,. \tag{4.6.66}$$

- *Random Pauli measurements.* This considerably simpler class of circuits trades an easier experimental implementation for a stronger dependence of the sample complexity on the structure of O. Under this unitary ensemble, equivalent to $U = U_1\otimes U_2\otimes\ldots\otimes U_n \sim \mathcal{U} = \mathrm{Cl}(2)^{\otimes n}$, we have [cf. Eq. (4.A.89)]

$$\hat{\rho} = \bigotimes_{j=1}^{n}\left(3U_j^\dagger|b_j\rangle\langle b_j|U_j - \mathbb{1}\right), \quad \|O - \tfrac{\operatorname{tr}O}{2^n}\mathbb{1}\|_{\text{shadow}}^2 \le 4^\kappa\|O\|_\infty^2\,, \tag{4.6.67}$$

  where κ is the *locality* of the observable O, that is, the number of qubits on which it acts nontrivially.

[11] Clifford circuits are a peculiar and important class of circuits. One of their key properties is that they are efficiently simulable, even though they contain entangling gates. The Clifford gate set must be complemented with a T gate to make it universal for quantum computation.

Appendix

4.A Explicit forms of $\hat{\rho}_{\mathrm{LS}}$

In this Appendix we derive explicit forms for the least squares estimator $\hat{\rho}_{\mathrm{LS}}$ under specific measurement schemes.

4.A.1 Covariant POVMs and two-designs

The *uniform* or *covariant* POVM in d dimensions is the set of all (re-normalized) rank-1 projectors $\{d\,|v\rangle\langle v|\,\mathrm{d}v\}_{v\in\mathcal{S}^d}$, where $\mathcal{S}^d$ denotes the d-dimensional complex unit hypersphere, and $\mathrm{d}v$ is the unique unitarily invariant measure (over the unitary group $\mathrm{U}(d)$) induced by the Haar measure. Define the operator

$$F_{(k)} = \int_{\mathcal{S}^d} (|v\rangle\langle v|)^{\otimes k}\ \mathrm{d}v \in \mathcal{H}_d^{\otimes k}\,. \tag{4.A.68}$$

Since it is unitarily invariant, it commutes with every k-fold tensor product of a unitary matrix $U \in \mathrm{U}(d)$, i.e.,

$$U^{\otimes k}F_{(k)} = \int_{\mathcal{S}^d} (U\,|v\rangle\langle v|)^{\otimes k}\ \mathrm{d}v = \int_{\mathcal{S}^d} (|\tilde{v}\rangle\langle\tilde{v}|\,U)^{\otimes k}\ \mathrm{d}\tilde{v} = F_{(k)}U^{\otimes k}\,, \tag{4.A.69}$$

where we have made the change of variables $\tilde{v} = Uv$, and note that $\mathrm{d}v = \mathrm{d}\tilde{v}$. Now, Schur's lemma tells us that a matrix that commutes with every element of a group representation must be proportional to a sum of projectors onto the associated irreducible representations (irreps). In this case, we are interested in the representation of the unitary group $\mathrm{U} \mapsto U^{\otimes k}$. This representation has many irreps associated to it (these can be determined using Schur-Weyl duality), but the relevant irrep for us is the *fully symmetric subspace* $\mathrm{Sym}^{(k)} \subset (\mathbb{C}^d)^{\otimes k}$, that is, the subspace of all vectors that are invariant under permuting tensor factors. Note that $F_{(k)}$ is an average over rank-1 projectors onto vectors $|v\rangle^{\otimes k}$, and therefore its range must be contained entirely in $\mathrm{Sym}^{(k)}$. Combining this with Schur's lemma yields

$$F_{(k)} = \int_{\mathcal{S}^d} (|v\rangle\langle v|)^{\otimes k}\ \mathrm{d}v = \begin{pmatrix} d+k-1 \\ k \end{pmatrix}^{-1} P_{\mathrm{Sym}^{(k)}}\,, \quad k \in \mathbb{N}\,, \tag{4.A.70}$$

where the prefactor is just the dimension of $\mathrm{Sym}^{(k)}$ and follows from the fact that $F_{(k)}$ has unit trace.

This expression is very useful. We can use it to write, for an arbitrary operator $X \in \mathcal{H}_d$,

$$\begin{aligned}(d+1)\int_{\mathcal{S}^d} d\langle v|X|v\rangle\,|v\rangle\langle v|\,\mathrm{d}v &= (d+1)d\,\mathrm{tr}_1\left(\int_{\mathcal{S}^d}(|v\rangle\langle v|)^{\otimes 2}\ \mathrm{d}v(X\otimes\mathbb{I})\right)\\ &= 2\mathrm{tr}_1\left(P_{\mathrm{Sym}^{(2)}}(X\otimes\mathbb{I})\right),\end{aligned} \tag{4.A.71}$$

where $\mathrm{tr}_1(A\otimes B) = (\mathrm{tr}\,A)B$ denotes the partial trace over the first tensor factor. The projector onto the totally symmetric subspace of two parties has an explicit representation: $P_{\mathrm{Sym}^{(2)}} = \frac{1}{2}(\mathbb{1}+\mathbb{F})$, where $\mathbb{F}$ denotes the flip (or swap) operator, i.e., $\mathbb{F}\,|a\rangle\otimes|b\rangle = |b\rangle\otimes|a\rangle$ for all $|a\rangle, |b\rangle \in \mathbb{C}^d$.[12] Using this in

[12] Note that $\mathbb{F}$ does NOT swap tensor product factors, i.e., $\mathbb{F}(X\otimes Y) \neq Y\otimes X$, for any operators X, Y.

Eq. (4.A.71) we have

$$(d+1)\int_{\mathcal{S}^d} d\langle v|X|v\rangle\,|v\rangle\langle v|\,\mathrm{d}v = \mathrm{tr}_1((\mathbb{1}+\mathbb{F})(X\otimes\mathbb{1})) = X + \mathrm{tr}(X)\mathbb{1}\,. \tag{4.A.72}$$

These observations motivate the definition of a *two-design*. A finite set of rank-1 projectors $\{|v_i\rangle\langle v_i|\}_{i=1}^m$ is called a two-design if

$$\frac{1}{m}\sum_{i=1}^m (|v_i\rangle\langle v_i|)^{\otimes 2} = \begin{pmatrix} d+1 \\ 2 \end{pmatrix}^{-1} P_{\mathrm{Sym}^{(2)}}\,. \tag{4.A.73}$$

In a way, we can think of two-designs as "discretizations" of the sphere as far as operators (actions) of degree 2 (in the projectors composing the design) are concerned. Each two-design is proportional to a POVM $\{\frac{d}{m}|v_i\rangle\langle v_i|\}_{i=1}^m$: if we take a partial trace on the above equation, we get the completeness relation $\frac{d}{m}\sum_{i=1}^m |v_i\rangle\langle v_i| = \mathbb{1}$. Viewing these POVMs as maps $\mathcal{T}:\mathcal{H}_d\mapsto\mathbb{R}^m$ and using Eq. (4.3.21), we have

$$\mathcal{T}^\dagger\mathcal{T}(X) = \frac{d^2}{m^2}\sum_{i=1}^m \langle v_i|\,X\,|v_i\rangle\,|v_i\rangle\langle v_i| = \frac{d}{m(d+1)}[X + (\mathrm{tr}\,X)\mathbb{1}]\,, \tag{4.A.74}$$

which can be readily inverted, yielding

$$(\mathcal{T}^\dagger\mathcal{T})^{-1}(X) = \frac{m}{d}[(d+1)X - (\mathrm{tr}\,X)\mathbb{1}]\,. \tag{4.A.75}$$

The least squares estimator becomes

$$\begin{aligned}\hat{\rho}_{\mathrm{LS}} &= \left(\mathcal{T}^\dagger\mathcal{T}\right)^{-1}\left(\mathcal{T}^\dagger(\boldsymbol{f})\right) = \left(\mathcal{T}^\dagger\mathcal{T}\right)^{-1}\left(\frac{d}{m}\sum_{i=1}^m f_i\,|v_i\rangle\langle v_i|\right)\\ &= \sum_{i=1}^m f_i\left[(d+1)\,|v_i\rangle\langle v_i| - \mathrm{tr}\left(|v_i\rangle\langle v_i|\right)\mathbb{1}\right] = (d+1)\sum_{i=1}^m f_i\,|v_i\rangle\langle v_i| - \mathbb{1}\,.\end{aligned} \tag{4.A.76}$$

The swap trick.

In deriving Eq. (4.A.74), it is useful to recall a general property of the operator $\mathbb{F}$ usually called "the swap trick". For any pair of operators X, Y,

$$\mathrm{tr}[\mathbb{F}(X\otimes Y)] = \mathrm{tr}(XY)\,. \tag{4.A.77}$$

4.A.2 Pauli observables

The set of Pauli observables W_i, $i=1,\ldots,d^2$ forms a unitary operator basis, that is, any operator $X\in\mathcal{H}_{2^n}$ can be written as

$$X = \frac{1}{d}\sum_{i=1}^{d^2}\mathrm{tr}(W_iX)W_i\,. \tag{4.A.78}$$

Associating to each W_i a two-outcome POVM with elements $P_i^x = \frac{1}{2}(\mathbb{1}+(-1)^xW_i)$, $x\in\{0,1\}$, we consider the union of all these POVMs to give rise to a map $\mathcal{T}\mapsto\mathbb{R}^{2d^2}$ which fulfills

$$\mathcal{T}^\dagger\mathcal{T}(X) = \sum_{i=1}^{d^2}\sum_x \mathrm{tr}\left(P_i^xX\right)P_i^x = \sum_{i=1}^{d^2}\frac{1}{2}\left[\mathrm{tr}(X)\mathbb{1} + \mathrm{tr}\left(W_iX\right)W_i\right] = \frac{1}{2}\left[d^2\,\mathrm{tr}(X)\mathbb{1} + X\right]\,. \tag{4.A.79}$$

Inverting this expression we have

$$\left(\mathcal{T}^\dagger\mathcal{T}\right)^{-1}(X) = \frac{2}{d}X - \frac{2\,\mathrm{tr}(X)}{d^2+1}\mathbb{1}\,. \tag{4.A.80}$$

Assuming that each frequency f_i^x is calculated over $N/(d^2-1)$ samples (i.e., we distribute equally the N copies of ρ among the nontrivial Pauli observables), the LS estimator takes the form

$$\begin{aligned}\hat{\rho}_{\mathrm{LS}} &= \left(\mathcal{T}^\dagger\mathcal{T}\right)^{-1}\left(\mathcal{T}^\dagger\left(\boldsymbol{f}\right)\right) = \left(\mathcal{T}^\dagger\mathcal{T}\right)^{-1}\left(\mathbb{1}\right) + \sum_{i=2}^{d^2}\sum_x f_i^x \left(\mathcal{T}^\dagger\mathcal{T}\right)^{-1}\left(P_i^x\right) \\ &= \frac{2}{d\left(d^2+1\right)}\mathbb{1} + \sum_{i=2}^{d^2} f_i^0\left(\frac{1}{d}\left(\mathbb{1}+W_i\right) - \frac{d}{d^2+1}\mathbb{1}\right) + \sum_{i=2}^{d^2} f_i^1\left(\frac{1}{d}\left(\mathbb{1}-W_i\right) - \frac{d}{d^2+1}\right) \\ &= \frac{1}{d}\sum_{i=2}^{d^2}\left(f_i^0 - f_i^1\right)W_i + \frac{2+\sum_{i=2}^{d^2}\left(f_i^0+f_i^1\right)}{d\left(d^2+1\right)}\mathbb{1}\,,\end{aligned} \tag{4.A.81}$$

which can be further simplified by noting that $f_i^0 + f_i^1 = 1$ and hence $\sum_{i=2}^{d^2}(f_i^0+f_i^1) = d^2-1$. We finally have

$$\hat{\rho}_{\mathrm{LS}} = \frac{1}{d}\sum_{i=2}^{d^2}\left(f_i^0 - f_i^1\right)W_i + \frac{1}{d}\mathbb{1} = \frac{1}{d}\sum_{i=1}^{d^2}\left(f_i^0 - f_i^1\right)W_i\,. \tag{4.A.82}$$

4.A.3 Local Pauli measurements

Let us first consider a single qubit, i.e., $d=2$. The measurement operators are given by the six projectors $\{|(s,\pm1)\rangle\langle(s,\pm1)| = \frac{1}{2}(\mathbb{1}\pm s)\}_{s=X,Y,Z}$. These projectors form a set of 3 *mutually unbiased bases*, i.e., for $s \neq s'$ they fulfill the condition

$$|\langle(s,\pm1)|(s',\pm1)\rangle|^2 = \frac{1}{4}\operatorname{tr}[(\mathbb{1}\pm s)(\mathbb{1}\pm s')] = \frac{1}{2}\,. \tag{4.A.83}$$

Sets of vectors that correspond to sets of mutually unbiased bases form two-designs, for which Eq. (4.3.21) takes the form

$$\mathcal{T}^\dagger\mathcal{T}(X) = \sum_{s,x}\langle(s,x)|\,X\,|(s,x)\rangle\,|(s,x)\rangle\langle(s,x)| \tag{4.A.84}$$

$$= X + (\operatorname{tr}X)\mathbb{1} = 3\left(\frac{1}{3}X + \frac{2}{3}\frac{\operatorname{tr}X}{2}\mathbb{1}\right) = 3\mathcal{D}_{1/3}(X)\,, \tag{4.A.85}$$

where $\mathcal{D}_{1/3}(X) = \frac{1}{3}X + (1-\frac{1}{3})\frac{\operatorname{tr}X}{2}\mathbb{1}$ is a single-qubit depolarizing channel with loss parameter $p=1/3$.

This behavior extends nicely to multiqubit systems. If $d=2^n$, each projector of the total POVM associated to a measurement setting $\boldsymbol{s}$ and measurement outcomes $\boldsymbol{x}$ has the tensor product form $|(\boldsymbol{s},\boldsymbol{x})\rangle\langle(\boldsymbol{s},\boldsymbol{x})| = \bigotimes_{i=1}^n |(s_i,x_i)\rangle\langle(s_i,x_i)|$. The map $\mathcal{T}: \mathcal{H}_{2^n} \mapsto \mathbb{R}^{3^n}\times\mathbb{R}^{2^n}$ resulting of the union of all such projectors fulfills, for any operator $X = \bigotimes_{i=1}^n X_i$,

$$\begin{aligned}\mathcal{T}^\dagger\mathcal{T}(X) &= \sum_{\boldsymbol{s},\boldsymbol{x}}\langle(\boldsymbol{s},\boldsymbol{x})|X|(\boldsymbol{s},\boldsymbol{x})\rangle\,|(\boldsymbol{s},\boldsymbol{x})\rangle\langle(\boldsymbol{s},\boldsymbol{x})| = \bigotimes_{i=1}^n\left(\sum_{s_i,x_i}\langle(s_i,x_i)|\,X_i\,|(s_i,x_i)\rangle\,|(s_i,x_i)\rangle\langle(s_i,x_i)|\right) \\ &= 3^n\bigotimes_{i=1}^n\mathcal{D}_{1/3}\left(X_i\right) = 3^n\mathcal{D}_{1/3}^{\otimes n}(X)\,,\end{aligned} \tag{4.A.86}$$

and extends to any $X\in\mathcal{H}_d$ by linearity. Since $\mathcal{D}_{1/3}(X)$ is invertible, its n-fold tensor product also is, hence

$$\left(\mathcal{T}^\dagger\mathcal{T}\right)^{-1}(X) = \frac{1}{3^n}\left(\mathcal{D}_{1/3}^{\otimes n}\right)^{-1}(X)\,. \tag{4.A.87}$$

We can now write the LS estimator as

$$\hat{\rho}_{\mathrm{LS}} = \left(\mathcal{T}^\dagger\mathcal{T}\right)^{-1}\left(\mathcal{T}^\dagger\left(\boldsymbol{f}\right)\right) = \frac{1}{3^n}\sum_{\boldsymbol{s},\boldsymbol{x}} f_{\boldsymbol{s},\boldsymbol{x}}\left(\mathcal{D}_{1/3}^{\otimes n}\right)^{-1}\left(|(\boldsymbol{s},\boldsymbol{x})\rangle\langle(\boldsymbol{s},\boldsymbol{x})|\right)\,. \tag{4.A.88}$$

What remains is to find the inversion of the qubit depolarizing channel. This is straightforwardly $\mathcal{D}_{1/3}^{-1}(X) = 3X - (\operatorname{tr} X)\mathbb{1}$, hence $\mathcal{D}_{1/3}^{-1}(|(s_i, x_i)\rangle\langle(s_i, x_i)|) = 3\,|(s_i, x_i)\rangle\langle(s_i, x_i)| - \mathbb{1}$, and

$$\hat{\rho}_{\mathrm{LS}} = \frac{1}{3^n} \sum_{\boldsymbol{s},\boldsymbol{x}} f_{\boldsymbol{s},\boldsymbol{x}} \bigotimes_{k=1}^{n} (3\,|(s_k, x_k)\rangle\langle(s_k, x_k)| - \mathbb{1})\,. \tag{4.A.89}$$

4.B Derivation of Eq. (4.5.53)

Excerpt from Ref. [33].

For structured measurements (two-designs) with m outcomes we may rewrite the LS estimator as

$$\hat{\rho}_{\mathrm{LS}} = (d+1) \sum_{i=1}^{m} f_i\,|v_i\rangle\langle v_i| - \mathbb{1} = \frac{1}{N} \sum_{i=1}^{N} X_i$$

where each X_i is an i.i.d. copy of the random matrix $X \in \mathcal{H}_d$ that assumes $(d+1)\,|v_i\rangle\langle v_i| - \mathbb{1}$ with probability $\frac{d}{m}\,\langle v_i|\rho|v_i\rangle$ for all $i \in [m]$. Unbiasedness with respect to the sample statistics ensures $\mathbb{E}\,[\hat{\rho}_{\mathrm{LS}}] = \mathbb{E}[X] = \rho$. Hence, $\hat{\rho}_{\mathrm{LS}} - \rho$ is a sum of i.i.d., centered random matrices $\frac{1}{N}\,(X_i - \mathbb{E}\,[X_i])$. These obey

$$\frac{1}{N}\,\|X_i - \mathbb{E}\,[X_i]\|_\infty = \frac{1}{N}\|(d+1)\,|v_i\rangle\langle v_i| - \mathbb{1} - \rho\|_\infty \leqslant \frac{d}{n} =: R$$

where $k \in [m]$ is arbitrary. Next, note that the random matrix X obeys

$$\mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \mathbb{E}\left[X^2\right] - \rho^2$$

and also

$$\begin{aligned}\mathbb{E}\left[X^2\right] &= \sum_{k=1}^{m} \frac{d}{m}\,\langle v_i|\rho|v_i\rangle\,((d+1)\,|v_i\rangle\langle v_i| - \mathbb{1})^2 = \frac{d\,(d^2-1)}{m} \sum_{k=1}^{m} \langle v_i|\rho|v_i\rangle\,|v_i\rangle\langle v_i| + \mathbb{1} \\ &= (d-1)(\rho + \mathbb{1}) + \mathbb{1}\end{aligned}$$

according to equation (4.A.72). This allows us to bound the variance parameter:

$$\left\| \sum_{i=1}^{N} \mathbb{E}\left[\frac{1}{n}(X - \mathbb{E}[X])\right]^2 \right\|_\infty = \frac{1}{n}\left\|(d-1)\rho + d\mathbb{1} - \rho^2\right\|_\infty \leqslant \frac{2d}{N} =: \sigma^2$$

The ratio $\frac{\sigma^2}{R} = 2$ indicates that any choice of $t \in [0, 2]$ will fall into the subgaussian regime of the matrix Bernstein inequality and Theorem 4.5.1 yields

$$\Pr\left[\|\hat{\rho}_{\mathrm{LS}} - \rho\|_\infty \geq t\right] = \Pr\left[\left\| \frac{1}{N} \sum_{i=1}^{N} (X_i - \mathbb{E}\,[X_i]) \right\|_\infty \geq t\right] \leq d\,\mathrm{e}^{-\frac{3Nt^2}{16d}}\,.$$

4.C Proof of Eq. (4.5.55)

Clopper-Pearson confidence intervals and Chernoff-Hoeffding theorem.

4.D The shadow norm in classical shadows

We saw that the variance of the estimator $\hat{o}$ of a linear property $o = \operatorname{tr} O\rho$ based on a single classical shadow obeys

$$\operatorname{Var}[\hat{o}] = \mathbb{E}\left[(\hat{o} - \mathbb{E}(\hat{o}))^2\right] \leq \|O - \tfrac{\operatorname{tr} O}{2^n}\mathbb{1}\|_{\mathrm{shadow}}^2\,. \tag{4.D.90}$$

The shadow norm in the above expression is defined as

$$\|O\|^2_{\text{shadow}} = \max_{\sigma:\sigma\geq 0,\operatorname{tr}\sigma=1} \mathbb{E}_{U\sim\mathcal{U}} \sum_{b\in\{0,1\}^n} \langle b|\, U\sigma U^\dagger\, |b\rangle\langle b|\, U\mathcal{M}^{-1}(O)U^\dagger\, |b\rangle^2 \,.$$

Proof.
Note that the variance only depends on the traceless part of O, since $\operatorname{tr}\hat{\rho} = 1$. Indeed, defining $O_0 = O - \frac{\operatorname{tr} O}{2^n}\mathbb{1}$, we have

$$\hat{o} - \mathbb{E}[\hat{o}] = \operatorname{tr}(O\hat{\rho}) - \operatorname{tr}(O\rho) = \operatorname{tr}(O_0\hat{\rho}) - \operatorname{tr}(O_0\rho)\,.$$

Also note that the inverse of $\mathcal{M}$ [cf. Eq. (4.6.63)] is self-adjoint: $\operatorname{tr}(\mathcal{M}^{-1}(X)Y) = \operatorname{tr}(X\mathcal{M}^{-1}(Y))$ for any pair of matrices X, Y with compatible dimensions. These two observations allow us to write

$$\operatorname{Var}[\hat{o}] = \mathbb{E}\left[(\hat{o} - \mathbb{E}[\hat{o}])^2\right] = \mathbb{E}\left[\left(\operatorname{tr}\left(O_0\hat{\rho}\right)\right)^2\right] - \left(\operatorname{tr}\left(O_0\mathbb{E}[\hat{\rho}]\right)\right)^2 = \mathbb{E}\left[\langle b|\, U\mathcal{M}^{-1}\left(O_0\right)U^\dagger\, |b\rangle^2\right] - \left(\operatorname{tr}\left(O_0\rho\right)\right)^2 \,.$$

Ignoring the last term in the above equation leads us to an upper bound on $\operatorname{Var}[\hat{o}]$, that we can write as

$$\operatorname{Var}[\hat{o}] \leq \mathbb{E}\left[\langle b|\, U\mathcal{M}^{-1}\left(O_0\right)U^\dagger\, |b\rangle\right]^2 = \mathbb{E}_{U\sim\mathcal{U}} \sum_{b\in\{0,1\}^n} \langle b|\, U\rho U^\dagger\, |b\rangle\, \langle b|\, U\mathcal{M}^{-1}\left(O_0\right)U^\dagger\, |b\rangle^2 \,.$$

Now we just have to maximize over all states to derive a state-independent worst-case upper bound to the above expression, and we end up with Eq. (4.D.90). ■

Bibliography

[1] Valentin Amrhein, Sander Greenland, and Blake McShane. Scientists rise up against statistical significance. *Nature*, 567:305–307, 2019.

[2] Christopher A Fuchs. Distinguishability and accessible information in quantum theory. *arXiv preprint quant-ph/9601020*, 1996.

[3] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[4] John Watrous. Lecture notes 7, https://cs.uwaterloo.ca/∼watrous/lecturenotes.html, 2017.

[5] A. Holevo. *Journal of Multivariate Analysis*, 3:337, 1973.

[6] R.S. Kennedy. On the optimum quantum receiver for the m-ary linearly independent pure state problem. *Res. Lab. Electron. MIT QPR*, 10:142, 1973.

[7] Nicola Dalla Pozza and Gianfranco Pierobon. Optimality of square-root measurements in quantum state discrimination. *Physical Review A*, 91(4):042334, apr 2015.

[8] Gael Sentís, Emilio Bagan, John Calsamiglia, Giulio Chiribella, and Ramon Muñoz-Tapia. Quantum Change Point. *Physical Review Letters*, 117(15):150502, oct 2016.

[9] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory 2nd Edition*. Wiley-Interscience, Hoboken, N.J, 2 edition edition, July 2006.

[10] K. M. R Audenaert, M. Nussbaum, A. Szkola, and F. Verstraete. Asymptotic Error Rates in Quantum Hypothesis Testing. *0708.4282*, August 2007.

[11] Rajendra Bhatia. *Matrix Analysis*. Springer, New York, 1997 edition edition, November 1996.

[12] Horn and Johnson. *Matrix Analysis 2nd Edition Hardback*. Cambridge ; New York, 2012.

[13] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, June 1994. Google-Books-ID: LeuNXB2bl5EC.

[14] Ligong Wang and Renato Renner. One-Shot Classical-Quantum Capacity and Hypothesis Testing. *Physical Review Letters*, 108(20):200501, May 2012. Publisher: American Physical Society.

[15] Fumio Hiai and Dénes Petz. The proper formula for relative entropy and its asymptotics in quantum probability. *Communications in Mathematical Physics*, 143(1):99–114, December 1991.

[16] T. Ogawa and H. Nagaoka. Strong converse and Stein's lemma in quantum hypothesis testing. *Information Theory, IEEE Transactions on*, 46(7):2428–2433, 2000.

[17] Tom Cooney, Milán Mosonyi, and Mark M. Wilde. Strong Converse Exponents for a Quantum Channel Discrimination Problem and Quantum-Feedback-Assisted Communication. *Communications in Mathematical Physics*, 344(3):797–829, June 2016.

[18] Mario Berta, Fernando G. S. L. Brandão, and Christoph Hirche. On Composite Quantum Hypothesis Testing. *Communications in Mathematical Physics*, 385(1):55–77, July 2021. Company: Springer Distributor: Springer Institution: Springer Label: Springer Number: 1 Publisher: Springer Berlin Heidelberg.

[19] Masahito Hayashi. Asymptotics of quantum relative entropy from a representation theoretical viewpoint. *Journal of Physics A: Mathematical and General*, 34(16):3413–3419, April 2001.

[20] Y. C. Eldar and G. D. Forney. On quantum detection and the square-root measurement. *IEEE Transactions on Information Theory*, 47(3):858–872, March 2001. Conference Name: IEEE Transactions on Information Theory.

[21] H. Barnum and E. Knill. Reversing quantum dynamics with near-optimal quantum and classical fidelity. *Journal of Mathematical Physics*, 43(5):2097–2106, April 2002.

[22] Giulio Chiribella. Optimal networks for quantum metrology: semidefinite programs and product rules. *New Journal of Physics*, 14(12):125008, 2012.

[23] G. Chiribella, G. M. D'Ariano, and P. Perinotti. Quantum Circuit Architecture. *Physical Review Letters*, 101(6):060401, August 2008. Publisher: American Physical Society.

[24] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerin, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S. Humble, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung, Matthew D. Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M. Martinis. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, October 2019.

[25] Christopher A Fuchs and Rüdiger Schack. Unknown Quantum States and Operations, a Bayesian View. In Matteo Paris and Jaroslav Řeháček, editors, *Quantum State Estimation*, volume 187, chapter 5, pages 147–187. Springer, Berlin/Heidelberg, 2004.

[26] N. Bent, H. Qassim, A. A. Tahir, D. Sych, G. Leuchs, L. L. Sánchez-Soto, E. Karimi, and R. W. Boyd. Experimental Realization of Quantum Tomography of Photonic Qudits via Symmetric Informationally Complete Positive Operator-Valued Measures. *Physical Review X*, 5(4):041006, oct 2015.

[27] Z. Hradil. Quantum-state estimation. *Physical Review A*, 55(3):R1561–R1564, mar 1997.

[28] Matteo Paris and Jaroslav Řeháček, editors. *Quantum State Estimation*, volume 649 of *Lecture Notes in Physics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[29] Robin Blume-Kohout. Optimal, reliable estimation of quantum states. *New Journal of Physics*, 12(4):043034, apr 2010.

[30] Travis L. Scholten and Robin Blume-Kohout. Behavior of the maximum likelihood in quantum state tomography. *New Journal of Physics*, 20(2):023050, feb 2018.

[31] Philippe Faist and Renato Renner. Practical and Reliable Error Bars in Quantum Tomography. *Physical Review Letters*, 117(1):010404, jul 2016.

[32] Jiangwei Shang, Hui Khoon Ng, Arun Sehrawat, Xikun Li, and Berthold-Georg Englert. Optimal error regions for quantum state estimation. *New Journal of Physics*, 15(12):123026, dec 2013.

[33] M Guţă, J. Kahn, R. Kueng, and J. A. Tropp. Fast state tomography with optimal error bounds. *Journal of Physics A: Mathematical and Theoretical*, 53(20):204001, may 2020.

[34] Jinzhao Wang, Volkher B Scholz, and Renato Renner. Confidence Polytopes in Quantum State Tomography. *Physical Review Letters*, 122(19):190401, may 2019.

[35] Jessica Oliveira De Almeida, Matthias Kleinmann, and Gael Sentís. Comparison of confidence regions for quantum state tomography. *New Journal of Physics*, October 2023.

[36] Ryan O'Donnell and John Wright. Efficient quantum tomography. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing - STOC 2016*, pages 899–912, New York, New York, USA, nov 2016. ACM Press.

[37] Jeongwan Haah, Aram W. Harrow, Zhengfeng Ji, Xiaodi Wu, and Nengkun Yu. Sample-optimal tomography of quantum states. *IEEE Transactions on Information Theory*, 63(9):5628–5641, 2017.

[38] David Gross, Yi-Kai Liu, Steven T. Flammia, Stephen Becker, and Jens Eisert. Quantum State Tomography via Compressed Sensing. *Physical Review Letters*, 105(15):150401, oct 2010.

[39] Steven T. Flammia, David Gross, Yi-Kai Liu, and Jens Eisert. Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators. *New Journal of Physics*, 14(9):095022, sep 2012.

[40] Scott Aaronson. Shadow tomography of quantum states. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing - STOC 2018*, pages 325–338, New York, New York, USA, nov 2018. ACM Press.

[41] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, October 2020. arXiv: 2002.08953.

Bibliography

[1] *Probability and Statistics II*. Notes by G. Deligiannidis.

[2] *Quantum estimation for quantum technology*, M. G. A. Paris, arXiv:0804.2981v3

[3] *Precision bounds in noisy quantum metrology*, J. Kolodynski, arXiv:1409.0535v2

[4] *Nonlinear quantum metrology*, S. Boixo, PhD thesis, University of New Mexico (2008).