



## 5. Quantum Hypothesis Testing and State Discrimination

**Prabhu Tej, Syed Raunaq Ahmed, I. Reena**

*Department of Physics, Bangalore University,  
Inana Bharathi, Bengaluru, Karnataka.*

**A. R. Usha Devi, A. K. Rajagopal**

*Inspire Institute Inc., Alexandria, Virginia, USA.*

### ABSTRACT

*This lecture notes gives a basic survey of the theory of hypothesis testing of quantum states in finite dimensional Hilbert spaces. Optimal measurement strategy for testing binary quantum hypotheses, which result in minimum error probability, is discussed. Collective and individual adaptive measurement strategies in testing hypotheses in the multiple copy scenario, with various upper and lower bounds on error probability, are outlined. A brief account on quantum channel discrimination and the role of entangled states in achieving enhanced precision in the task of channel discrimination is given.*

### **1. Introduction:**

Given two quantum states  $\rho_0$  and  $\rho_1$ , estimating the true state, based on an optimal decision strategy, in favor of one of the binary hypotheses  $H_0$  or  $H_1$  is referred to as *quantum(binary) hypothesis testing*. The first step towards the mathematical description of quantum hypotheses was formulated by Helstrom [1,2].

Further progress in testing quantum hypotheses made by Yuen, Kennedy and Lax [3,4], Holevo [5], Parthasarathy [6], Hayashi [7], Kargin [8], Nussbaum and Szkola [9], Audenaert et. al., [10, 11].

A quantum system is described by a density operator  $\rho$ , which is a non-negative operator in a complex Hilbert space  $\mathcal{H}$ , with unit trace. A set consisting of finite number of positive operators  $\{E_\alpha\}$  obeying

$$E_\alpha \geq 0, \quad \sum_\alpha E_\alpha = I, \quad (1.1)$$

characterize measurement with a countable number of outcomes  $\alpha = 0, 1, 2, \dots, d$ . This set is referred to as positive operator valued measure (POVM) [12]. Every element  $E_\alpha$  of the

POVM corresponds to a measurement outcome  $\alpha$ . Measurement in a quantum state  $\rho$  results in an outcome  $\alpha$  with probability

$$p_\alpha = \text{Tr}(\rho E_\alpha). \quad (1.2)$$

In this article we confine our discussion only to finite dimensional complex Hilbert spaces.

In binary hypothesis testing, the problem is to decide, which of the two density matrices  $\rho_0$  and  $\rho_1$  is true, based on a measurement strategy leading to minimum probability of error. Suppose the hypotheses  $H_0, H_1$  are given by the quantum states  $\rho_0$  and  $\rho_1$ , with respective prior probabilities  $\Pi_0$  and  $\Pi_1$ ;  $\Pi_0 + \Pi_1 = 1$ . Then, probabilities of making incorrect decision are given by

$$p(\beta|H_\alpha) = \text{Tr}(\rho_\alpha E_\beta), \quad \alpha \neq \beta = 0,1. \quad (1.3)$$

Type I error  $p(1|H_0) = \text{Tr}(\rho_0 E_1)$  is the error of accepting the *alternative hypothesis*  $H_1$ , when the *null hypothesis* is true. Type II error  $p(0|H_1) = \text{Tr}(\rho_1 E_0)$  occurs when *alternative hypothesis*  $H_1$  is the true one in reality, but *null hypothesis* is accepted.

An optimal decision strategy requires one to recognize a measurement POVM  $\{E_\alpha^{\text{opt}}, \alpha = 0, 1\}$ , such that the *average probability of error*

$$P_e = \Pi_0 p(1|H_0) + \Pi_1 p(0|H_1) = \Pi_0 \text{Tr}(\rho_0 E_1) + \Pi_1 \text{Tr}(\rho_1 E_0) \quad (1.4)$$

is minimum. It may be noted that when  $\rho_0$  and  $\rho_1$  commute with each other, the problem reduces to the testing of hypotheses based on classical statistical decision strategy. The optimal decision in the classical hypothesis test is realized by the *maximum-likelihood decision rule* [2].

In the case when null hypothesis  $H_0$  is assigned to  $\rho_0^{\otimes M}$  (i.e., tensor product of  $M$  copies of the state  $\rho_0$ ), and the alternative hypothesis to the tensor product  $\rho_1^{\otimes M}$ , the *asymptotic error rate*, realized in the limit of  $M \rightarrow \infty$ , is of interest [10, 13]. In the classical setting, the error probability in distinguishing two probability distributions  $p_0(\alpha)$  and  $p_1(\alpha)$  decreases exponentially with the increase of the number  $M$  of statistical trials i.e.,

$$P_e^{(M)} \sim e^{-M\xi(p_0, p_1)}. \quad (1.5)$$

Here,  $\xi(p_0, p_1) > 0$  denotes the error rate exponent. More specifically, in an optimal hypothesis test, the probability of error  $P_e^{(M)}$  decreases exponentially with the increase of the number  $M$  of statistical trials. Chernoff [14] derived the following expression

$$\xi_{\text{CB}} = -\lim_{M \rightarrow \infty} \left( \frac{1}{M} \log P_{e, \text{CB}}^{(M)} \right) = -\log \inf_{s \in [0,1]} \sum_{\alpha} [p_0^s(\alpha) p_1^{1-s}(\alpha)], \quad (1.6)$$

for the error rate exponent, which holds exactly in the asymptotic limit of  $M \rightarrow \infty$ . The error rate exponent  $\xi_{\text{CB}}$  gives the asymptotic efficiency of testing classical hypotheses. Moreover, for finite number of trials, one obtains a Chernoff upper bound  $P_{e,\text{CB}}^{(M)} \geq P_e^{(M)}$  on the probability of error  $P_e^{(M)}$ .

A quantum generalization of the Chernoff's result remained unsolved for long time. Various lower and upper bounds on the optimal error exponent in terms of fidelity between the two density operators  $\rho_0, \rho_1$  were identified [8]. Nussbaum and Szkola [9], and Audenaert et. al. [13] settled the issue by identifying the quantum Chernoff bound

$$\xi_{\text{QCB}} = -\lim_{M \rightarrow \infty} \left( \frac{1}{M} \log P_{e,\text{QCB}}^{(M)} \right) = -\log \inf_{s \in [0,1]} \text{Tr}(\rho_0^s \rho_1^{1-s}), \quad (1.7)$$

where  $P_{e,\text{QCB}}^{(M)}$  offers an lower bound on probability of error  $P_e^{(M)}$ .

In order to arrive at a decision with minimum error probability one has to choose optimal measurements for discriminating the states  $\rho_0^{\otimes M}$  and  $\rho_1^{\otimes M}$ . Different measurement strategies employed have been classified into (i) collective measurements, where a single POVM is employed to distinguish  $M$  copies of the states  $\rho_0$  and  $\rho_1$  and (ii) individual measurements [15] performed on each copy of state. As collective measurements, with large number of copies  $M$ , are hard to achieve in experimental implementation, individual measurement strategies are preferred. It has been shown [15, 16] that individual adaptive measurements, where a sequence of individual measurements designed such that a measurement on any copy is optimized based on the outcome obtained in previous measurement on the previous copy of the sequence. Such adaptive individual measurement strategies are shown to result in the same precision as that of the collective strategy [15].

In this paper, we present an overview of quantum state discrimination based on binary hypothesis testing both in the single copy and the multiple copy scenario. We illustrate, with the help of an example, an alternate approach termed as *unambiguous state discrimination*, which is employed for quantum state discrimination.

A discussion on collective and adaptive measurements in the multiple copy situation, with various upper and lower bounds on error probability is given in Sec. 3. In Sec. 4 an overview of quantum channel discrimination and the role of entangled states in enhancing precision in the task of channel discrimination is presented. A brief summary is given in Sec. 5.

## 2. Quantum Hypothesis Testing and State Discrimination:

Suppose the hypotheses  $H_\alpha, \alpha = 0,1$  are assigned to the quantum states characterized by their density operators  $\rho_\alpha, \alpha = 0,1$  respectively and measurements  $\{E_\beta, \beta = 0, 1, 2 \dots\}$  are employed to identify which is the *true* state. Let  $p(\beta|H_\alpha), \beta \neq \alpha$  denote the probability with which the hypothesis  $\beta$  is declared to be correct, while in fact  $\alpha$  is the true one. Associating an outcome  $\beta$  with the measurement  $E_\beta$ , the probability of error in discriminating the states  $\rho_0, \rho_1$  is given by (see (1.1),(1.2) )

$$p(\beta|H_\alpha) = \text{Tr}(\rho_\alpha E_\beta), \quad \sum_{\beta=0,1} p(\beta|H_\alpha) = 1.$$

If, with optimal measurements, one can achieve

$$p(\beta|H_\alpha) = \delta_{\alpha\beta} = \begin{cases} 0, & \text{if } \alpha \neq \beta \\ 1, & \text{if } \alpha = \beta, \end{cases} \quad (2.1)$$

then it is possible to arrive at a *correct* decision and discriminate the two quantum states  $\rho_0$ , and  $\rho_1$  with *no error*. In the special case of orthogonal quantum states  $\text{Tr}(\rho_0\rho_1) = 0$ , the conditions (1.3) can be expressed in the form of a  $2 \times 2$  matrix,

$$\mathbb{P} = \begin{pmatrix} \text{Tr}(\rho_0 E_0) & \text{Tr}(\rho_0 E_1) \\ \text{Tr}(\rho_1 E_0) & \text{Tr}(\rho_1 E_1) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

and one concludes that *orthogonal quantum states can be discriminated perfectly*. On the other hand, discrimination of non-orthogonal states can only be done with an error. In order to illustrate this, we consider an example of two pure non-orthogonal states

$$\rho_i = |\psi_\alpha\rangle\langle\psi_\alpha|, \quad \alpha = 0,1,$$

with  $\langle\psi_0|\psi_1\rangle \neq 0$ . Let  $E_0$  and  $E_1$  be the measurement operators used to discriminate these states. Suppose

$$\text{Tr}(\rho_0 E_0) = \langle\psi_0|E_0|\psi_0\rangle = 1 \quad (2.2a)$$

$$\text{Tr}(\rho_1 E_1) = \langle\psi_1|E_1|\psi_1\rangle = 1 \quad (2.2b)$$

Based on the condition (see (1.1))

$$\sum_{\alpha=0,1} E_\alpha = I$$

on measurement operators, it is readily seen that  $\langle\psi_0|E_1|\psi_0\rangle = 0 \Rightarrow \sqrt{E_1}|\psi_0\rangle = 0$ . Then, by expressing  $|\psi_1\rangle$  as

$$|\psi_1\rangle = a|\psi_0\rangle + b|\psi_0^\perp\rangle, \\ \langle\psi_0|\psi_0^\perp\rangle = 0, |a|^2 + |b|^2 = 1, 0 < |b| < 1$$

one obtains

$$\sqrt{E_1}|\psi_1\rangle = b\sqrt{E_1}|\psi_0^\perp\rangle. \quad (2.3)$$

This in turn implies that

$$\langle \psi_1 | E_1 | \psi_1 \rangle = |b|^2 \neq 1$$

in contradiction with (2.2b).

*Remark:* For a set of orthogonal states, there exists an optimum measurement scheme leading to perfect discrimination, i.e., with zero probability of error. It is not possible to achieve perfect discrimination of non-orthogonal states in the *single copy* scenario.

In a more general setting of testing multiple hypotheses, a set of states  $\rho_\alpha$  ( $\alpha = 0, 1, \dots$ ) are given with a priori probabilities  $\Pi_\alpha$  and a *true* state is to be identified from the set of states, by using an adequate measurement strategy. Define *average cost* associated with a given strategy as follows [2]:

$$\bar{C} = \sum_{\alpha, \beta} \Pi_\alpha C_{\alpha\beta} \text{Tr}(\rho_\alpha E_\beta), \quad \sum_{\beta} E_\beta = I, \quad (2.4)$$

where  $C_{\alpha\beta}$  denotes the *cost* incurred when one arrives at a *wrong decision* (i.e., reaching a conclusion that  $\rho_\beta$  is the *true* state when, in fact,  $\rho_\alpha$  happens to be the correct one). Task is to minimize the average cost  $\bar{C}$  by adapting an optimal decision strategy.

Defining *risk operator* as,

$$R_\alpha = \sum_{\beta} C_{\alpha\beta} \text{Tr}(\rho_\alpha E_\beta), \quad (2.5)$$

one can express the average cost (2.4) as,

$$\bar{C} = \sum_{\alpha} \Pi_\alpha \text{Tr}(\rho_\alpha R_\alpha). \quad (2.6)$$

Bayes' strategy [17] is to assign the costs

$$C_{\alpha\beta} = \begin{cases} 1, & \text{if } \alpha \neq \beta, \\ 0, & \text{if } \alpha = \beta \end{cases} \quad (2.7)$$

following which the average cost reduces to the *minimum average probability of error*:

$$P_e = \min_{\{E_\beta\}} P_{\text{err}} = \min_{\{E_\beta\}} \sum_{\alpha} \Pi_\alpha \text{Tr}(\rho_\alpha E_\beta) \quad (2.8)$$

Reverting back to the case of binary hypothesis testing, we define the *Helstrom matrix* [2]:

$$\Gamma = \Pi_1 \rho_1 - \Pi_0 \rho_0. \quad (2.9)$$

Substituting  $\sum_{\alpha=0,1} E_{\alpha} = I$ , the minimum average probability of error (2.8) can be expressed as,

$$P_e = \min_{\{E_0, E_1=I-E_0\}} \frac{1}{2} \{1 + \text{Tr}[\Gamma(E_0 - E_1)]\} \quad (2.10)$$

From the spectral decomposition of the Hermitian Helstrom matrix  $\Gamma$ ,

$$\Gamma = \sum_{k_+=1}^r \lambda_{k_+} |\phi_{k_+}\rangle\langle\phi_{k_+}| + \sum_{k_-=r+1}^n \lambda_{k_-} |\phi_{k_-}\rangle\langle\phi_{k_-}| \quad (2.11)$$

in terms of the eigenstates  $|\phi_{k_+}\rangle$ , corresponding to the real positive/negative eigenvalues  $\lambda_{k_+}$ ,  $k_+ = 1, 2, \dots, r$ ;  $k_- = r + 1, r + 2, \dots, n$ , we obtain,

$$P_e = \min_{\{E_0, E_1\}} \frac{1}{2} \left[ 1 + \sum_{k_+=1}^r \lambda_{k_+} \langle\phi_{k_+}|(E_0 - E_1)|\phi_{k_+}\rangle + \sum_{k_-=r+1}^n \lambda_{k_-} \langle\phi_{k_-}|(E_0 - E_1)|\phi_{k_-}\rangle \right] \quad (2.12)$$

An optimal choice of measurement  $\{E_0, E_1 = I - E_0\}$  turns out to be,

$$E_0 = \sum_{k_+=1}^r |\phi_{k_+}\rangle\langle\phi_{k_+}|, \quad E_1 = I - E_0. \quad (2.13)$$

Thus, one obtains the *minimum* average error probability as

$$P_e = \min_{\{E_0, E_1\}} P_{\text{err}} = \frac{1}{2} (1 - \|\Gamma\|) \quad (2.14)$$

where  $\|A\|_1 = \text{Tr} \sqrt{A^\dagger A}$  denotes the trace norm of the operator  $A$ . *This result on single copy minimum probability of error* (given by (2.14)) in testing quantum binary hypotheses is attributed to Holevo & Helstrom [2, 5].

In the symmetric case of equal a priori probabilities, i.e.,  $\Pi_0 = \Pi_1 = \frac{1}{2}$ , the minimum error probability is given by

$$P_e = \frac{1}{2} \left[ 1 - \frac{1}{2} \|\rho_1 - \rho_0\| \right]. \quad (2.15)$$

- If  $\rho_0 = \rho_1$ , then  $\|\rho_1 - \rho_0\| = 0 \Rightarrow P_e = \frac{1}{2}$ , i.e., decision is completely random when the states are identical.
- Minimum probability of error  $P_e = 0$  for orthogonal states  $\rho_0$  and  $\rho_1$  for which  $\|\rho_1 - \rho_0\| = 0$  i.e., the states can be discriminated with zero error.
- For pure states  $\rho_0 = |\psi_0\rangle\langle\psi_0|$  and  $\rho_1 = |\psi_1\rangle\langle\psi_1|$ , the error probability (2.15) gets simplified:

$$P_e = \frac{1}{2} \left( 1 - \sqrt{1 - |\langle\psi_0|\psi_1\rangle|^2} \right). \quad (2.16)$$

## 2.1 Unambiguous State Discrimination:

An *unambiguous discrimination* of two quantum states with a measurement involving two elements  $E_0, E_1$  is possible only when the states are orthogonal.

In an alternative approach, termed as *unambiguous state discrimination*, introduced by Ivanovic [18], the attempt is to discriminate non-orthogonal states *unambiguously* (i.e., with zero error), but the cost that one has to pay in this scheme is due to *inconclusive result* that one ends up with.

Here, a POVM consisting of three elements  $\{E_0, E_1, E_2 = I - E_0 - E_1\}$  is chosen. Then, one identifies

$$\text{Tr}(\rho_0 E_1) = 0, \quad \text{Tr}(\rho_1 E_0) = 0. \quad (2.17)$$

But this requires an additional *inconclusive result* arising from the measurement element  $E_2 = I - E_0 - E_1$  i.e., one ends up with uncertainty because  $\text{Tr}(\rho_0 E_2) \neq 0, \text{Tr}(\rho_1 E_2) \neq 0$ . The errors arising due to inconclusive outcomes are expressed by

$$\text{Tr}(\rho_0 E_2) = 1 - q_0, \quad \text{Tr}(\rho_1 E_2) = 1 - q_1, \quad 0 \leq q_0, q_1 \leq 1 \quad (2.18)$$

Using (2.17), and substituting  $E_0 + E_1 + E_2 = I$ , it follows that,

$$\begin{aligned} \text{Tr}(\rho_0 E_0) &= \text{Tr}(\rho_0 \{I - E_1 - E_2\}) = q_0, \\ \text{Tr}(\rho_1 E_1) &= \text{Tr}(\rho_1 \{I - E_0 - E_2\}) = q_1. \end{aligned} \quad (2.19)$$

Now, consider two pure non-orthogonal states  $\rho_0 = |\psi_0\rangle\langle\psi_0|$ ,  $\rho_1 = |\psi_1\rangle\langle\psi_1|$ , occurring with a priori probabilities  $\Pi_0, \Pi_1$  respectively. A measurement scheme with *zero discrimination error*, obeying the condition (2.17) can be explicitly constructed as follows:

$$\begin{aligned} E_0 &= \frac{q_0}{|\langle\psi_0|\psi_1^\perp\rangle|^2} |\psi_1^\perp\rangle\langle\psi_1^\perp| \\ E_1 &= \frac{q_1}{|\langle\psi_0^\perp|\psi_1\rangle|^2} |\psi_0^\perp\rangle\langle\psi_0^\perp| \end{aligned} \quad (2.20)$$

where  $|\psi_0^\perp\rangle$  and  $|\psi_1^\perp\rangle$  are states orthogonal to  $|\psi_0\rangle$  and  $|\psi_1\rangle$  respectively. Then we obtain,

$$\begin{aligned} P_{\text{inconclusive}} &= \Pi_0 \text{Tr}(\rho_0 E_2) + \Pi_1 \text{Tr}(\rho_1 E_2) \\ &= \Pi_0 q_0 + \Pi_1 q_1 \end{aligned} \quad (2.21)$$

as the probability of inconclusive result. With the choice

$$\begin{aligned} q_0 &= \sqrt{\frac{\Pi_1}{\Pi_0}} |\langle \psi_0 | \psi_1 \rangle|, \\ q_1 &= \sqrt{\frac{\Pi_0}{\Pi_1}} |\langle \psi_0 | \psi_1 \rangle|. \end{aligned}$$

it may be seen that the associated probability of inconclusive result (2.21) reduces to

$$P_{\text{inconclusive}} = 2\sqrt{\Pi_0 \Pi_1} |\langle \psi_0 | \psi_1 \rangle|. \quad (2.22)$$

Furthermore, in the symmetric case  $\Pi_0 = \Pi_1 = \frac{1}{2}$ , one ends up with  $P_{\text{inconclusive}} = |\langle \psi_0 | \psi_1 \rangle|$  i.e., the error arising due to inconclusive measurement outcome is proportional to the overlap between the states and is zero only when the states are orthogonal.

**Comparison of unambiguous state discrimination with Holevo-Helstrom minimum error strategy:** Consider a simple example of discriminating two non-orthogonal states

$$|\psi_0\rangle = |0\rangle, \quad |\psi_1\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}.$$

occurring with equal a priori probabilities  $\Pi_0 = \Pi_1 = \frac{1}{2}$ .

- The error probability of inconclusive outcomes (see (2.22)) is given by

$$P_{\text{inconclusive}} = \frac{1}{\sqrt{2}} \approx 0.707.$$

- The minimum probability of error (see (2.16)) in the Holevo-Helstrom single copy discrimination scheme is given by,

$$P_e = \frac{1}{2} \left( 1 - \sqrt{\frac{1}{2}} \right) \approx 0.146$$

Thus, an experimenter testing which of the given two states is true one, ends up with 70% error if he/she adapts the unambiguous state discrimination approach. In contrast, using

Baysean strategy (which leads to the Holevo-Helstrom result (2.15) for discrimination), leads to around 15% error. This example reveals that price to be paid for an error-free or unambiguous discrimination is high, compared to that for the minimum error strategy.

### 3. Multiple Copy State Discrimination:

Testing hypotheses with multiple copies of quantum states is known to reduce error incurred [8, 11, 13]. We discuss some known results on error probabilities when  $M$  copies of the quantum states, i.e.,  $\rho_0^{\otimes M}$  and  $\rho_1^{\otimes M}$  are available for quantum hypothesis testing.

Measurement strategies with multiple copies of quantum states are broadly divided into two categories:

1. **Collective measurements:** A single measurement is performed on all the  $M$  copies of the quantum states.
2. **Individual measurements:** Each of the measurements (which may not be the same) are performed separately on individual copies.

We proceed to outline the different measurement strategies.

#### 3.1. Collective Measurements:

The *Holevo-Helstrom* result leading to the error probability (2.15) holds in the multiple copy situation too, when an optimal collective measurement is performed on  $\rho_0^{\otimes M}$  and  $\rho_1^{\otimes M}$  i.e.,

$$P_e^{(M)} = \frac{1}{2} \left[ 1 - \frac{1}{2} \|\rho_1^{\otimes M} - \rho_0^{\otimes M}\|_1 \right], \quad (3.1)$$

where we have chosen equal a priori probabilities  $\Pi_0 = \Pi_1 = \frac{1}{2}$  for the states  $\rho_0^{\otimes M}$  and  $\rho_1^{\otimes M}$  for simplicity.

We consider some special cases:

- Restricting to pure states  $\rho_0 = |\psi_0\rangle\langle\psi_0|$  and  $\rho_1 = |\psi_1\rangle\langle\psi_1|$ , the  $M$ -copy error probability (3.1) reduces to the form,

$$P_e^{(M)} = \frac{1}{2} \left[ 1 - \sqrt{1 - |\langle\psi_0|\psi_1\rangle|^{2M}} \right] \quad (3.2)$$

As  $0 < |\langle\psi_0|\psi_1\rangle| < 1$  the error probability (3.2) declines by increasing the number of copies  $M$ . For  $M \gg 1$  and  $|\langle\psi_0|\psi_1\rangle|^{2M} \ll 1$ , we obtain,

$$P_e^{(M)} \approx \frac{1}{2} \left[ 1 - \left( 1 - \frac{1}{2} |\langle\psi_0|\psi_1\rangle|^{2M} \right) \right] = \frac{1}{4} |\langle\psi_0|\psi_1\rangle|^{2M}. \quad (3.3)$$

- In the asymptotic limit of  $M \rightarrow \infty$ , the  $M$ -copy error probability declines exponentially [10, 11, 13]

$$P_e^{(M)} \sim e^{-M \xi_{\text{QCB}}} \quad \text{as } M \rightarrow \infty. \quad (3.4)$$

where the optimal error exponent  $\xi_{\text{QCB}}(\rho_0, \rho_1)$  is given by,

$$\xi_{\text{QCB}} = \inf_{s \in [0,1]} \log \text{Tr}\{\rho_0^s \rho_1^{1-s}\}. \quad (3.5)$$

- An upper bound on the  $M$ -copy error probability  $P_e^{(M)}$  of (3.1) has been established [10, 11, 13],

$$P_e^{(M)} \leq P_{\text{QCB}}^{(M)}, \quad (3.6)$$

based on the quantum Chernoff error exponent  $\xi_{\text{QCB}}$ , where the error upper bound given by

$$P_{\text{QCB}}^{(M)} = \frac{1}{2} \left( \inf_{0 \leq s \leq 1} \text{Tr}\{\rho_0^s \rho_1^{1-s}\} \right)^M \quad (3.7)$$

is referred to as the Quantum Chernoff Bound (QCB).

### 3.2. Individual Measurements:

It is known that collective measurements perform better than separate measurements done on individual copies of the  $M$ -copy state resulting in optimal state discrimination when multiple copies of the states  $\rho_0, \rho_1$  are given [19]. But, when the number of copies  $M$  is large, collective measurements are hard to implement experimentally. Thus, it is of interest to explore how far one may be able to approach results of optimal state discrimination (realized based on collective measurement strategy) by confining to individual measurements i.e., to measurements performed separately on each copy of the collective  $M$ -copy states  $\rho_0^{\otimes M}, \rho_1^{\otimes M}$ .

**Fixed individual measurements:** Consider  $\rho_0^{\otimes M}$  and  $\rho_1^{\otimes M}$  occurring with equal probabilities  $\Pi_0 = \frac{1}{2}$  and  $\Pi_1 = \frac{1}{2}$ . Consider an individual measurement scheme, where same measurement is performed *individually* on each copy of  $\rho_0^{\otimes M}, \rho_1^{\otimes M}$ . In the specific case with measurements  $E_0^{\otimes M}$  and  $E_1^{\otimes M}$ , with individual measurement operators  $E_0 = |\psi_0\rangle\langle\psi_0|$  and  $E_1 = I - |\psi_0\rangle\langle\psi_0|$ , the error probability  $P_{\text{ind}}^{(M)}$  is given by,

$$P_{\text{ind}}^{(M)} = \frac{1}{2} (\text{Tr}[\rho_0^{\otimes M} E_1^{\otimes M}] + \text{Tr}[\rho_1^{\otimes M} E_0^{\otimes M}]). \quad (3.8)$$

- If one of the states, say  $\rho_0$  is pure, i.e.,  $\rho_0^{\otimes M} = (|\psi_0\rangle\langle\psi_0|)^{\otimes M}$ , the error probability (3.8) can be simplified:

$$\begin{aligned}
 P_{\text{ind}}^{(M)} &= \frac{1}{2} (\text{Tr}[\rho_0^{\otimes M} E_1^{\otimes M}] + \text{Tr}[\rho_1^{\otimes M} E_0^{\otimes M}]) \\
 &= \frac{1}{2} (\text{Tr}[\rho_0^{\otimes M} (I - |\psi_0\rangle\langle\psi_0|^{\otimes M})] + \text{Tr}[\rho_1^{\otimes M} |\psi_0\rangle\langle\psi_0|^{\otimes M}]) \\
 &= \frac{1}{2} |\langle\psi_0|\rho_1|\psi_0\rangle|^M.
 \end{aligned} \tag{3.9}$$

- If both the states are pure, then the error probability simplifies to

$$P_{\text{ind}}^{(M)} = \frac{1}{2} |\langle\psi_0|\psi_1\rangle|^{2M}. \tag{3.10}$$

Note that the approximate value  $P_e^{(M)} \approx \frac{1}{4} |\langle\psi_0|\psi_1\rangle|^{2M}$  of  $M$ -copy error probability realized using collective measurement strategy (see (3.3)) is less than  $P_{\text{ind}}^{(M)}$ . In other words, error probability obtained using collective measurements provides a lower bound on that realized from individual fixed measurements. They both match (i.e., they approach the value 0) only in the asymptotic limit  $M \rightarrow \infty$ .

**Adaptive Measurements:** In an adaptive measurement scheme, the restriction on fixed measurement on each copy of the state is relaxed. The strategy here is to optimize the next consequent measurement by using the information gathered from the results of previous measurement. This is done in a step-by-step manner. It has been shown [15] that local adaptive measurements can reveal equally good performance as that of collective optimized measurements. Further details about adaptive measurement strategy can be found in references [15, 16, 20, 21].

**Bounds on error probability:** Recall that quantum fidelity  $F(\rho_0, \rho_1)$  defined by [12, 22, 23]

$$F(\rho_0, \rho_1) = \left[ \text{Tr} \left( \sqrt{\sqrt{\rho_0} \rho_1 \sqrt{\rho_0}} \right) \right]^2, \tag{3.11}$$

serves as a quantitative measure of how close are the states  $\rho_0$  and  $\rho_1$ . It is known that the trace norm  $\|\rho_1 - \rho_0\|_1$  is bounded by the fidelity  $F(\rho_0, \rho_1)$  as follows [12]:

$$1 - \sqrt{F(\rho_0, \rho_1)} \leq \frac{1}{2} \|\rho_1 - \rho_0\|_1 \leq \sqrt{1 - F(\rho_0, \rho_1)}. \tag{3.12}$$

Using the property  $F(\rho_0^{\otimes M}, \rho_1^{\otimes M}) = [F(\rho_0, \rho_1)]^M$ , the following upper and lower bounds [8] are realized on the optimal  $M$ -copy error probability (see (2.15)):

$$\frac{1}{2} \left( 1 - \sqrt{1 - [F(\rho_0, \rho_1)]^M} \right) \leq P_e^{(M)} \leq \frac{1}{2} \left( \sqrt{F(\rho_0, \rho_1)} \right)^M. \tag{3.13}$$

- If one of the states is pure, say  $\rho_0 = |\psi_0\rangle\langle\psi_0|$ , a strict upper bound on optimal error probability  $P_e^{(M)}$  follows:

$$P_e^{(M)} \leq \frac{1}{2} |\langle\psi_0|\psi_1\rangle|^M. \quad (3.14)$$

- When both the states are pure i.e.,  $\rho_0 = |\psi_0\rangle\langle\psi_0|$  and  $\rho_1 = |\psi_1\rangle\langle\psi_1|$ , the lower bound in (3.13) matches with the exact expression (3.2) on  $M$ -copy error probability.

Another pair of computable upper and lower bounds, referred to as quantum Bhattacharya bounds [13, 24] are found to be useful in identifying the asymptotic limit of the  $M$ -copy error probability (2.15):

$$\frac{1}{2} \left( 1 - \sqrt{1 - \left[ \text{Tr} \left( \rho_0^{\frac{1}{2}} \rho_1^{\frac{1}{2}} \right) \right]^{2M}} \right) \leq P_e^{(M)} \leq \frac{1}{2} \left[ \text{Tr} \left( \rho_0^{\frac{1}{2}} \rho_1^{\frac{1}{2}} \right) \right]^M. \quad (3.15)$$

- The upper bounds of (3.13) and (3.15) are related to each other as,

$$\begin{aligned} F(\rho_0, \rho_1) &= \left( \text{Tr} \left[ \sqrt{\sqrt{\rho_0} \rho_1 \sqrt{\rho_0}} \right] \right)^2 = \left( \text{Tr} \left[ \sqrt{\sqrt{\rho_0} \sqrt{\rho_1} \sqrt{\rho_1} \sqrt{\rho_0}} \right] \right)^2 \\ &= \left\| \rho_0^{\frac{1}{2}} \rho_1^{\frac{1}{2}} \right\|_1^2, \end{aligned}$$

(3.16)  
leading to [13]

$$\text{Tr}[\sqrt{\rho_0} \sqrt{\rho_1}] \leq \left\| \sqrt{\rho_0} \sqrt{\rho_1} \right\|_1 \equiv \sqrt{F(\rho_0, \rho_1)}. \quad (3.17)$$

Thus, one obtains the inequality constraining the  $M$ -copy error probability:

$$P_e^{(M)} \leq \frac{1}{2} \left[ \text{Tr} \left( \rho_0^{\frac{1}{2}} \rho_1^{\frac{1}{2}} \right) \right]^M \leq \frac{1}{2} \left( \sqrt{F(\rho_0, \rho_1)} \right)^M \quad (3.18)$$

#### 4. Quantum Channel Discrimination:

Suppose that an input quantum state  $\rho$  goes through channels  $\Phi_\alpha$ ,  $\alpha = 0, 1, 2, \dots$ . The channels  $\Phi_\alpha$  acting on  $\rho$  result in the output states  $\rho_\alpha$  of the channel:

$$\Phi_\alpha(\rho) = \rho_\alpha. \quad (4.1)$$

The task is to ascertain which of the channels  $\{\Phi_\alpha\}$  the state  $\rho$  went through. We confine here to discrimination of two channels  $\Phi_0, \Phi_1$ .



Fig. 1. Discrimination of quantum channels  $\Phi_0, \Phi_1$

The question of distinguishing channels  $\Phi_\alpha, \alpha = 0, 1$  by choosing an input state  $\rho$  reduces to that of detecting the output states  $\rho_0$  and  $\rho_1$  (see Fig. 1) with an appropriate measurement strategy. The single copy error-probability for binary channel discrimination is given by

$$\begin{aligned}
 P_e &= \frac{1}{2} \left( 1 - \frac{1}{2} \|\Phi_0(\rho) - \Phi_1(\rho)\|_1 \right) \\
 &= \frac{1}{2} \left( 1 - \frac{1}{2} \|\rho_0 - \rho_1\|_1 \right). \tag{4.2}
 \end{aligned}$$

An optimization over a set of all input states  $\rho$  leads to the minimum error probability of discriminating the two channels  $\Phi_0, \Phi_1$  i.e.,

$$\min_{\{\rho \in \mathcal{H}\}} P_e = \frac{1}{2} \left( 1 - \max_{\{\rho \in \mathcal{H}\}} \frac{1}{2} \|\Phi_0(\rho) - \Phi_1(\rho)\|_1 \right). \tag{4.3}$$

#### 4.1. Entanglement As a Resource for Channel Discrimination:

Consider a composite bipartite state  $\rho_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$  as an input to the channel(s)  $\Phi_0$  ( $\Phi_1$ ). The channels  $\Phi_0, \Phi_1$  are designed so as to act only on one of the subsystems, say  $\rho_A = \text{Tr}_B(\rho_{AB})$ . It is convenient to employ the notation  $\Phi_0 = \Phi_0^{(A)}$  and  $\Phi_1 = \Phi_1^{(A)}$ . Action of the channels on the input state  $\rho_{AB}$  is expressed as follows:

$$\begin{aligned}
 [\Phi_0 \otimes \mathbb{1}](\rho_{AB}) &= \rho_{AB}^{(0)} \\
 [\Phi_1 \otimes \mathbb{1}](\rho_{AB}) &= \rho_{AB}^{(1)}.
 \end{aligned}$$

Here  $\mathbb{1}$  denotes the identity channel.

The error probability  $P_e$  of discriminating binary channels, in a single evaluation, is given by,

$$\begin{aligned}
 P_e &= \frac{1}{2} \left( 1 - \frac{1}{2} \left\| \rho_{AB}^{(0)} - \rho_{AB}^{(1)} \right\| \right) \\
 &= \frac{1}{2} \left( 1 - \frac{1}{2} \left\| [\Phi_0 \otimes \mathbb{1}](\rho_{AB}) - [\Phi_1 \otimes \mathbb{1}](\rho_{AB}) \right\| \right). \tag{4.4}
 \end{aligned}$$

and the minimum error probability is obtained by optimizing over the set of all input bipartite states  $\rho_{AB}$ ,

$$\begin{aligned}
 & \min_{\{\rho_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B\}} P_e \\
 &= \frac{1}{2} \left( 1 - \max_{\{\rho_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B\}} \frac{1}{2} \left\| [\Phi_0^{(A)} \otimes I^{(B)}](\rho_{AB}) - [\Phi_1^{(A)} \otimes I^{(B)}](\rho_{AB}) \right\| \right) \tag{4.5}
 \end{aligned}$$

$$= \frac{1}{2} \left( 1 - \frac{1}{2} \|\Phi_0 - \Phi_1\|_\diamond \right) \tag{4.6}$$

where

$$\|\Phi_0 - \Phi_1\|_\diamond = \max_{\{\rho_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B\}} \left\| [\Phi_0^{(A)} \otimes I^{(B)}](\rho_{AB}) - [\Phi_1^{(A)} \otimes I^{(B)}](\rho_{AB}) \right\|$$

is referred to as the *diamond* norm [25, 26].

An optimization over the set of all pure bipartite states is enough [27, 28] for achieving minimum error probability in (4.5). In the case of single-shot channel discrimination, Piani and Watrous [28] have shown that

$$\begin{aligned}
 & \max_{\{\rho_{AB}^{(\text{sep})} \in \mathcal{H}_A \otimes \mathcal{H}_B\}} \left\| [\Phi_0^{(A)} \otimes I^{(B)}](\rho_{AB}^{(\text{sep})}) - [\Phi_1^{(A)} \otimes I^{(B)}](\rho_{AB}^{(\text{sep})}) \right\| = \max_{\rho \in \mathcal{H}_A} \left\| \Phi_0^{(A)}(\rho) - \Phi_1^{(A)}(\rho) \right\| \tag{4.7}
 \end{aligned}$$

Here an optimization is carried out by restricting only to the set of all separable states

$$\rho_{AB}^{(\text{sep})} = \sum_i p_i \rho_{A,i} \otimes \rho_{B,i}, \quad 0 \leq p_i \leq 1, \quad \sum_i p_i = 1.$$

In other words, there is *no advantage* in employing a *separable* composite bipartite state  $\rho_{AB}^{(\text{sep})}$  as input of the channels, because the probability of error does not get reduced beyond the one achievable using any input state  $\rho$  belonging to the Hilbert space  $\mathcal{H}_A$  itself. On the

otherhand, it has been identified that entangled input states help in channel discrimination with improved precision [27, 29–34], where it has been established that with a choice of entangled input state, it is possible to reduce channel discrimination error probability. More specifically, Piani and Watrous [28] proved

$$\|\Phi_0 - \Phi_1\|_* \geq \max_{\{\rho_{AB}^{(\text{sep})} \in \mathcal{H}_A \otimes \mathcal{H}_B\}} \left\| \left[ \Phi_0^{(A)} \otimes I^{(B)} \right] \left( \rho_{AB}^{(\text{sep})} \right) - \left[ \Phi_1^{(A)} \otimes I^{(B)} \right] \left( \rho_{AB}^{(\text{sep})} \right) \right\| \quad (4.8)$$

in the case of single-shot discrimination of the channels. In other words, given an entangled state, it is always possible to find a pair of quantum channels such that the error probability of single-shot channel discrimination gets minimized. For a detailed mathematical treatment on channel discrimination see Ref. [26].

**Discrimination of identity and completely depolarizing channels:** Let us consider an example[27, 32], where a completely depolarizing channel and an identity channel, labelled respectively as channel 0 and channel 1, are to be discriminated based on their action on a pure input state  $|\psi\rangle$  belonging to a finite  $d$  dimensional Hilbert space  $\mathcal{H}_d$ . The output states of the channels are given by,

$$\rho_0 = \Phi_0(\rho) = \frac{I}{d}$$

$$\rho_1 = \Phi_1(\rho) = |\psi\rangle\langle\psi|. \quad (4.9)$$

The single copy error probability in distinguishing  $\rho_0$  and  $\rho_1$  is readily found to be,

$$P_{e,|\psi\rangle} = \frac{1}{2} \left( 1 - \frac{1}{2} \left\| \frac{I}{d} - |\psi\rangle\langle\psi| \right\|_1 \right)$$

$$= \frac{1}{2} \left( 1 - \frac{1}{2} \left[ \left| \frac{1}{d} - 1 \right| + \frac{d-1}{d} \right] \right) = \frac{1}{2d}. \quad (4.10)$$

Let us consider a maximally entangled  $d \times d$  state,

$$|\psi_{AB}\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d |i_A, i_B\rangle, \quad (4.11)$$

as input of the channels  $\Phi_0 \otimes \mathbb{1}, \Phi_1 \otimes \mathbb{1}$ . The output states are then found to be,

$$\rho_{AB}^{(0)} = (\Phi_0 \otimes \mathbb{1}) |\psi_{AB}\rangle\langle\psi_{AB}|$$

$$= \frac{I}{d} \otimes \text{Tr}_A[|\psi_{AB}\rangle\langle\psi_{AB}|] = \frac{I \otimes I}{d^2}$$

$$\rho_{AB}^{(1)} = (\Phi_1 \otimes \mathbb{1})|\psi_{AB}\rangle\langle\psi_{AB}| = |\psi_{AB}\rangle\langle\psi_{AB}|. \quad (4.12)$$

The error-probability in discriminating the two channels, with a maximally entangled state is equal to

$$P_{e,|\psi_{AB}\rangle} = \frac{1}{2d^2}. \quad (4.13)$$

This clearly shows that maximally entangled state (4.11) is advantageous in the discrimination of completely depolarizing and identity channels [27, 32].

## 5. Summary:

This article presents an overview of quantum state discrimination based on binary hypothesis testing. A brief outline on *Unambiguous state discrimination*, an alternate approach developed for quantum state discrimination, is given, with the help of an illustrative example. Collective and adaptive measurements strategies employed in the case of multiple copy hypothesis testing are described. A discussion on computable upper and lower bounds on error probability in the multiple copy scenario and the error rate exponent in the asymptotic limit is given. Furthermore, quantum channel discrimination and the role of entangled states in enhancing precision in the task of channel discrimination are presented.

## 6. Acknowledgements:

ARU & SRA acknowledge the support of University Grants Commission (UGC) Major Research Project (Grant No. MRP-MAJOR-PHYS-2013-29318), Government of India and JPT thanks UGC-BSR, India for financial support when this work was in initial stages. ARU & IR are supported by the Department of Science and Technology, India (Project No. DST/ICPS/QUST/Theme-2/2019).

## 7. References:

1. W. Helstrom, *Information and Control* **10**, 254 (1967).
2. W. Helstrom, *Quantum detection and estimation theory* (Academic, New York, 1976).
3. H. P. Yuen, *Communication theory of quantum systems*, Ph.D. thesis, Submitted to the Dep. Elec. Eng., M.I.T., Res. Lab. Electron., Tech. Rep. 482 (1970).
4. H. P. Yuen, R. S. Kennedy, and M. Lax, *Proc. IEEE (Lett.)* **21** (1970).
5. S. Holevo, *Theor. Prob. Appl.* **23** (1978).
6. K. R. Parthasarathy, in *Stochastics in Finite and Infinite Dimensions*, edited by T. H. et. al. (Birkhäuser, Boston, 2001) pp. 361–377.
7. M. Hayashi, *Quantum Information: An Introduction* (Springer-Verlag, Berlin Heidelberg, 2006).
8. V. Kargin, *Ann. Stat.* **33**, 959 (2005).
9. M. Nussbaum and A. Szkoła, *Ann. Stat.* **37** (2009).
10. K. M. R. Audenaert, J. Calsamiglia, R. M. noz Tapia, E. Bagan, L. Masanes, A. Acin, and F. Verstraete, *Phys. Rev. Lett.* **98**, 160501 (2007).

11. K. M. R. Audenaert, M. Nussbaum, A. Szkoła, and F. Verstraete, *Comm. Math. Phys.* **279**, 251 (2008).
12. M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge Series on Information and the Natural Sciences (Cambridge University Press, 2000).
13. J. Calsamiglia, R. M. noz Tapia, L. Masanes, A. Acin, and E. Bagan, *Phys. Rev. A* **77**, 032311 (2008).
14. H. Chernoff, *Ann. Math. Statist.* **23**, 493 (1952).
15. Acin, E. Bagan, M. Baig, L. Masanes, and R. M. noz Tapia, *Phys. Rev. A* **71**, 032338 (2005).
16. L. Higgins, A. C. Doherty, S. D. Bartlett, G. J. Pryde, and H. M. Wiseman, *Phys. Rev. A* **83**, 052314 (2011).
17. T. Bayes, *Philosophical Transactions of Royal Society (London)* **53** (1763), reprinted in *Biometrika*, **45**, 293-315 (1958).
18. Ivanovic, *Physics Letters A* **123**, 257 (1987).
19. Peres and W. K. Wootters, *Physical Review Letters* **66** (1991).
20. W. Harrow, A. Hassidim, D. W. Debbie, and J. Watrous, *Phys. Rev. A* **81**, 032339 (2010).
21. S. Slussarenko, M. M. Weston, J.-G. Li, N. Campbell, H. M. Wiseman, and G. J. Pryde, *Phys. Rev. Lett.* **118**, 030502 (2017).
22. Uhlmann, *Rep. Math. Phys.* **9** (1976).
23. R. Josza, *Journal of Modern Optics* **41** (1994).
24. M. H. T. Ogawa, *IEEE Trans. Inform. Theory* **50** (2004).
25. Y. Kitaev, *Russian Mathematical Surveys* **52**, 1191 (1997).
26. Watrous, *Quantum Inf. Comput* **8** (2008).
27. M. F. Sacchi, *Phys. Rev. A* **71**, 062340 (2005).
28. M. Piani and J. Watrous, *Phys. Rev. Lett.* **102**, 250501 (2009).
29. M. F. Sacchi, *Phys. Rev. A* **72**, 014305 (2005).
30. S. Lloyd, *Science* **321** (2008).
31. S.-H. Tan, B. I. Erkmén, V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, S. Pirandola, and J. H. Shapiro, *Physical Review Letters* **101** (2008).
32. R. U. Devi and A. K. Rajagopal., *Phys. Rev. A* **79**, 062320 (2009).
33. Prabhu Tej, A. R. Usha Devi, and A. K. Rajagopal, *Phys. Rev. A* **87**, 052308 (2013).
34. J. Prabhu Tej, A. R. Usha Devi, H. S. Karthik, Sudha, and A. K. Rajagopal, *Phys. Rev. A* **89**, 062116 (2014).