Chapter 5
Quantum Decision Theory: Analysis and Optimization

5.1 Introduction

We consider **the transmission of classical information through a quantum channel**, where the information carrier is given by quantum states. A system that achieves this target is called *Quantum Communications system*. Like in classical communications, in quantum communications the usual configuration applies: transmitter, channel, and receiver. Analog quantum transmission systems have been considered too [1], but, as seen in the previous chapter, according to the current trend, we limit ourselves exclusively to digital systems. So Fig. 5.1 illustrates a quantum digital system, emphasizing its essential components.

In this chapter we will develop the theory of decision applied to the combination of the quantum measure and the decision element, without any specification on the nature of the quantum states. In the following chapters the quantum decision theory will be applied to the systems in which the states are physically produced by a coherent monochromatic radiation (coherent or Glauber states).

5.1.1 General Description of a Digital Transmission System

We consider the transmission of a single[1] classical symbol $A \in \mathcal{A}$. Thus, a classical source emits a symbol among K possible symbols, $A \in \mathcal{A} = \{0, 1, \ldots, K-1\}$, with assigned *a priori* probabilities

$$q_i := \mathrm{P}[A = i], \qquad i \in \mathcal{A}. \tag{5.1}$$

[1] In Sect. 4.2 we justified the advantage of dealing with a single symbol instead of a sequence of symbols.

G. Cariolaro, *Quantum Communications*, Signals and Communication Technology,
DOI 10.1007/978-3-319-15600-2_5

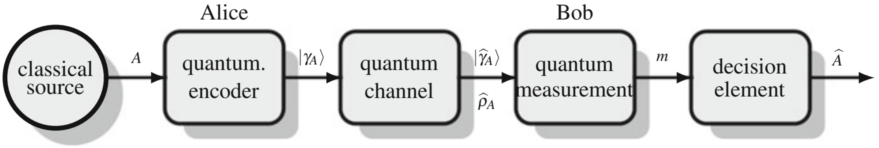


Fig. 5.1 Quantum communication system for the transmission of *classical information* through a *quantum* channel. The transmission of a single digital symbol $A \in \mathcal{A}$ is considered. In transmission a pure state $|\gamma_A\rangle$ is assumed, while in reception the quantum state may be a pure state, $|\widehat{\gamma}_A\rangle$, or a mixed state, $\widehat{\rho}_A$

The transmitter (Alice) encodes the symbol A into a quantum state $|\gamma_A\rangle$ of a Hilbert space $\mathcal{H}_T$, thus realizing the classical-to-quantum mapping $A \to |\gamma_A\rangle$. This implies that Alice is capable of preparing the quantum system $\mathcal{H}_T$ in K distinct quantum states

$$|\gamma_0\rangle, |\gamma_1\rangle, \ldots, |\gamma_{K-1}\rangle \tag{5.2}$$

which must be considered as *pure*, since they are known to Alice, because she prepares the specific state $|\gamma_i\rangle$ when the source emits the symbol $A = i$. The pure state (ket) prepared by Alice is alternatively described by the density operator $\rho_i = |\gamma_i\rangle\langle\gamma_i|$.

The channel, be it an optical fiber or the free space, modifies the density operators, introducing noise and distortion, so that the received state is in general a mixed state described by a density operator $\widehat{\rho}_A$. Then the channel performs the quantum-to-quantum mapping $\rho_A \to \widehat{\rho}_A$. As we shall see in Chap. 12, a quite general model to represent explicitly this mapping is given by the the Kraus representation [2]

$$\widehat{\rho}_A = \sum_k V_k^* \; \rho_A \; V_k \tag{5.3}$$

where $\{V_k\}$ is a class of operators.

The receiver (Bob) performs a quantum measurement on the received state $\widehat{\rho}_A$, and, to this end, he must choose a system of measurement operators $\{P_k, k \in \mathcal{M}\}$, which in general are POVM, and, in particular, projectors (see Sect. 3.8). The outcome of the measurement m is a new discrete random variable with alphabet $\mathcal{M}$, which can be seen as the *received signal*, or better, in the language of telecommunications, the *signal at the decision point*. Finally, according to the outcome m of the measurement, a decision must be made, based on a *decision criterion*, to select the symbol $\widehat{A} \in \mathcal{A}$ that was presumably transmitted. Globally, the quantum measurement combined with the decision element provides the quantum-to-classical mapping $\widehat{\rho}_A \to \widehat{A}$.

Note on symbolism. The alphabet $\mathcal{A}$ of the symbols is indicated in the form

$$\mathcal{A} = \{0, 1, \ldots, K-1\}$$

but it can take other forms (also with complex symbols) related to the modulation format. The alphabet of the measurements $\mathcal{M}$ can be different, even in cardinality, from the alphabet of the source, but as will be seen in the next section, it is not restrictive to assume that the two alphabets coincide, then, in the analysis of specific systems, we will assume $\mathcal{M} = \mathcal{A}$.

Guidelines and Preview of the Chapter

As we want to adopt a very general and complete approach, the chapter may appear long and complex. We encourage the reader to tackle it gradually, restricting the first reading to the concepts related to decision with pure states, and skipping decision based on mixed states. So the study of Chap. 7 is quite feasible, as, for its comprehension, the decision theory based on pure states is sufficient. Later on, the study can be resumed and completed, going through the decision with mixed states, a subject necessary for a full understanding of Chap. 8. Another suggestion is to read this chapter again after viewing the applications of quantum decision to quantum communications systems, developed in Chaps. 7 and 8.

We now detail the line followed in this chapter for the Quantum Decision Theory, but before we remark that this theory, here presented in the language of Telecommunications, is an important and autonomous field of Quantum Mechanics, which could be presented independently of quantum communications systems (and, in fact, in Quantum Mechanics the quantum communications systems are often ignored).

The chapter is organized in four topics.

Analysis of Quantum Decision (Sects. 5.2 to 5.7)

We begin with the *Analysis* of a general quantum communications system, where the target is the evaluation of the system's performance in terms of probabilities. Then, we deal with a specific case to let the reader become familiar with the main concepts introduced: the optimization of a binary system following *Helstrom's theory*.

Optimization of Quantum Decision (Sects. 5.8 to 5.11)

We give a general formulation of *Quantum Optimization*, which has the target of finding the measurement operators that ensure the "best performance," that is, the *maximum correct decision probability*. Optimization may be viewed in the framework of *convex linear programming* and appears to be a formidable problem because the unknowns are the measurement operators, which have severe constraints. Two main results are Holevo's and Kennedy's theorems, which provide conditions that the measurement operators must meet to be optimal.

Geometrically Uniform Symmetry (GUS) (Sects. 5.13 and 5.14)

The GUS is verified in several quantum communications systems and facilitates, in general, analysis and performance evaluation, in particular, optimization and suboptimization. We first consider the GUS for pure states and then for mixed states.

State Compression in Quantum Detection (Sect. 5.15)

In general, quantum states and measurement operators are "redundant," but it is possible and convenient to perform a compression onto a "compressed" space, where redundancy is removed. Quantum detection can be reformulated in the "compressed" space, getting properties simpler than in the original Hilbert space.

5.2 Analysis of a Quantum Communications System

In the *Analysis* of a general quantum communications system, it is assumed that both the transmitter (Alice) and the receiver (Bob) are assigned and the target consists in finding the statistical description of the system's behavior **in terms of probabilities**, exactly as in a classical communications system.

In a quantum system, probabilities come into play in two ways, and, in fact we have two sources of randomness. One is related to the source of information, which emits a symbol $A = i \in \mathcal{A}$ with a given probability $q_i = \mathrm{P}[A = i]$, which is called *a priori probabilities*. Therefore, we have a probability distribution q_i, $i \in \mathcal{A}$ of the random variable A. The other form of randomness is related to the quantum measurement, which produces another random variable $m \in \mathcal{M}$, whose statistical description is provided by Postulate 3 of Quantum Mechanics seen in Sect. 3.5. Then the Analysis of the system will be necessarily based on Probability Theory.

Next, we have to study the viewpoint of Bob, who receives a "signal" and performs the measurement. About this we can make two different hypotheses:

(1) The signal has not been contaminated, so that Bob receives the state $|\gamma_A\rangle$ that Alice associated to the symbol A.
(2) The signal has been contaminated by the channel and by *thermal noise* (also called *background noise*), and therefore Bob does not see the pure state $|\gamma_A\rangle$ any more, but instead a mixed state represented by a density operator ρ_A.

The two cases are illustrated in Fig. 5.2.

Case (1) corresponds to a transmission with an ideal noiseless channel, whereas case (2) accounts for the fact that the channel can fail to be ideal and noiseless. It is important to observe that also in case (1) Bob will not be able to make with certainty a correct decision, because it would be based on quantum measurements, which, as already seen, do not give error-free results; in the classical case, the randomness of the measurement with pure states corresponds to *shot noise*.

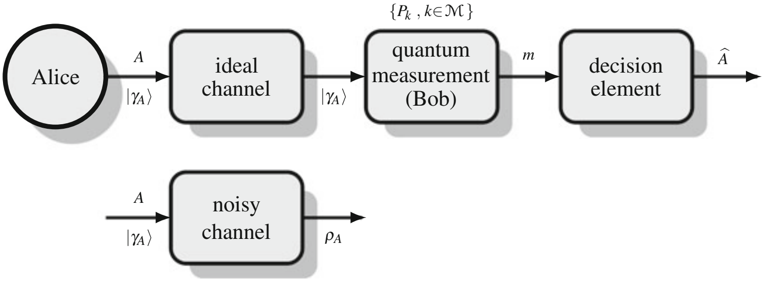


Fig. 5.2 Transmission of a classical symbol A through a quantum channel. At reception Bob performs the measurement in a quantum system in a pure state $|\gamma_A\rangle$ (ideal channel) or in a quantum system in a mixed state ρ_A (noisy channel)

The two cases can be unified considering that also in case (1), to the pure state $|\gamma_A\rangle$ one can associate the degenerate density operator $\rho_A = |\gamma_A\rangle\langle\gamma_A|$.

5.2.1 Quantum Measurement

To perform a quantum measurement, Bob chooses a *measurement operator system*

$$\{P_k, k \in \mathcal{M}\}.$$

From Postulate 3, if we know that the system under measurement is in the state $|\gamma_A\rangle$, the probability that the result of the measurement be $m = k$, is given by (see (3.26) and (3.50))

$$\boxed{\mathrm{P}[m = k|\ \gamma_A] = \langle\gamma_A|P_k|\gamma_A\rangle, \qquad k \in \mathcal{M}.} \tag{5.4}$$

Clearly, this result holds if the state $|\gamma_A\rangle$ is known with certainty (pure state). If, instead, the system state is only statistically known through the density operator ρ_A (mixed state), the probability that the result of the measurement be $m = k$ is calculated according to (see (3.32) and (3.51))

$$\boxed{\mathrm{P}[m = k|\ \rho_A] = \mathrm{Tr}[\rho_A P_k], \qquad k \in \mathcal{M}.} \tag{5.5}$$

Relation (5.5) includes relation (5.4), because it holds even when the system state is known, thus $\rho_A = |\gamma_A\rangle\langle\gamma_A|$ and then it suffices to recall the identity on the trace (2.37), to obtain (5.4) from (5.5).

In quantum communications systems, we must apply (5.4) when we neglect *thermal noise*, and (5.5) when we take it into account.

5.2.2 The Digital Channel from the Source to the Measurement

The steps that go from the transmitted symbol $A \in \mathcal{A}$ to the outcome of the measurement $m \in \mathcal{M}$ identify a *digital channel*, as shown in Fig. 5.3. The alphabet at the input of this channel is that of the possible symbols of the source

$$\mathcal{A} = \{0, 1, \ldots, K - 1\}, \tag{5.6}$$

whereas at the output we have the alphabet $\mathcal{M}$, which gives the possible outcomes of the measurements and can be indicated in the form

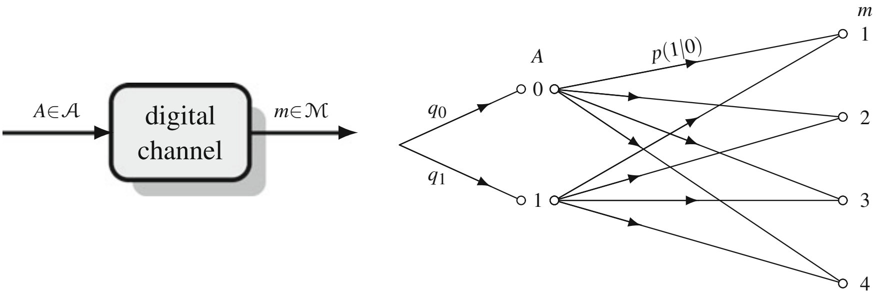


Fig. 5.3 Source-to-measurement digital channel with transition probabilities $p(k|i) = \mathrm{P}[m = k|A = i]$. In the graph, the source alphabet is $\mathcal{A} = \{0, 1\}$ and the measurement alphabet is $\mathcal{M} = \{1, 2, 3, 4\}$

$$\mathcal{M} = \{1, 2, \ldots, K'\} \tag{5.7}$$

where the cardinality K' may be different from K.

The transition probabilities of this channel are given by (5.4) or by (5.5). In fact, in the former case, thinking in terms of Probability Theory, the *event* $\{|\gamma_A\rangle = |\gamma_i\rangle\}$ coincides with the *event* $\{A = i\}$, because Alice has "prepared" the quantum system in the state $|\gamma_i\rangle$, having observed that $A = i$. Therefore, $\mathrm{P}[m = k|\ A = i] = \mathrm{P}[m = k|\ |\gamma\rangle = |\gamma_i\rangle]$ and the transition probabilities of the channel become

$$p(k|i) := \mathrm{P}[m = k|\ A = i] = \langle\gamma_i|P_k|\gamma_i\rangle, \qquad k \in \mathcal{M}, i \in \mathcal{A}. \tag{5.8a}$$

Even in the latter case, Alice has prepared the system in the state $|\gamma_A\rangle$. However, because of the presence of noise, the state is not pure any more, but it is described by the density operator ρ_A. However, at the level of events, we still have that to $\{A = i\}$ it uniquely corresponds $\{\rho_A = \rho_i\}$, thus

$$p(k|i) := \mathrm{P}[m = k|\ A = i] = \mathrm{Tr}[\rho_i P_k], \qquad k \in \mathcal{M}, i \in \mathcal{A}. \tag{5.8b}$$

As usual, (5.8b) represents the general case, as it yields (5.8a) assuming $\rho_i = |\gamma_i\rangle\langle\gamma_i|$.

It remains to observe that, for the sake of generality, we have chosen a measurement alphabet $\mathcal{M}$, in general different from the source alphabet $\mathcal{A}$ of the symbols. For example, in Fig. 5.3 we have $\mathcal{A} = \{0, 1\}$ and $\mathcal{M} = \{1, 2, 3, 4\}$. The important constraint is that the cardinality of $\mathcal{M}$ must not be smaller than that of $\mathcal{A}$

$$|\mathcal{M}| \geq |\mathcal{A}| \qquad \rightarrow \qquad K' \geq K.$$

As we will see, the two alphabets are often chosen coincident.

5.2.3 Post-measurement Decision. Correct Decision Probability

Remaining in the general case, the decision criterion after the measurement must be expressed by partitioning the measurement alphabet in correspondence with the symbol alphabet, i.e., by finding a *partition* of $\mathcal{M}$ of the type

$$\mathcal{M}_0, \mathcal{M}_1, \ldots, \mathcal{M}_{K-1}. \tag{5.9}$$

Then the *decision criterion* becomes

$$m \in \mathcal{M}_i \quad \Longleftrightarrow \quad \widehat{A} = i. \tag{5.10}$$

For example, in Fig. 5.4, where $\mathcal{A} = \{0, 1\}$ and $\mathcal{M} = \{1, 2, 3, 4\}$, we have chosen the partitions $\mathcal{M}_0 = \{1, 2\}$ and $\mathcal{M}_1 = \{3, 4\}$.

Once chosen the decision criterion, we complete the global digital channel of the quantum system, whose input is the symbol $A \in \mathcal{A}$, and output the symbol $\widehat{A} \in \mathcal{A}$ obtained after the decision (Fig. 5.4). The transition probabilities of this global channel become

$$\begin{aligned} p_c(j|i) &= \mathrm{P}[\widehat{A} = j|\ A = i] = \mathrm{P}[m \in \mathcal{M}_j|\ A = i] \\ &= \sum_{k \in \mathcal{M}_j} \mathrm{P}[m = k|\ A = i]. \end{aligned} \tag{5.11}$$

Therefore, using (5.8b), we have

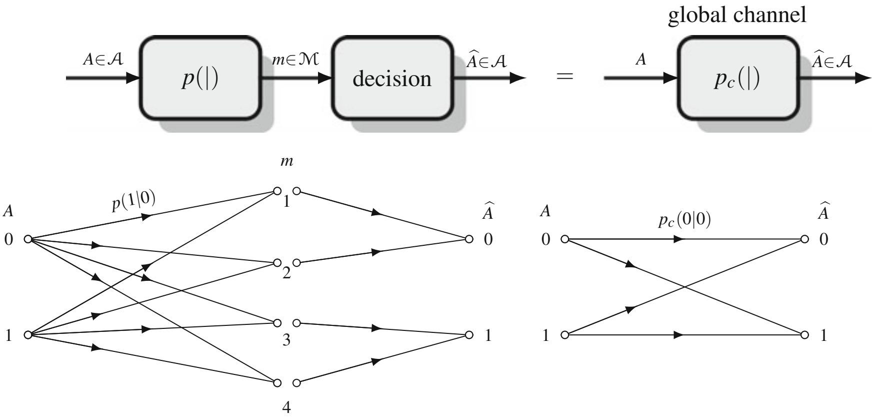


Fig. 5.4 Global digital channel with transition probability $p_c(j|i) = \mathrm{P}[\hat{A} = j|A = i]$

$$\boxed{p_c(j|i) = \sum_{k\in\mathcal{M}_j} \mathrm{Tr}[\rho_i P_k], \qquad i, j \in \mathcal{A}.} \tag{5.12}$$

From the global transition probabilities, being also known the a priori probabilities $q_i = \mathrm{P}[A = i]$, we can calculate the *correct decision probability* as

$$\begin{aligned} P_c = \mathrm{P}[\widehat{A} = A] &= \sum_{i\in\mathcal{A}} q_i \; p_c(i|i) \\ &= \sum_{i\in\mathcal{A}} \sum_{k\in\mathcal{M}_i} q_i \; \mathrm{Tr}[\rho_i P_k] \end{aligned} \tag{5.13}$$

from which we obtain the *error probability*[2] as $P_e = 1 - P_c$.

5.2.4 Combination of Measurement and Post-measurement Decision

To the purpose of optimization, the decision criterion can be combined with the system of the measurement operators.

Then, given the system of the measurement operators $\{P_k, k \in \mathcal{M}\}$, and the decision criterion determined by the partition (5.9), a set of new operators is defined as follows:

$$Q_i = \sum_{k\in\mathcal{M}_i} P_k, \qquad i \in \mathcal{A}. \tag{5.14}$$

The set of the operators $\{Q_i, i \in \mathcal{A}\}$ forms a system of POVMs, that is, with the properties (see Sect. 3.7):

(1) they are Hermitian operators, $Q_i^* = Q_i$,
(2) they are PSD, $Q_i \geq 0$,
(3) they resolve the identity, $\sum_{i\in\mathcal{A}} Q_i = I_{\mathcal{H}}$.

The proof of these properties is based on the fact that the initial operators P_k also have such properties; in particular, (3) is obtained according to

$$\sum_{i\in\mathcal{A}} Q_i = \sum_{i\in\mathcal{A}} \sum_{k\in\mathcal{M}_i} P_k = \sum_{k\in\mathcal{M}} P_k = I_{\mathcal{H}}.$$

Substituting the new operators (5.14) in (5.12) for the transition probabilities, we obtain simply

[2] In practice, the performance of a telecommunication system (classical or quantum) is often expressed in terms of the *error probability*, but in theoretical formulation it is more convenient to refer to the *correct decision probability*.

$$p_c(j|\,i) = \mathrm{Tr}[\rho_i Q_j], \qquad j, i \in \mathcal{A}. \tag{5.15}$$

Analogously, the *correct decision probability* becomes

$$\boxed{P_c = \sum_{i\in\mathcal{A}} q_i \mathrm{Tr}[\rho_i Q_i].} \tag{5.16}$$

In particular, when the system in reception is in a pure state (absence of thermal noise), letting $\rho_i = |\gamma_i\rangle\langle\gamma_i|$ we obtain

$$\boxed{P_c = \sum_{i\in\mathcal{A}} q_i \langle\gamma_i|Q_i|\gamma_i\rangle.} \tag{5.16a}$$

At this point, conceptually, the quantum measurement can be performed directly with the new measurement operators Q_i (global measurement operators), and we obtain directly, as its result, the decided symbol $\widehat{A}$, as illustrated in Fig. 5.5.

In conclusion, we have seen that in principle, in reception, we perform a quantum measurement, followed by a decision, but **it is not restrictive** to include in the measurement also the final post-measurement decision, therefore the choice to make for a good performance affects only the *global measurement operators*.

We finally remark the following statement:

Proposition 5.1 *If the measurement operators $\{P_k, k \in \mathcal{M}\}$ form a projector system, also the global operators $\{Q_i, i \in \mathcal{A}\}$ form a projector system.*

Problem 5.1 ⋆⋆ Prove Proposition 5.1. *Hint*: see Sect. 3.6.4.

Problem 5.2 ⋆⋆ *Optimization of decision element.* In a post-measurement decision the decision element is a mapping: $\mathcal{M} \to \mathcal{A}$, where $|\mathcal{M}| \geq |\mathcal{A}|$, in which every point $k \in \mathcal{M}$ must be associated to a symbol $a \in \mathcal{A}$, thus creating a partition of $\mathcal{M}$ into K sets $\mathcal{M}_a, a \in \mathcal{A}$. For given a priori probabilities $\{q_i\}$ and transition probabilities $\{p_c(j|i)\}$, one can optimize the decision element with the criterion to

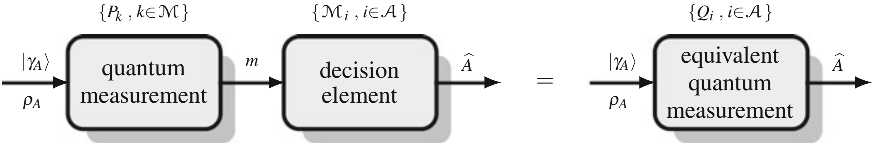

$\{P_k, k\in\mathcal{M}\}$
$\{\mathcal{M}_i, i\in\mathcal{A}\}$
$\{Q_i, i\in\mathcal{A}\}$
$|\gamma_A\rangle$
ρ_A
quantum measurement
m
decision element
$\widehat{A}$
=
$|\gamma_A\rangle$
ρ_A
equivalent quantum measurement
$\widehat{A}$

Fig. 5.5 The quantum measurement with the system of measurement operators $\{P_k, k \in \mathcal{M}\}$, followed by the decision element, is equivalent to the measurement with the system of global measurement operators $\{Q_i, i \in \mathcal{A}\}$

get the maximum correct decision probability. Prove the following statement: *Define the K decision functions as*

$$f_a(k) := q_a\, p_c(k|a), \qquad a \in \mathcal{A}, k \in \mathcal{M}.$$

Then, for each $k \in \mathcal{M}$, find the decision function $f_a(k)$ such that

$$f_a(k) \geq f_b(k), \qquad \forall b \neq a. \tag{5.17}$$

The value of a that verifies (5.17) is placed in $\mathcal{M}_a$. This defines the sets $\mathcal{M}_a$ that determine the optimum decision element.

Problem 5.3 ⋆⋆ In a binary system $\{0, 1\}$, where the a priori probabilities are $q(0) = 1/3$ and $q(1) = 2/3$, the quantum measurement, obtained with a photon counting, gives two Poisson variables with averages $\Lambda_0 = \mathrm{E}[m|A = 0] = 5$ and $\Lambda_1 = \mathrm{E}[m|A = 1] = 20$.

Apply the statement of the previous problem to find the optimum decision element.

Problem 5.4 ⋆ As in the previous problem but with $\Lambda_0 = 0$ and $\Lambda_1 = 20$ and equally likely symbols.

5.3 Analysis and Optimization of Quantum Binary Systems

To become familiar with the problem, before proceeding with the general theory, it seems useful to develop explicitly the decision theory in a binary quantum communications system, following the well-known *Helstrom theory* [1]. This theory represents one of the few cases in which explicit closed-form results are obtained.

In a quantum binary system with symbols $A \in \{0, 1\}$ the modulator (Alice) puts the system in one of the two states $|\gamma_0\rangle$ and $|\gamma_1\rangle$. We assume that the measurement alphabet $\mathcal{M}$ is still binary and coincident with the source alphabet, $\mathcal{A} = \mathcal{M} = \{0, 1\}$, and therefore we omit the post-measurement decision element. Then, for the measurement, we need two measurement operators (Hermitian and PSD) Q_0 and Q_1 that maximize the correct decision probability (optimal decision). Given that $Q_0 + Q_1 = I$, we can restrict our search to a single operator, for example, to Q_1.

5.3.1 Optimization with Mixed States (General Case)

We now proceed with the case in which the system is specified by two density operators ρ_0 and ρ_1. To calculate the probability of correct decision we use (5.16), which, as $Q_0 = I - Q_1$, yields

$$\begin{aligned} P_c &= q_0\,\mathrm{Tr}[\rho_0 Q_0] + q_1\,\mathrm{Tr}[\rho_1 Q_1] \\ &= q_0\,\mathrm{Tr}[\rho_0 I] + \mathrm{Tr}[(q_1\rho_1 - q_0\rho_0)Q_1] \\ &= q_0 + \mathrm{Tr}[(q_1\rho_1 - q_0\rho_0)Q_1] \end{aligned} \tag{5.18}$$

where we have taken into account the fact that the trace of a density operator is always unitary (see Sect. 3.3.2). The correct decision probability becomes

$$P_c = q_0 + \mathrm{Tr}[D\,Q_1]$$

where

$$D := q_1\rho_1 - q_0\rho_0 = \widehat{\rho}_1 - \widehat{\rho}_0 \tag{5.19}$$

is called for convenience *decision operator* ($\widehat{\rho}_i = q_i\rho_i$ are *weighted* density operators).

Then, to maximize the correct decision probability, we must find the measurement operator Q_1 such that

$$\max_{Q_1}\mathrm{Tr}[(q_1\rho_1 - q_0\rho_0)Q_1] = \max_{Q_1}\mathrm{Tr}[D\,Q_1] \qquad q_0 + q_1 = 1.$$

To this end, let us consider the eigendecomposition (EID) of the decision operator

$$D = q_1\rho_1 - q_0\rho_0 = \sum_k \eta_k|\eta_k\rangle\langle\eta_k| \tag{5.20}$$

where η_k is the generic eigenvalue, and $|\eta_k\rangle$ the corresponding eigenvector (the η_k are assumed as distinct, so the corresponding vectors $|\eta_k\rangle$ are orthonormal). Note that D is Hermitian but not PSD, so that the η_k are real, but may be either positive or negative. We then have

$$\begin{aligned} \mathrm{Tr}[D\,Q_1] &= \sum_k \eta_k \mathrm{Tr}[|\eta_k\rangle\langle\eta_k|Q_1] \\ &= \sum_k \eta_k\langle\eta_k|Q_1|\eta_k\rangle, \end{aligned} \tag{5.21}$$

where we have used the notable identity (2.37).

Now the crucial point for optimization is to observe that the quantity

$$\varepsilon_k := \langle\eta_k|Q_1|\eta_k\rangle$$

represents the **probability of a measurement obtained through the measurement operator** Q_1 when the system is in the state $|\eta_k\rangle$, and therefore $0 \le \varepsilon_k \le 1$. Then the maximum of the expression (5.21) is obtained by choosing, if possible, the terms with $\eta_k > 0$ and $\varepsilon_k = 1$. This choice is actually possible if we define the measurement operator Q_1 in the following way

$$\boxed{Q_1 = \sum_{\eta_k > 0} |\eta_k\rangle\langle\eta_k|.} \tag{5.22}$$

In fact, with this operator we obtain $\varepsilon_k = \langle\eta_k|Q_1|\eta_k\rangle = 1$ and the required maximum is

$$\mathrm{Tr}[(q_1\rho_1 - q_0\rho_0)Q_1] = \sum_{\eta_k > 0} \eta_k,$$

i.e., it is given by the sum of the positive eigenvalues. With this choice, the maximum correct decision probability becomes

$$\boxed{P_c = q_0 + \sum_{\eta_k > 0} \eta_k.} \tag{5.23}$$

It remains to verify that the two operators obtained through the optimization

$$Q_1 = \sum_{\eta_k > 0} |\eta_k\rangle\langle\eta_k|, \qquad Q_0 = I - Q_1 = \sum_{\eta_k < 0} |\eta_k\rangle\langle\eta_k| \tag{5.24}$$

really form a *measurement operator system*. What is more, it can be shown that Q_1 and Q_0 form a *projector system* (see Problem 5.5).

In conclusion, to obtain the maximum correct decision probability in a binary system, we must perform a *projective* measurement with projectors given by (5.24).

Summary of the Optimization Procedure

We summarize the steps required to find the optimal measurement operators in a quantum binary system:

(1) we start from the EID (5.20) of the decision operator

$$D = q_1\rho_1 - q_0\rho_0 = \sum_k \eta_k |\eta_k\rangle\langle\eta_k|; \tag{5.25}$$

(2) the optimal measurement operators (projectors) Q_0 and Q_1 are calculated from (5.24);

(3) the maximum probability of a correct decision is simply given by q_0 plus the sum of the positive eigenvalues of the operator D.

It is important to observe that this result is *totally general*, in the sense that no hypothesis has been made on the density operators ρ_0 and ρ_1, which can describe even mixed states. This general result will be applied in Chap. 8 to binary quantum communications systems in the presence of thermal noise.

Problem 5.5 ⋆⋆ Prove that the operators Q_1 and Q_0, defined by (5.24), form a projector system.

Problem 5.6 ⋆⋆ Consider the following density operators:

$$\rho_0 = \frac{1}{208}\begin{bmatrix} 46 & 13-37\mathrm{i} & -16 & 13+37\mathrm{i} \\ 13+37\mathrm{i} & 58 & 13-37\mathrm{i} & -32 \\ -16 & 13+37\mathrm{i} & 46 & 13-37\mathrm{i} \\ 13-37\mathrm{i} & -32 & 13+37\mathrm{i} & 58 \end{bmatrix}$$

$$\rho_1 = \frac{1}{208}\begin{bmatrix} 58 & 29-29\mathrm{i} & 8 & 21+29\mathrm{i} \\ 29+29\mathrm{i} & 58 & 29-21\mathrm{i} & -8 \\ 8 & 29+21\mathrm{i} & 46 & 21-21\mathrm{i} \\ 21-29\mathrm{i} & -8 & 21+21\mathrm{i} & 46 \end{bmatrix}$$

First verify that they are "true" density operators. Then, assuming that they are the states in a binary transmission with a priori probabilities $q_0 = 1/5$ and $q_1 = 4/5$, find the correct decision probability P_c.

5.4 Binary Optimization with Pure States

The general theory of the previous section is now applied to a binary quantum system prepared in one of the two pure states $|\gamma_0\rangle$ and $|\gamma_1\rangle$, therefore described by the density operators

$$\rho_0 = |\gamma_0\rangle\langle\gamma_0| \qquad \rho_1 = |\gamma_1\rangle\langle\gamma_1|. \tag{5.26}$$

We will find explicit and very important results, which be applied in Chap. 7 to quantum binary communications systems in the absence of thermal noise.

5.4.1 Helstrom's Bound

To find the optimal measurement operators, we must evaluate the EID of the decision operator, which with pure state is given by

$$D = q_1\rho_1 - q_0\rho_0 = q_1|\gamma_1\rangle\langle\gamma_1| - q_0|\gamma_0\rangle\langle\gamma_0|. \tag{5.27}$$

To comprehend the nature of this operator, consider its image

$$\mathcal{D} = \mathrm{im}\ D = D\,\mathcal{H}$$

which is a subspace generated by the linear combination of the two kets $|\gamma_0\rangle$ and $|\gamma_1\rangle$, assumed (geometrically) independent, and whose dimension is $\dim \mathcal{D} = 2$. Then the EID of D is limited to two terms (only two eigenvalues are different from zero) and the two eigenvectors $|\eta_)\rangle$ and $|\eta_1\rangle$ of D must belong to the subspace $\mathcal{D}$ and therefore are linear combinations of two states[3]

$$|\eta_0\rangle = a_{00}|\gamma_0\rangle + a_{01}|\gamma_1\rangle, \qquad |\eta_1\rangle = a_{10}|\gamma_0\rangle + a_{11}|\gamma_1\rangle. \tag{5.28}$$

Now, the coefficients a_{ij} are obtained by applying the definition of eigenvector, that is,

$$D\,|\eta_0\rangle = \eta_0\,|\eta_0\rangle, \qquad D\,|\eta_1\rangle = \eta_1\,|\eta_1\rangle \tag{5.29}$$

where η_0 and η_1 are the eigenvalues. Substituting (5.27) and (5.28) in (5.29), recalling that $\langle\gamma_1|\gamma_1\rangle = \langle\gamma_0|\gamma_0\rangle = 1$ and letting $X = \langle\gamma_0|\gamma_1\rangle$, we obtain

$$q_1(a_{0i}X + a_{1i})|\gamma_1\rangle - q_0(a_{0i} + a_{1i}X^*)|\gamma_0\rangle = \eta_{0i}(a_{0i}|\gamma_0\rangle + a_{1i}|\gamma_1\rangle), \qquad i = 0, 1. \tag{5.30}$$

But, because of the assumed *independence*, in (5.30) the coefficients of $|\gamma_1\rangle$ and $|\gamma_0\rangle$ must be equal to zeo. Hence

$$q_1(a_{i\,0}X^* + a_{i\,1}) = \eta_i\,a_{i\,1}, \qquad -q_0(a_{i\,0} + a_{i\,1}X) = \eta_i\,a_{i\,0}, \qquad i = 0, 1\ . \tag{5.31}$$

Solving with respect to η_i we get the equation

$$\eta_i^2 - \eta_i(q_1 - q_0) - q_0q_1(1 - |X|^2) = 0$$

from which

$$\eta_{0,1} = \tfrac{1}{2}\,(q_1 - q_0 \mp R)\,, \qquad R := \sqrt{1 - 4q_0q_1|X|^2} \tag{5.32}$$

where $\eta_1 > 0$ and $\eta_0 < 0$.

We have only one positive eigenvalue, and (5.23) gives

$$\boxed{\begin{aligned} P_c &= \tfrac{1}{2}\left(1 + \sqrt{1 - 4q_0q_1|X|^2}\right) \\ P_e &= \tfrac{1}{2}\left(1 - \sqrt{1 - 4q_0q_1|X|^2}\right) \end{aligned}} \tag{5.33}$$

where the parameter

$$|X|^2 = |\langle\gamma_0|\gamma_1\rangle|^2 \tag{5.33a}$$

[3] This point will be clarified in Sect. 5.11, Proposition 5.4. The eigenvectors $|\eta_i\rangle$ are called *measurement vectors* because they form the measurement operators as $Q_i = |\eta_i\rangle\langle\eta_i|$.

represents the *(quadratic) superposition degree* between the two states. In the literature expressions (5.33) are universally known as **Helstrom's bound**.

The optimal projectors derive from (5.24) and become

$$Q_0 = |\eta_0\rangle\langle\eta_0|, \qquad Q_1 = |\eta_1\rangle\langle\eta_1| \tag{5.34}$$

and therefore they are of the *elementary* type, with **measurement vectors** given by the eigenvectors $|\eta_0\rangle$ and $|\eta_1\rangle$ of the decision operator D.

It remains to complete the computation of these two eigenvectors, identified by the linear combinations (5.28). Considering (5.31) we find

$$|\eta_0\rangle = a_{00}\left(|\gamma_0\rangle + \frac{q_1 X^*}{\eta_0 - q_1}|\gamma_1\rangle\right), \qquad |\eta_1\rangle = a_{11}\left(-\frac{q_0 X}{\eta_1 + q_0}|\gamma_0\rangle + |\gamma_1\rangle\right) \tag{5.35}$$

where a_{00} and a_{11} are calculated by imposing the normalization $\langle\eta_i|\eta_i\rangle = 1$. In the general case, the calculation of the eigenvectors is very complicated[4] and we prefer to carry out the evaluation with the geometric approach developed below.

To consolidate the ideas on quantum detection we anticipate a few definitions and properties on quantum detection and optimization. The linear combination (5.28) can be written in the matrix form[5]

$$M = \Gamma A \quad \text{with} \quad \Gamma = [|\gamma_0\rangle, |\gamma_1\rangle],\ M = [|\mu_0\rangle, |\mu_1\rangle,\ A = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix}$$

where Γ is called *state matrix* and M is called *measurement matrix* (see Sect. 5.6). The target of optimization is to find the (optimal) measurement matrix M_{opt} that maximizes the correct decision probability. In Sect. 5.11 we shall see that the optimal measurement vectors are always orthogonal. This property can be written in the form $M_{\text{opt}}^* M_{\text{opt}} = I_2$, where I_2 is the 2×2 identity matrix.

Finally, we note that a quantum system with pure states, say $\mathcal{S}(q, \Gamma)$, is completely specified by the vector of the a priori probabilities q and by the state matrix Γ. The optimization is specified by the measurement matrix M_{opt}, which allows us to find the maximum correct decision probability as

$$P_{e,\max} = \sum_{i=0}^{K-1} |\langle\mu_i|\gamma_i\rangle|^2 = \sum_{i=0}^{K-1} |\langle M_{\text{opt}}(i)|\Gamma(i)\rangle|^2 \tag{5.36}$$

where $|\mu_i\rangle = M_{\text{opt}}(i)$ is the ith element of M_{opt}.

[4] To the author's knowledge the general expression of the eigenvectors (with X complex and not equally likely symbols) does not seem to be available in the literature.

[5] The measurement vectors, previously obtained as eigenvectors and denoted by $|\eta_i\rangle$, are hereafter denoted by $|\mu_i\rangle$.

5.4.2 *Optimization by Geometric Method*

The binary optimization with pure states can be conveniently developed by a geometric approach with several advantages. We first assume that the inner product $Y := \langle\gamma_0|\gamma_1\rangle$ is **real** and then we generalize the approach to the complex case.

The geometry of decision with two pure states $|\gamma_0\rangle$ and $|\gamma_1\rangle$ is developed in the subspace $\mathcal{D}$ generated by two states. In this hyperplane, the states are written in terms of an appropriate orthonormal basis $\{|u_0\rangle, |u_1\rangle\}$ as (Fig. 5.6)

$$\begin{aligned} |\gamma_0\rangle &= \cos\theta|u_0\rangle + \sin\theta|u_1\rangle \\ |\gamma_1\rangle &= \cos\theta|u_0\rangle - \sin\theta|u_1\rangle \end{aligned} \tag{5.37}$$

where

$$\cos 2\theta = \langle\gamma_0|\gamma_1\rangle = Y. \tag{5.38}$$

In (5.37) we have assumed that the basis vector $|u_0\rangle$ lies in the bisection determined by the state vectors, which does not represent a restriction. For now we assume the two measurement vectors $|\mu_0\rangle$ and $|\mu_1\rangle$ not necessarily optimal, but satisfying the conditions of being orthonormal, in addition to belonging to the Hilbert subspace $\mathcal{D}$. Then they can be written as

$$\begin{aligned} |\mu_0\rangle &= \cos\phi|u_0\rangle + \sin\phi|u_1\rangle \\ |\mu_1\rangle &= \sin\phi|u_0\rangle - \cos\phi|u_1\rangle. \end{aligned} \tag{5.39}$$

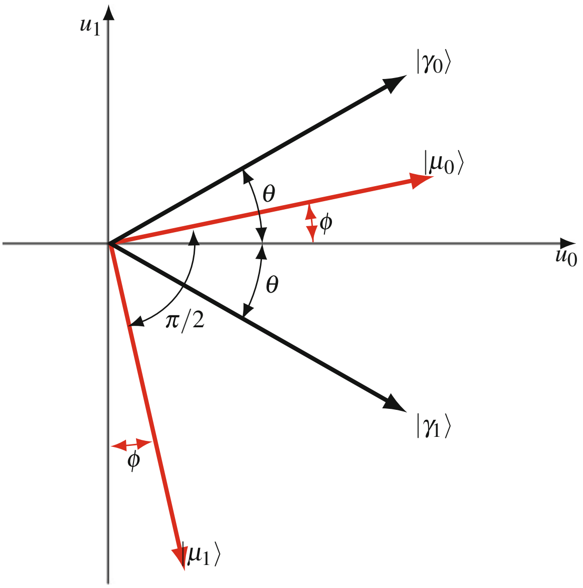


Fig. 5.6 Binary decision with generic state vectors and measurement vectors

Note that trigonometric functions take automatically into account the ket normalization and allow us to express the ket geometry of the four vectors involved by only two angles.

Considering that $\langle\mu_0|\gamma_0\rangle = \cos(\phi - \theta)$ and $\langle\mu_1|\gamma_1\rangle = \sin(\phi + \theta)$, the transition probabilities $p(j|i) := \mathrm{P}[\hat{A}_0 = j|A_0 = i]$ are given by

$$\begin{aligned} p(0|0) &= \cos^2(\phi - \theta) = \tfrac{1}{2}\,[1 + \sin 2\theta \sin 2\phi + \cos 2\theta \cos 2\phi] \\ p(1|1) &= \sin^2(\phi + \theta) = \tfrac{1}{2}\,[1 + \sin 2\theta \sin 2\phi - \cos 2\theta \cos 2\phi] \end{aligned} \tag{5.40}$$

and the correct detection probability turns out to be

$$\begin{aligned} P_c &= q_0 \cos^2(\phi - \theta) + q_1 \sin^2(\phi + \theta) \\ &= \tfrac{1}{2}\,[1 + (q_0 - q_1)(\cos 2\theta\ \cos 2\phi + \sin 2\theta\ \sin 2\phi)]\,. \end{aligned} \tag{5.41}$$

Here the angle θ is given through the inner product Y (see (5.38)), while the angle ϕ is unknown and is evaluated by optimization. It is immediate to see that the angle ϕ giving the maximum of P_c is given by

$$\tan 2\phi = \frac{1}{q_0 - q_1} \tan 2\theta = \frac{1}{q_0 - q_1} \frac{\sqrt{1 - Y^2}}{Y}, \tag{5.42}$$

which gives

$$\sin 2\phi = \frac{1}{R} \sin 2\theta, \quad \cos 2\phi = \frac{q_0 - q_1}{R} \cos 2\theta \tag{5.43}$$

where $R = \sqrt{1 - 4q_0q_1Y^2}$. The corresponding optimal correct decision probability is

$$P_c = \tfrac{1}{2}(1 + R) = \tfrac{1}{2}\left(1 + \sqrt{1 - 4q_0q_1Y^2}\right), \tag{5.44}$$

i.e., the Helstrom bound.

The transition probabilities (5.40), with the optimal decision, become

$$\begin{aligned} p(0|0) &= \tfrac{1}{2}\left[1 + (1 - Y^2 + (q_0 - q_1)Y^2)/R\right] \\ p(1|1) &= \tfrac{1}{2}\left[1 + (1 - Y^2 - (q_0 - q_1)Y^2)/R\right]. \end{aligned} \tag{5.45}$$

Finally, we consider the explicit evaluation of the optimal measurement vectors. The first step is finding in (5.39) the expression of the basis vectors $|u_0\rangle$ and $|u_1\rangle$ in terms of the given quantum states. For the particular choice made for these vectors we have that $|u_0\rangle$ is proportional to $|\gamma_0\rangle + |\gamma_1\rangle$ and $|u_1\rangle$ is proportional to $|\gamma_0\rangle - |\gamma_1\rangle$ (see Fig. 5.6), that is, $|u_0\rangle = H_0(|\gamma_0\rangle + |\gamma_1\rangle)$ and $|u_1\rangle = H_1(|\gamma_0\rangle - |\gamma_1\rangle)$.

The normalization gives $H_0 = 1/\sqrt{2+2Y}$ and $H_1 = 1/\sqrt{2-2Y}$. Next, in (5.39) the optimal angle is given by (5.42) and then we find the optimal measurement matrix as[6]

$$M_{\text{opt}} = \varGamma A \quad \text{with} \quad A = \frac{1}{2}\begin{bmatrix} \dfrac{\sqrt{1-L}}{\sqrt{1-Y}} + \dfrac{\sqrt{1+L}}{\sqrt{1+Y}} & \dfrac{\sqrt{1-L}}{\sqrt{1+Y}} - \dfrac{\sqrt{1+L}}{\sqrt{1-Y}} \\[2ex] \dfrac{\sqrt{1+L}}{\sqrt{1+Y}} - \dfrac{\sqrt{1-L}}{\sqrt{1-Y}} & \dfrac{\sqrt{1-L}}{\sqrt{1+Y}} + \dfrac{\sqrt{1+L}}{\sqrt{1-Y}} \end{bmatrix} \tag{5.46}$$

where $L = (q_0 - q_1)\, Y/R$. This completes the optimization with a real inner product.

In the general case of a **complex inner product**

$$X = |X|\, \mathrm{e}^{\mathrm{i}\,\beta}$$

we introduce the new quantum states

$$|\widetilde{\gamma}_0\rangle = |\gamma_0\rangle, \qquad |\widetilde{\gamma}_1\rangle = \mathrm{e}^{-\mathrm{i}\,\beta}|\gamma_1\rangle$$

which give the matrix relation

$$\widetilde{\varGamma} = \varGamma B, \qquad \text{with} \quad B = \begin{bmatrix} \mathrm{e}^{-\mathrm{i}\,\beta} & 0 \\ 0 & 1 \end{bmatrix}. \tag{5.47}$$

Now we have two binary systems, $\mathcal{S}(q, \varGamma)$ and $\mathcal{S}(q, \widetilde{\varGamma})$, with the same a priori probabilities, but different inner products, respectively $X = |X|\, \mathrm{e}^{\mathrm{i}\,\beta}$ and $\widetilde{X} = \langle\widetilde{\gamma}_0|\widetilde{\gamma}_1\rangle = \mathrm{e}^{-\mathrm{i}\,\beta}\langle\gamma_0|\gamma_1\rangle = |X|$. It is immediate to verify (see (5.36)) that if M_{opt} is the optimal measurement matrix for $\mathcal{S}(q, \varGamma)$, the optimal measurement matrix for $\mathcal{S}(q, \widetilde{\varGamma})$ is given by

$$\widetilde{M}_{\text{opt}} = M_{\text{opt}}\, B \quad \to \quad M_{\text{opt}} = \widetilde{M}_{\text{opt}}\, B^{-1}. \tag{5.48}$$

But the system $\mathcal{S}(q, \widetilde{\varGamma})$ has a real inner product and, with the replacement $Y \to |X|$, we can use the previous theory to find: (1) the Helstrom bound from (5.44), (2) the transition probabilities from (5.45) and (3) the optimal measurement matrix $\widetilde{M}_{\text{opt}} = \widetilde{\varGamma}\widetilde{A}$ from (5.46). Hence, from $\widetilde{M}_{\text{opt}}$ we can obtain the measurement matrix for the system $\mathcal{S}(q, \varGamma)$. In fact, by combination of (5.47) and (5.48) we find

$$M_{\text{opt}} = \varGamma A \quad \text{with} \quad A = B\widetilde{A}\, B^{-1}$$

[6] To express $\cos\phi$ and $\sin\phi$ from $\tan 2\phi$ we use the trigonometric identities

$$\sin\phi = 2^{-1/2}\sqrt{1 - 1/\sqrt{1+\tan^2 2\phi}}, \qquad \cos\phi = 2^{-1/2}\sqrt{1 + 1/\sqrt{1+\tan^2 2\phi}}$$

which hold for $0 \le \phi \le \pi/4$. This range of ϕ covers the cases of interest.

which gives

$$A = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} = \begin{bmatrix} \widetilde{a}_{00} & \mathrm{e}^{\mathrm{i}\,\beta}\widetilde{a}_{01} \\ \mathrm{e}^{-\mathrm{i}\,\beta}\widetilde{a}_{10} & \widetilde{a}_{11} \end{bmatrix}.$$

We summarize the general results as follows:

Proposition 5.2 *The optimization of the quantum decision in a binary system prepared in the pure states $|\gamma_0\rangle$ and $|\gamma_1\rangle$, having inner product $\langle\gamma_0|\gamma_1\rangle := X = |X|\mathrm{e}^{\mathrm{i}\beta}$ and a priori probabilities q_0 and q_1, gives the transition probabilities*

$$\begin{aligned} p(0|0) &= \tfrac{1}{2}\left[1 + (1 - |X|^2 + (q_0 - q_1)|X|^2)/R\right] \\ p(1|1) &= \tfrac{1}{2}\left[1 + (1 - |X|^2 - (q_0 - q_1)|X|^2)/R\right]. \end{aligned} \tag{5.49}$$

and the correct decision probability

$$P_c = \tfrac{1}{2}\left(1 + \sqrt{1 - 4q_0q_1|X|^2}\right). \tag{5.50}$$

The optimal measurement matrix is obtained as $M_{\text{opt}} = \Gamma A$, where

$$A = \frac{1}{2}\begin{bmatrix} \dfrac{\sqrt{1-L}}{\sqrt{1-|X|}} + \dfrac{\sqrt{1+L}}{\sqrt{1+|X|}} & \mathrm{e}^{\mathrm{i}\,\beta}\left(\dfrac{\sqrt{1-L}}{\sqrt{1+|X|}} - \dfrac{\sqrt{1+L}}{\sqrt{1-|X|}}\right) \\ \mathrm{e}^{-\mathrm{i}\beta}\left(\dfrac{\sqrt{1+L}}{\sqrt{1+|X|}} - \dfrac{\sqrt{1-L}}{\sqrt{1-|X|}}\right) & \dfrac{\sqrt{1-L}}{\sqrt{1+|X|}} + \dfrac{\sqrt{1+L}}{\sqrt{1-|X|}} \end{bmatrix}$$

with

$$R = \sqrt{1 - 4q_0q_1}, \qquad L = |(q_0 - q_1)X|/R.$$

5.4.3 Pure States with Equally Likely Symbols

With equally likely symbols ($q_0 = q_1 = \frac{1}{2}$) we find several simplifications. In the trigonometric approach the optimization is obtained by rotating the measurement vectors until they form the same angle with the corresponding state vectors, specifically, we have $\theta = \pi/4$, as shown in Fig. 5.7. The expressions of correct decision probabilities and of the error probability are simplified as

$$P_c = \tfrac{1}{2}\left(1 + \sqrt{1 - |X|^2}\right), \qquad P_e = \tfrac{1}{2}\left(1 - \sqrt{1 - |X|^2}\right). \tag{5.51}$$

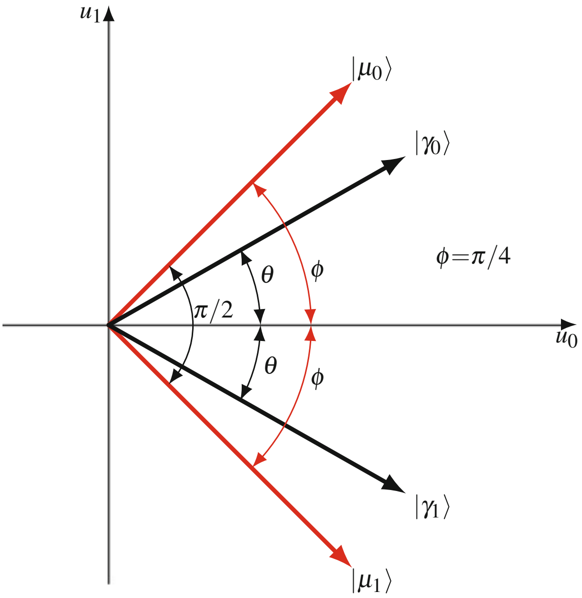


Fig. 5.7 Optimal binary decision with equally probable symbols ($q_0 = q_1 = \frac{1}{2}$). The optimization is obtained by rotating the measurement vectors until they form the same angle with the corresponding state vectors

The transition probabilities become equal

$$p(0|0) = p(1|1) = \tfrac{1}{2}\left(1 + \sqrt{1 - |X|^2}\right) = P_c$$

and hence we get a *binary symmetric channel.*

The measurement vectors become

$$|\mu_0\rangle = a\,|\gamma_0\rangle + b\,\mathrm{e}^{\mathrm{i}\beta}\,|\gamma_1\rangle, \qquad |\mu_1\rangle = b\,\mathrm{e}^{-\mathrm{i}\beta}\,|\gamma_0\rangle + a\,|\gamma_1\rangle \tag{5.52}$$

where $\beta = \arg X$ and

$$a = \frac{1}{2}\left[\frac{1}{\sqrt{1-|X|}} + \frac{1}{\sqrt{1+|X|}}\right], \qquad b = \frac{1}{2}\left[\frac{1}{\sqrt{1+|X|}} - \frac{1}{\sqrt{1-|X|}}\right]. \tag{5.53}$$

Problem 5.7 ⋆⋆ Find the coefficients a_{01} and a_{11} in the expression of the measurement vectors (5.35), assuming equally likely symbols and X real.

Problem 5.8 ⋆⋆ Write the fundamental relations of the geometrical approach in matrix form, using the matrices

$$\Gamma = [|\gamma_0\rangle, |\gamma_1\rangle], \qquad U = [|u_0\rangle, |u_1\rangle], \qquad M = [|\mu_0\rangle, |\mu_1\rangle].$$

5.5 System Specification in Quantum Decision Theory

After a thorough examination of decision in a binary system, we return to the general considerations, assuming a K-ary system.

From the general analysis of Sect. 5.2 and from the choice of proceeding with global measurements, we realize that the system specification in quantum decision can be limited to the following few parameters ("players").

On the transmitter side (Alice), the players are:

(a) the a priori probabilities q_i, $i \in \mathcal{A}$,
(b) the states $|\gamma_i\rangle$, $i \in \mathcal{A}$, or the density operator ρ_i, $i \in \mathcal{A}$.

The sets $\{|\gamma_i\rangle | i \in \mathcal{A}\}$ and $\{\rho_i | i \in \mathcal{A}\}$ will be called *constellations of states*.

At the receiver side (Bob) the players are the (global) measurement operators, which must form a *measurement operator system* $\{Q_i, i \in \mathcal{A}\}$ in the sense already underlined, but worthwhile recalling:

(1) they are Hermitian operators, $Q_i^* = Q_i$,
(2) they are PSD, $Q_i \geq 0$,
(3) they give a resolution of the identity, $\sum_{i\in\mathcal{A}} Q_i = I_{\mathcal{H}}$.

There are several ways to specify the above parameters, as we shall see in the next sections, making the usual distinction between pure and mixed states.

5.5.1 Weighted States and Weighted Density Operators

In the above, the transmitter specification is composed by two players, however, the same specification can be obtained by a single player with the introduction of weighted states (already used in Sect. 3.11).

The *weighted states* are defined by

$$|\widehat{\gamma}_i\rangle = \sqrt{q_i}|\gamma_i\rangle, \qquad i \in \mathcal{A} \tag{5.54}$$

and contain the information of both the probabilities q_i and the states $|\gamma_i\rangle$. In fact, considering that the states are normalized, $\langle\gamma_i|\gamma_i\rangle = 1$, one gets

$$q_i = \langle\widehat{\gamma}_i|\widehat{\gamma}_i\rangle, \qquad |\gamma_i\rangle = (1/\sqrt{q_i})\ |\widehat{\gamma}_i\rangle. \tag{5.55}$$

The *weighted density operators* are defined by

$$\widehat{\rho}_i = q_i\ \rho_i, \qquad i \in \mathcal{A}. \tag{5.56}$$

Then, considering that $\mathrm{Tr}[\rho_i] = 1$, one gets

$$q_i = \mathrm{Tr}[\widehat{\rho}_i], \qquad \rho_i = (1/q_i)\ \widehat{\rho}_i. \tag{5.57}$$

5.6 State and Measurement Matrices with Pure States

If the decision is taken from pure states, that is, from rank-one density operators, also the measurement operators may be chosen with rank-one, and therefore expressed by *measurement vectors* in the form $Q_i = |\mu_i\rangle\langle\mu_i|$. This was seen in Sect. 5.3 with a binary system, but it holds in general (see Kennedy's theorem in Sect. 5.11). Then, referring to an n-dimensional Hilbert space $\mathcal{H}$, the players become vectors (kets) of $\mathcal{H}$, which can be conveniently represented in the matrix form.

Now, K pure states $|\gamma_i\rangle$, interpreted as column vectors of dimension $n \times 1$, form *the state matrix*

$$\underset{n\times K}{\Gamma} = [|\gamma_0\rangle, |\gamma_1\rangle, ..., |\gamma_{K-1}\rangle]. \tag{5.58}$$

Analogously, the measurement vectors $|\mu_i\rangle$ form the *measurement matrix*

$$\underset{n\times K}{M} = [|\mu_0\rangle, |\mu_1\rangle, \dots, |\mu_{K-1}\rangle]. \tag{5.59}$$

In particular, the measurement matrix allows us to express the resolution of the identity $\sum_{i\in\mathcal{A}} |\mu_i\rangle\langle\mu_i| = I_{\mathcal{H}}$, in the compact form

$$M\,M^* = I_{\mathcal{H}}. \tag{5.60}$$

The specification of the source by the state matrix Γ is sufficient in the case of equally likely symbols. With generic a priori probabilities q_i we can introduce the matrix of weighted states [3]

$$\widehat{\Gamma} = \left[|\widehat{\gamma}_0\rangle, |\widehat{\gamma}_1\rangle_1, \dots, |\widehat{\gamma}_{K-1}\rangle\right], \tag{5.61}$$

where $|\widehat{\gamma}_i\rangle = \sqrt{q_i}|\gamma_i\rangle$.

5.7 State and Measurement Matrices with Mixed States ⇓

The state and the measurement matrices can be extended to mixed states but their introduction is less natural because the density and the measurement operators are not presented in a factorized form as in the case of pure states.

With pure states, the density operators have the factorized form $\rho_i = |\gamma_i\rangle\langle\gamma_i|$ and, with standard notation, $\rho_i = \gamma_i\gamma_i^*$, where the states $\gamma_i = |\gamma_i\rangle$ must be considered as column vectors. In the general case, the density operators do not appear as a product of two factors, but can be equally factorized in the form

$$\rho_i = \gamma_i\gamma_i^*, \qquad i = 0, 1, \dots, K-1 \tag{5.62}$$

where the γ_i become matrices of appropriate dimensions (and not simply column vectors). As we will see soon, if n is the dimension of the Hilbert space and h_i is the rank of ρ_i, the matrix γ_i can be chosen of dimensions $n \times h_i$. It must be observed that such factorization is not unique, and also the dimensions $n \times h_i$ are to some extent arbitrary, because h_i has the constraint $\text{rank}(\rho_i) \le h_i \le n$. However, the minimal choice $h_i = \text{rank}(\rho_i)$ is the most convenient (and in the following we will comply with this choice).

Similar considerations hold for the measurement operators Q_i, which, with unitary rank, have the factored form $Q_i = |\mu_i\rangle\langle\mu_i|$, but also with rank $h_i > 1$ can be factored in the form

$$Q_i = \mu_i \mu_i^* \tag{5.63}$$

where the factors μ_i are $n \times h_i$ matrices. Further on, we will realize (see Kennedy's theorem and its generalization in Sect. 5.11) that in the choice of the measurement operators it is not restrictive to assume that h_i be given by the same rank of the corresponding density operators.

By analogy with the pure states and with the measurement vectors, the factors γ_i will be called **state factors** and the factors μ_i **measurement factors** (this terminology is not standard and is introduced for the sake of simplicity). The factorization will be useful in various ways; first of all because, if the rank h_i is not full ($h_i < n$), it removes the redundancy of the operators, by gathering the information in an $n \times h_i$ rectangular matrix, instead of an $n \times n$ square matrix, and also because it often makes it possible to extend to the general case some results that are obtained with pure states.

5.7.1 How to Obtain a Factorization

The factorization of a density operator was developed in Sect. 3.11 in the context of the multiplicity of an ensemble of probabilities/states. Here the factorization is seen in a different context and, for clarity, some considerations will be repeated.

Consider a generic density operator ρ of dimensions $n \times n$ and rank h, which is always a PSD Hermitian operator. Then a factorization $\gamma\gamma^*$ can be obtained using its *reduced* EID (see Sect. 2.11 and Proposition 3.5)

$$\rho = Z_h\, D_h^2\, Z_h^* = \sum_{i=1}^{h} d_i^2\, |z_i\rangle\langle z_i| \tag{5.64}$$

where $D_h^2 = \text{diag}[d_1^2, \dots, d_h^2]$ is an $h \times h$ diagonal matrix containing the h positive eigenvalues of ρ and $Z_h = [|z_1\rangle \cdots |z_h\rangle]$ is $n \times h$. Letting $D_h = \sqrt{D_h^2} = \text{diag}[d_1, \dots, d_h]$, we see immediately that

$$\gamma = Z_h\, D_h \tag{5.65}$$

is a *factor* of ρ.

From the EID (5.64) it results that the density operator is decomposed into the sum of the elementary operators $d_i^2\, |z_i\rangle\langle z_i|$, where d_i^2 has the meaning of the probability that the quantum system described by the operator ρ be in the state $|z_i\rangle$, exactly in the form in which the density operator has been introduced (see (3.7)). Then the factor γ turns out to be a collection of h vectors

$$\gamma = [d_1\,|z_1\rangle, \ldots, d_h\,|z_h\rangle] \tag{5.66}$$

where the $|z_i\rangle$ are orthonormal (as taken from a unitary matrix Z of an EID).[7]

Reconsidering the theory developed in Sect. 3.11, we find that γ is a *minimum factor* of ρ and, more specifically, a *minimum orthogonal factor*.

Example 5.1 Consider the Hilbert space $\mathcal{H} = \mathbb{C}^4$, where we assume as basis

$$|b_1\rangle = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \quad |b_2\rangle = \begin{bmatrix} \frac{1}{2} \\ -\frac{\mathrm{i}}{2} \\ -\frac{1}{2} \\ \frac{\mathrm{i}}{2} \end{bmatrix}, \quad |b_3\rangle = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}, \quad |b_4\rangle = \begin{bmatrix} \frac{1}{2} \\ \frac{\mathrm{i}}{2} \\ -\frac{1}{2} \\ -\frac{\mathrm{i}}{2} \end{bmatrix}.$$

From this basis we build the density operator

$$\rho = \frac{1}{3}|b_1\rangle\langle b_1| + \frac{2}{3}|b_2\rangle\langle b_2| = \begin{bmatrix} \frac{1}{4} & \frac{1}{12} - \frac{\mathrm{i}}{6} & -\frac{1}{12} & \frac{1}{12} + \frac{\mathrm{i}}{6} \\ \frac{1}{12} + \frac{\mathrm{i}}{6} & \frac{1}{4} & \frac{1}{12} - \frac{\mathrm{i}}{6} & -\frac{1}{12} \\ -\frac{1}{12} & \frac{1}{12} + \frac{\mathrm{i}}{6} & \frac{1}{4} & \frac{1}{12} - \frac{\mathrm{i}}{6} \\ \frac{1}{12} - \frac{\mathrm{i}}{6} & -\frac{1}{12} & \frac{1}{12} + \frac{\mathrm{i}}{6} & \frac{1}{4} \end{bmatrix}$$

which has eigenvalues $\left\{\frac{2}{3}, \frac{1}{3}, 0, 0\right\}$ and therefore has rank $h = 2$. Its reduced EID $\rho = Z_h\, D_h^2\, Z_h^*$ is specified by the matrices

$$Z_h = \begin{bmatrix} \frac{\mathrm{i}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \\ -\frac{\mathrm{i}}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \qquad D_h^2 = \begin{bmatrix} \frac{2}{3} & 0 \\ 0 & \frac{1}{3} \end{bmatrix} \qquad Z_h^* = \begin{bmatrix} -\frac{\mathrm{i}}{2} & -\frac{1}{2} & \frac{\mathrm{i}}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2}. \end{bmatrix}$$

Now, to obtain a factor γ of ρ we use (5.65), which gives the 4 × 2 matrix

[7] Another way to obtain a factorization is given by Choleski's decomposition (see Sect. 2.12.5).

$$\gamma = Z_h \sqrt{D^2} = \begin{bmatrix} \frac{\mathrm{i}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \\ -\frac{\mathrm{i}}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \sqrt{\frac{2}{3}} & 0 \\ 0 & \sqrt{\frac{1}{3}} \end{bmatrix} = \begin{bmatrix} \frac{\mathrm{i}}{\sqrt{6}} & \frac{1}{2\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} \\ -\frac{\mathrm{i}}{\sqrt{6}} & \frac{1}{2\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{2\sqrt{3}}. \end{bmatrix}$$

As regards the factorization of the measurement operator, say $Q = \mu\mu^*$, similar considerations hold, provided that the operator Q is known. However, in quantum detection Q is not known and it should be determined by optimization. In this context the unknown may become μ, then giving Q as $\mu\mu^*$ and the factorization is no more required.

5.7.2 State and Measurement Matrices

The definition of these matrices can be extended to mixed states by expressing the density operators and the corresponding measurement operators through their factors. The state matrix Γ is obtained by juxtaposing the factors γ_i, intended as blocks of columns of dimensions $n \times h_i$

$$\underset{n\times H}{\Gamma} = \left[\gamma_0, \gamma_1, \ldots, \gamma_{K-1}\right] \tag{5.67}$$

where the number of columns H is given by the global number of the columns of the factors γ_i

$$H = h_0 + h_1 + \cdots + h_{K-1}.$$

We can make explicit Γ bearing in mind that each factor γ_i is a collection of h_i kets (see (5.66)). For example, for $K = 2$, $h_0 = 2$, $h_1 = 3$ we have

$$\Gamma = [\gamma_0, \gamma_1] = [|\gamma_{01}\rangle, |\gamma_{02}\rangle, |\gamma_{11}\rangle, |\gamma_{12}\rangle, |\gamma_{13}\rangle] \tag{5.68}$$

where $|\gamma_{0i}\rangle$ are the kets of γ_0 and $|\gamma_{1i}\rangle$ are the kets of γ_1.

Analogously, the measurement matrix M is obtained by juxtaposing the factors μ_i, intended as blocks of columns

$$\underset{n\times H}{M} = \left[\mu_0, \mu_1, \ldots, \mu_{K-1}\right]. \tag{5.69}$$

Even the resolution of the identity (5.60) is extended to mixed states. In fact,

$$M\,M^* = \sum_{i\in\mathcal{A}} \mu_i \mu_i^* = \sum_{i\in\mathcal{A}} Q_i = I_{\mathcal{H}}. \tag{5.70}$$

Clearly, these last definitions are the most general and include the previous ones when the ranks are unitary ($h_i = 1$ and $H = K$).

Also the definition of the *matrix of weighted states*, given by (5.61) for pure states, can be extended to mixed states [3], namely

$$\widehat{\Gamma} = \left[\widehat{\gamma}_0, \widehat{\gamma}_1, \ldots, \widehat{\gamma}_{K-1}\right] = \left[\sqrt{q_0}\gamma_0, \sqrt{q_1}\gamma_1, \ldots, \sqrt{q_{K-1}}\gamma_{K-1}\right], \tag{5.71}$$

where the weighted states can be obtained as a factorization of weighted density operators, namely $\widehat{\rho}_i = q_i \rho_i = \sqrt{q_i}\gamma_i \sqrt{q_i}\gamma_i^* = \widehat{\gamma}_i \widehat{\gamma}_i^*$.

5.7.3 Probabilities Expressed Through Factors

In quantum decision, probabilities can be computed from the factors γ_i and μ_i of the density operators and of the measurement operators. Recalling the expression of the transition probabilities, given by (5.15), we obtain explicitly

$$p_c(j|i) = \mathrm{Tr}[Q_j\, \rho_i] = \mathrm{Tr}[\mu_j \mu_j^*\, \gamma_i \gamma_i^*]. \tag{5.72}$$

Analogously, from (5.16) we obtain the correct decision probability

$$P_c = \sum_{i \in \mathcal{A}} q_i\, \mathrm{Tr}[Q_i\, \rho_i] = \sum_{i \in \mathcal{A}} q_i\, \mathrm{Tr}[\mu_i \mu_i^*\, \gamma_i \gamma_i^*]. \tag{5.73}$$

In the evaluation of these probabilities it is convenient to introduce the *matrix of mixed products*

$$\underset{H \times H}{B} := M^*\, \Gamma = \begin{bmatrix} b_{0,0} & \cdots & b_{0,K-1} \\ \vdots & \ddots & \vdots \\ b_{k-1,0} & \cdots & b_{K-1,K-1} \end{bmatrix}, \qquad b_{ij} := \mu_i^*\, \gamma_i \tag{5.74}$$

where dim $b_{ij} = h_i \times h_j$. Then, using the cyclic property of the trace, we find

$$\boxed{p_c(j|i) = \mathrm{Tr}[b_{ji}^*\, b_{ji}], \qquad P_c = \sum_{i \in \mathcal{A}} q_i \mathrm{Tr}[b_{ii}^*\, b_{ii}].} \tag{5.75}$$

Finally, it must be observed that state and measurement factors are not uniquely determined by the corresponding operators. In fact, if γ_i is a factor of ρ_i, also $\tilde{\gamma}_i = \gamma_i\, Z$, where Z is any matrix with the property $ZZ^* = I_{h_i}$, is a factor of ρ_i, as follows from $\tilde{\gamma}_i \tilde{\gamma}_i^* = \gamma_i\, ZZ^* \gamma_i^* = \gamma_i\, \gamma_i^* = \rho_i$. However, the multiplicity of the factors has no influence on the computation of the probabilities, as can be verified from the above expressions.

Problem 5.9 ⋆⋆ From the following normalized states of $\mathcal{H} = \mathbb{C}^4$

$$|\gamma_1\rangle = \begin{bmatrix} \frac{2}{\sqrt{13}} \\ \frac{2}{\sqrt{13}} \\ \frac{2}{\sqrt{13}} \\ \frac{1}{\sqrt{13}} \end{bmatrix} \quad |\gamma_2\rangle = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \quad |\gamma_3\rangle = \begin{bmatrix} \frac{1}{2} \\ -\frac{\mathrm{i}}{2} \\ -\frac{1}{2} \\ \frac{\mathrm{i}}{2} \end{bmatrix} \quad |\gamma_4\rangle = \begin{bmatrix} \frac{2}{\sqrt{13}} \\ -\frac{2\mathrm{i}}{\sqrt{13}} \\ -\frac{1}{\sqrt{13}} \\ \frac{2\mathrm{i}}{\sqrt{13}} \end{bmatrix} \quad |\gamma_5\rangle = \begin{bmatrix} \frac{1}{\sqrt{13}} \\ -\frac{2\mathrm{i}}{\sqrt{13}} \\ -\frac{2}{\sqrt{13}} \\ \frac{2\mathrm{i}}{\sqrt{13}} \end{bmatrix}$$

form the density operators

$$\rho_1 = \frac{3}{4}|\gamma_1\rangle\langle\gamma_1| + \frac{1}{4}|\gamma_2\rangle\langle\gamma_2|, \qquad \rho_2 = \frac{3}{4}|\gamma_3\rangle\langle\gamma_3| + \frac{1}{8}|\gamma_4\rangle\langle\gamma_4|\frac{1}{8}|\gamma_5\rangle\langle\gamma_5|$$

and find their *minimum* factors γ_1 and γ_2. Find also factorizations in which the matrices γ_1 and γ_2 have the same dimensions.

Problem 5.10 ⋆ Consider the transition probabilities given by (5.72). Prove that, if γ_i is replaced by $\gamma_i Z$, with $ZZ^* = I_h$, and μ_j by $\mu_j W$, with $WW^* = I_h$, the transition probabilities do not change.

Problem 5.11 ⋆⋆ Prove that the measurement matrix M defined by (5.59) and its generalization to mixed states (5.69), allows us to express the resolution of the identity in the form $MM^* = I_{\mathcal{H}}$.

5.8 Formulation of Optimal Quantum Decision

The viewpoint for the Optimal Quantum Decision is the following: the a priori probabilities and the constellation (of pure states or of mixed stated) are assumed as given, whereas the measurement operator system is unknown and should be determined to meet the decision criterion, given by **the maximization of the correct decision probability**.

Then, considering the general expression of the correct decision probability, given by (see (5.16))

$$P_c = \sum_{i \in \mathcal{A}} q_i \operatorname{Tr}[\rho_i Q_i]$$

the *optimal measurement operators* Q_i must be determined from

$$\max_{\{Q_i\}} \sum_{i=0}^{K-1} q_i \operatorname{Tr}[\rho_i Q_i]. \tag{5.76}$$

If the operators are expressed through their factors (see (5.62) and (5.63)), (5.76) becomes

$$\max_{\{\mu_i\}} \sum_{i=0}^{K-1} q_i \operatorname{Tr}[\gamma_i \gamma_i^* \mu_i \mu_i^*]. \tag{5.77}$$

Finally, if the states are pure, we have the simplification

$$\max_{\{|\mu_i\rangle\}} \sum_{i=0}^{K-1} q_i \operatorname{Tr}[|\gamma_i\rangle\langle\gamma_i|\mu_i\rangle\langle\mu_i|] = \max_{\{|\mu_i\rangle\}} \sum_{i=0}^{K-1} q_i \, |\langle\gamma_i|\mu_i\rangle|^2. \tag{5.78}$$

In the last relation we have used the identity (2.37) over the trace, $\operatorname{Tr}[A|u\rangle\langle u|] = \langle u|A|u\rangle$, with $A = |\gamma_i\rangle\langle\gamma_i|$ and $|u\rangle = |\mu_i\rangle$.

5.8.1 *Optimization as Convex Semidefinite Programming (CSP)*

Starting from the Hilbert space $\mathcal{H}$ on which the quantum decision is defined, it is convenient to introduce the following classes (Fig. 5.8):

- the class $\mathcal{B}$ of the Hermitian operators defined on $\mathcal{H}$,
- the subclass $\mathcal{B}_0$ of the PSD Hermitian operators,
- the class $\mathcal{M}$ of the K-tuples $\mathbf{Q} = [\, Q_0, \ldots, Q_{K-1}\,]$, $Q_i \in \mathcal{B}$ of Hermitian operators,
- the subclass $\mathcal{M}_0$ of $\mathcal{M}$ consisting of the K-tuples $\mathbf{Q}$, whose elements Q_i are PSD Hermitian, $Q_i \in \mathcal{B}_0$, and, globally, resolve the identity on $\mathcal{H}$, that is, $\sum_i Q_i = I_{\mathcal{H}}$.

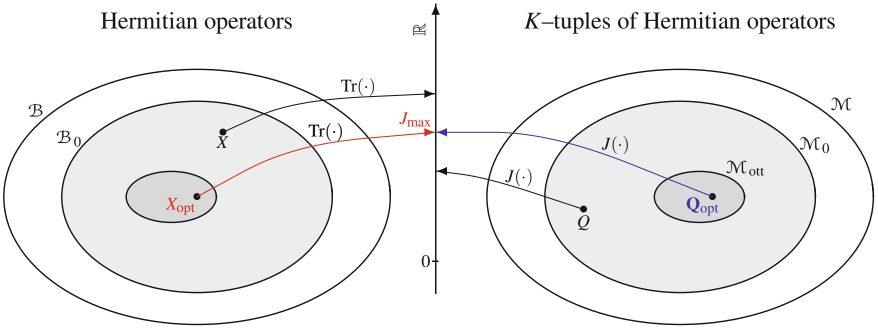


Fig. 5.8 Classes in the quantum decision for the determination of optimal measurement operators. On the *right*, the class $\mathcal{M}$ formed by the K-tuples of Hermitian operators and the subclass $\mathcal{M}_0$ constituted by systems of measurement operators $\mathbf{Q}$; in $\mathcal{M}_0$ the functional $J(\mathbf{Q})$ is defined, which has a maximum $J_{\max}$ when $\mathbf{Q}$ becomes optimal. On the *left*, the class $\mathcal{B}$ of the Hermitian operators X and the subclass $\mathcal{B}_0$ of the positive semidefinite X; in general $\operatorname{Tr}(X) \geq J_{\max}$, but for particular $X = X_{\text{opt}}$ it results $\operatorname{Tr}(X_{\text{opt}}) = J_{\max}$

In other words, each K-tuple $\mathbf{Q} \in \mathcal{M}_0$ identifies a valid *measurement operator system.*

With these premises, the problem of the optimal decision can be treated in the framework of *convex programming*. Starting from the data specified by the *weighted* density operators

$$\widehat{\rho}_i = q_i \rho_i , \qquad i = 0, \ldots, K-1, \tag{5.79}$$

we must determine a measurement operator system $Q \in \mathcal{K}_0$ that maximizes the quantity

$$J(\mathbf{Q}) = \sum_{i=0}^{K-1} \mathrm{Tr}[\widehat{\rho}_i Q_i], \qquad \mathbf{Q} \in \mathcal{M}_0. \tag{5.80}$$

We are dealing with a problem of *convex semidefinite optimization* because the K-tuple $\mathbf{Q}$ must be found on a convex set: in fact, given two K-tuples $\mathbf{P}$ and $\mathbf{Q}$ of $\mathcal{M}_0$ and given any λ with $0 < \lambda < 1$, it can be easily shown that the convex linear combination $\lambda\mathbf{P} + (1-\lambda)\mathbf{Q}$ is still formed by a K-tuple of $\mathcal{M}_0$. Therefore, by definition, $\mathcal{M}_0$ is a convex set. Within such set, it results:

Proposition 5.3 *The functional $J(\mathbf{Q})$, which gives the correct decision probability P_c, in $\mathcal{M}_0$ admits the maximum*

$$J_{\max} = \max_{\mathbf{Q}\in\mathcal{M}_0} J(\mathbf{Q}) = J(\mathbf{Q}_{\mathrm{opt}}).$$

This maximum gives the maximum of the correct decision probability, $P_{c,\max} = J_{\max} = J(\mathbf{Q}_{\mathrm{opt}})$, and $\mathbf{Q}_{\mathrm{opt}}$ is by definition an optimal system of measurement operators.

This proposition will be proved in Appendix section “Proof of Holevo’s Theorem”.

5.9 Holevo’s Theorem

The following theorem, stated by Holevo in (1972) [4], completely characterizes the optimal solution, and is probably one of the most important results of the theory of quantum decision in the last decades.

Theorem 5.1 (Holevo’s Theorem) *In a K-ary system characterized by the weighted density operators $\widehat{\rho}_i = q_i \rho_i$, the measurement operators Q_i are optimal if and only if, having defined the operator*

$$L = \sum_{i=0}^{K-1} Q_i \widehat{\rho}_i , \tag{5.81}$$

it follows that the operators $L - \widehat{\rho}_i$ are PSD, that is,

$$L - \widehat{\rho}_i \in \mathcal{B}_0 \tag{5.82}$$

and, for each $i = 0, \ldots, K-1$,

$$(L - \widehat{\rho}_i)Q_i = 0_{\mathcal{H}}. \tag{5.83}$$

Holevo's theorem, which will be proved in Appendix section "Proof of Holevo's Theorem", determines the conditions that must be verified by an optimal system of measurement operators $\mathbf{Q}_{\text{opt}}$, but does not provide any clue on how to identify it.

An equivalent form of Holevo's theorem, but, as we will see, more appropriate for numerical computation, has been proved by Yuen et al. [5] and, recently, in a detailed form, by Eldar et al. [3]. The result is obtained by transforming the original problem into a *dual problem*, according to a well-known technique of linear programming.

Theorem 5.2 (Dual theorem) *In a K-ary system characterized by the weighted density operators $\widehat{\rho}_i = q_i \rho_i$, the measurement operators Q_i are optimal if and only if there exists a PSD operator, $X \in \mathcal{B}_0$, such that* $\text{Tr}[X]$ *is minimal,*

$$T_{\min} = \min_{X \in \mathcal{B}_0} \text{Tr}[X] \tag{5.84}$$

and for every $j = 0, \ldots, K-1$ the operators $X - \widehat{\rho}_j$ are PSD, $X - \widehat{\rho}_j \in \mathcal{B}_0$. The optimal operators Q_i satisfy the conditions

$$(X - \widehat{\rho}_i)Q_i = 0_{\mathcal{H}}. \tag{5.85}$$

The minimum obtained for $\text{Tr}[X]$ *coincides with the requested maximum of $J(\mathbf{Q})$*

$$T_{\min} = J_{\max} = P_{c,\max}. \tag{5.86}$$

Notice that the conditions imposed on the operators for optimality $X - \widehat{\rho}_i$ are the same as those indicated in Holevo's theorem, imposed on operators $X - \widehat{\rho}_i$. To understand why the dual theorem leads to a lower computational complexity, suppose that the Hilbert space be of finite dimensions n. In Holevo's theorem we must look for a K-tuple of Hermitian operators Q_i, for a total of Kn^2 unknowns; instead, in the dual theorem we must look for the Hermitian matrix X, for a total of n^2 unknowns.

Example 5.2 We want to check that the projectors Q_0 and Q_1, evaluated in Sect. 5.3 with Helstrom's theory, satisfy the conditions of Holevo's theorem. For $K = 2$, the operator (5.81) becomes, bearing in mind the resolution constraint of the identity $Q_0 + Q_1 = I$,

$$L = Q_0\widehat{\rho}_0 + Q_1\widehat{\rho}_1 = (I - Q_1)\widehat{\rho}_0 + Q_1\widehat{\rho}_1 = \widehat{\rho}_0 + Q_1 D$$

where $D = \widehat{\rho}_1 - \widehat{\rho}_0$ is the decision operator introduced in Helstrom's theory (see (5.20)). The conditions (5.83) give

$$Q_1 D Q_0 = 0, \qquad Q_0 D Q_1 = 0$$

which are mutually equivalent. We can verify them using the expressions (5.25) and (5.24), and the orthonormality. We obtain

$$Q_1 D Q_0 = \sum_{\eta_k > 0} |\eta_k\rangle\langle\eta_k| \sum_m \eta_m |\eta_m\rangle\langle\eta_m| \sum_{\eta_h < 0} |\eta_h\rangle\langle\eta_h| = 0.$$

The conditions (5.83) become

$$L - \widehat{\rho}_0 = Q_1 D \geq 0, \qquad L - \widehat{\rho}_1 = -Q_0 D \geq 0.$$

We have

$$Q_1 D = \sum_{\eta_h > 0} |\eta_h\rangle\langle\eta_h| \sum_m \eta_m |\eta_m\rangle\langle\eta_m| = \sum_{\eta_h > 0} \eta_h |\eta_h\rangle\langle\eta_h|$$

which is PSD because $\eta_h > 0$ and $|\eta_h\rangle\langle\eta_h|$ are elementary projectors. Analogously, it can be proved that $-Q_0 D \geq 0$.

5.10 Numerical Methods for the Search for Optimal Operators

As already said, only in some particular cases the problem of the determination of the optimal measurement operators and of the maximum correct decision probability has closed-form solutions. In the other cases, we either restrict ourselves to search for near-optimal solutions, with the SRM measurements, or we must resort to numerical computation. As we are dealing with problems of convex programming, which fall under a very general class of problems, we can use existing very sophisticated software packages, like LMI (linear matrix inequalities) and CSP (convex semidefinite programming), both operating in the MatLab© environment [6, 7].

5.10.1 The MatLab Procedure `Cvx`

The use of this procedure is conceptually very simple. For the application of Holevo's theorem, in the general case, all it takes is to provide, as input data, the K weighted density operators $\widehat{\rho}_i$, with the constraints

$$Q_i \geq 0, \qquad i = 0, \dots, K-1, \qquad \sum_{i=0}^{K-1} Q_i = I$$

and to request as output the K measurement operators Q_i that *maximize*

$$J(\mathbf{Q}) = \sum_{i=0}^{K-1} \mathrm{Tr}[\widehat{\rho}_i Q_i].$$

Resorting to the dual theorem reduces the computational complexity. Inputting the $\widehat{\rho}_i$, with the constraints

$$X - \widehat{\rho}_i \geq 0, \qquad i = 0, \dots, K-1$$

the user asks for the operator X of *minimal trace*. From X we obtain the optimal measurement operators as solutions of the equations $(X - \widehat{\rho}_i)Q_i = 0$. Clearly, the computation is simplified because the search is limited to the single operator X.

We write the MatLab procedure in the binary case, which is easily extended to an arbitrary K.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% cvx  procedure applied to Holevo’s theorem
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
cvx_begin SDP
   variable Q0(dim, dim) hermitian
   variable Q1(dim, dim) hermitian
   maximize(trace(Q0*R0+Q1*R1))
   subject to
    Q0>0;
    Q1>0;
    Q0==eye(dim)-Q1;
cvx_end

Pc_Holevo=trace(Q0*R0+Q1*R1);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% cvx procedure apply to the dual problem
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

cvx_begin SDP
   variable Q(dim, dim) hermitian
   minimize(trace(Q))
   subject to
    Q>R0;
    Q>R1;
cvx_end

Pc_dual=trace(Q);
```

5.10.2 Example

We input the weighted density operators

$$\widehat{\rho}_0=\frac{1}{2}\begin{bmatrix}0.29327 & 0.29327 & 0.29327 & 0.17788\\ 0.29327 & 0.29327 & 0.29327 & 0.17788\\ 0.29327 & 0.29327 & 0.29327 & 0.17788\\ 0.17788 & 0.17788 & 0.17788 & 0.12019\end{bmatrix}$$

$$\widehat{\rho}_1=\frac{1}{2}\begin{bmatrix}0.23558 & 0.24519\mathrm{i} & -0.22596 & -0.24519\mathrm{i}\\ -0.24519\mathrm{i} & 0.26442 & 0.24519\mathrm{i} & -0.26442\\ -0.22596 & -0.24519\mathrm{i} & 0.23558 & 0.24519\mathrm{i}\\ 0.24519\mathrm{i} & -0.26442 & -0.24519\mathrm{i} & 0.26442\end{bmatrix}.$$

the “Holevo” procedure gives as output

$$P_e = 0.009316144.$$

$$Q_0=\begin{bmatrix}0.502788 & 0.259735 & 0.247032 & -0.0141425\\ 0.259735 & 0.278855 & 0.259735 & 0.256145\\ 0.247032 & 0.259735 & 0.502788 & -0.0141425\\ -0.0141425 & 0.256145 & -0.0141425 & 0.715569\end{bmatrix}$$

$$Q_1=\begin{bmatrix}0.497212 & -0.259735 & -0.247032 & 0.0141425\\ -0.259735 & 0.721145 & -0.259735 & -0.256145\\ -0.247032 & -0.259735 & 0.497212 & 0.0141425\\ 0.0141425 & -0.256145 & 0.0141425 & 0.284431\end{bmatrix}$$

The “dual” procedure gives as the output

$$P_e = 0.009316139$$

$$X=\begin{bmatrix}0.263349 & 0.138595 & 0.0314089 & 0.0971245\\ 0.138595 & 0.264951 & 0.138595 & -0.0387083\\ 0.0314089 & 0.138595 & 0.263349 & 0.0971245\\ 0.0971245 & -0.0387083 & 0.0971245 & 0.199034\end{bmatrix}$$

Hence we find the same minimum error probability, as expected (the Helstrom procedure gives $P_e = 0.00936141$). The negligible differences are due to the different way of numerical computations.

5.11 Kennedy's Theorem

Holevo's theorem has general validity, because it is concerned with optimal decision in a system specified through density operators, which does not rule out the possibility that the states may be pure. Instead, Kennedy's theorem [8] is about a system in which there is a constellation of K pure states

$$|\gamma_0\rangle, |\gamma_1\rangle, \ldots, |\gamma_{K-1}\rangle. \tag{5.87}$$

Theorem 5.3 (Kennedy's theorem) *In a K-ary system specified by K* **pure states** *$|\gamma_0\rangle, \ldots, |\gamma_{K-1}\rangle$, the optimal projectors (which maximize the correct decision probability) are always elementary, that is, they have the form*

$$Q_i = |\mu_i\rangle\langle\mu_i|, \qquad i = 0, 1, \ldots, K-1 \tag{5.88}$$

where the measurement vectors $|\mu_i\rangle$ must be **orthonormal**.

The theorem is proved in Appendix section "Proof of Kennedy's Theorem".

5.11.1 Consequences of Kennedy's Theorem

With Kennedy's Theorem the search for the optimal decision is substantially simplified, as it is restricted to the search for K *orthonormal measurement vectors*

$$|\mu_0\rangle, |\mu_1\rangle, \ldots, |\mu_{K-1}\rangle$$

from which the optimal projectors are built, using (5.88). The simplification lies in the fact that, instead of searching for K matrices, it suffices to search for K vectors.

Example 5.3 In the binary case, we have seen that the optimal projectors are given by (5.34), where both Q_0 and Q_1 are elementary projectors. In addition, $|\mu_0\rangle$ and $|\mu_1\rangle$ are orthonormal.

From now on, the Hilbert space $\mathcal{H}$ will be assumed of finite dimension n, even though, in the applications to quantum communications systems, the dimensions become infinite ($n = \infty$). When the decision is made starting from K pure states, a fundamental role is played by the *subspace* generated from the states

$$\mathcal{U} = \text{span}(|\gamma_0\rangle, |\gamma_1\rangle, \ldots, |\gamma_{K-1}\rangle) \subseteq \mathcal{H}. \tag{5.89}$$

The dimension of this space, $r = \dim \mathcal{U}$, is equal to K if the states $|\gamma_i\rangle$ are linearly independent (not necessarily orthonormal), and lower than K if the states are linearly independent; so, in general

$$r = \dim \mathcal{U} \le K \le \dim \mathcal{H} = n.$$

In any case, it is very important to observe that:

Proposition 5.4 *It is not restrictive to suppose that the measurement vectors $|\mu_i\rangle$ belong to the space generated by the states*

$$|\mu_i\rangle \in \mathcal{U} \tag{5.90}$$

because any component of the $|\mu_i\rangle$ belonging to the complementary $\mathcal{U}^\perp$ has no influence on the decision probabilities.

In fact, if we decompose $|\mu_j\rangle$ into the sum

$$|\mu_j\rangle = |\mu_j'\rangle + |\mu_j''\rangle, \qquad |\mu_j'\rangle \in \mathcal{U}, \qquad |\mu_j''\rangle \in \mathcal{U}^\perp$$

the transition probabilities become

$$p_c(j|i) = |\langle \mu_j|\gamma_i\rangle|^2 = |\langle \mu_j'|\gamma_i\rangle|^2$$

where $\langle \mu_j''|\gamma_i\rangle = 0$ as $|\mu_j''\rangle \in \mathcal{U}^\perp$ is orthogonal to $|\gamma_i\rangle \in \mathcal{U}$.

Proposition 5.4 is illustrated in Fig. 5.9, where it is evidenced that the states and the measurement vectors belong to the common subspace $\mathcal{U}$. In harmony with Proposition 5.4, we have:

Proposition 5.5 *For the measurement operators, the resolution of the identity can be substituted by the resolution of the generalized identity*

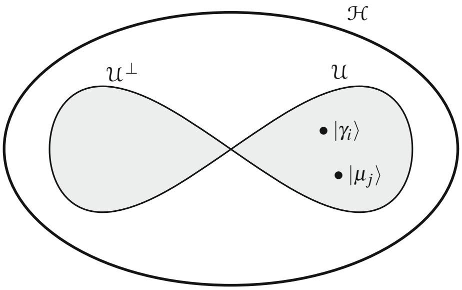


Fig. 5.9 The measurement vectors $|\mu_j\rangle$ belong to the subspace $\mathcal{U}$ generated by the constellation of the states $|\gamma_i\rangle$

$$\sum_{i=0}^{K-1} |\mu_i\rangle\langle\mu_i| = P_{\mathcal{U}} \tag{5.91}$$

where $P_{\mathcal{U}}$ is the projector of $\mathcal{H}$ onto $\mathcal{U}$.

For the proof of this proposition see Sect. 3.7.2. A consequence of the (5.90) is the following:

Proposition 5.6 *The measurement vectors are given by a linear combination of the states*

$$|\mu_i\rangle = \sum_{j=0}^{K-1} a_{ij} |\gamma_j\rangle, \tag{5.92}$$

where the coefficients a_{ij} are in general complex.

Proposition 5.7 *With decision from pure states, the transition probabilities become*

$$p_c(j|i) = |\langle\mu_i|\gamma_j\rangle|^2 \tag{5.93}$$

and the correct decision probability is given by

$$P_c = \sum_{i=0}^{K-1} q_i \, |\langle\mu_i|\gamma_j\rangle|^2. \tag{5.94}$$

5.11.2 Applications of Kennedy's Theorem to Holevo's Theorem

In a decision starting from pure states, the optimal measurement vectors must satisfy Holevo's theorem with $Q_i = |\mu_i\rangle\langle\mu_i|$ and $\widehat{\rho}_i = q_i \, |\gamma_i\rangle\langle\gamma_i|$. Then, assuming that the $|\mu_i\rangle$ belong to the same subspace $\mathcal{U}$ of the states, the geometry relative to the two vector systems is determined by the inner products

$$b_{ij} = \langle\mu_i|\gamma_j\rangle, \qquad i, j = 0, 1, \ldots, K-1. \tag{5.95}$$

Assuming that the $|\mu_i\rangle$ form an orthonormal basis of $\mathcal{U}$ (Fig. 5.10), the inner product b_{ij} can be seen as the projection of $|\gamma_i\rangle$ along the axis $|\mu_j\rangle$. We observe also that the b_{ij} have the important probabilistic meaning

$$p_c(j|i) = |b_{ij}|^2.$$

Using the mixed inner products b_{ij}, from Holevo's theorem we obtain:

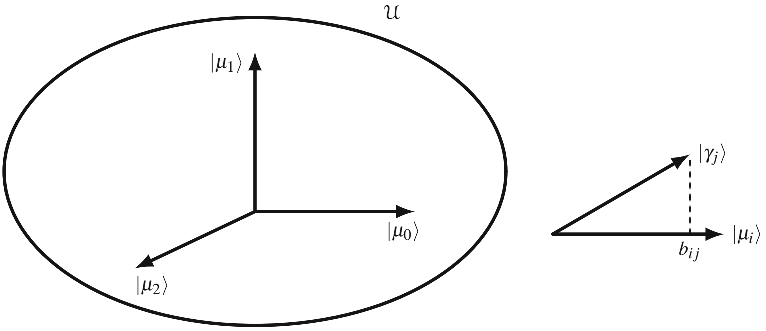


Fig. 5.10 Coordinate systems of $\mathcal{U}$ done by the measurement vectors $|\mu_i\rangle$ and meaning of the mixed inner product $b_{ij} = \langle\gamma_j|\mu_i\rangle$

Corollary 5.1 *In a K-ary system with a constellation of pure states $|\gamma_0\rangle, \ldots, |\gamma_{K-1}\rangle$, the optimal measurement vectors $|\mu_i\rangle$ must verify the conditions*

$$(q_j\, b_{ij} b_{jj}^* - q_i\, b_{ii} b_{ji}^*)|\mu_i\rangle\langle\mu_j| = 0, \qquad \forall i, \forall j \tag{5.96a}$$

$$\sum_{j=0}^{K-1} q_j b_{jj} |\mu_j\rangle\langle\gamma_j| - q_i|\gamma_i\rangle\langle\gamma_i| \geq 0, \qquad \forall i. \tag{5.96b}$$

Relation (5.96a) allows us to write the following conditions on the inner products

$$q_j\, b_{ij} b_{jj}^* - q_i\, b_{ii} b_{ji}^* = 0 \tag{5.97}$$

which can be seen as a nonlinear system of $(K-1)K/2$ equations in the K^2 unknowns b_{ij}. We can add to this other equations derived from the Fourier expansion of the states $|\gamma_i\rangle$ with basis $|\mu_j\rangle$ (see (2.51)), which assumes the form

$$|\gamma_i\rangle = \sum_{j=0}^{K-1} (\langle\mu_j|\gamma_i\rangle)|\mu_j\rangle = \sum_{j=0}^{K-1} b_{ji}|\mu_j\rangle.$$

Then, expressing the inner products $\langle\gamma_i|\gamma_j\rangle$, which we assumed as known, we obtain the relations

$$\sum_{k=0}^{K-1} b_{ki}^* b_{kj} = \langle\gamma_i|\gamma_j\rangle \tag{5.98}$$

which constitute the $(K+1)K/2$ equations.

In principle, we can try to solve this nonlinear system, which admits solutions if the states are linearly independent, and eventually we can verify whether, with these solutions, even the conditions (5.96b) are verified. However, we can see that even in the binary case the search for an exact solution turns out to be rather complicated. We could proceed in numerical form, but in this case it is more convenient to adopt the method derived from the geometric interpretation, as we are going to illustrate.

5.11.3 Geometric Interpretation of Optimization

We consider the subspace $\mathcal{U}$ generated by the states $|\gamma_i\rangle$ in which an orthogonal system of coordinate has been introduced, made of the measurement vectors $|\mu_j\rangle$. The correct decision probability can be expressed in the forms

$$P_c = \sum_{i=0}^{K-1} q_i \; p_c(i|i) = \sum_{i=0}^{K-1} q_i \; |b_{ii}|^2$$

where b_{ij} are the inner products (5.95). If such products are real numbers, we can define the angle θ_i between $|\gamma_i\rangle$ and $|\mu_i\rangle$ from

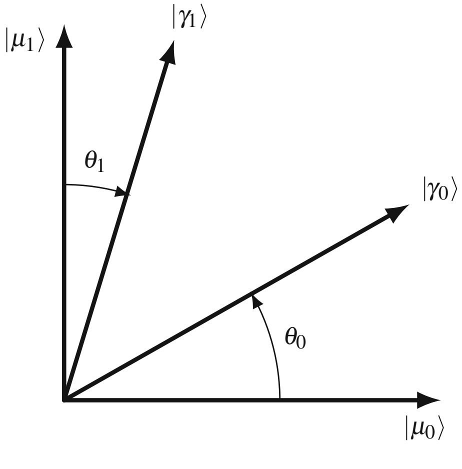
$$\sin^2\theta_i = 1 - b_{ii}^2$$

and then the error probability can be written as

$$P_e = 1 - P_c = \sum_{i=0}^{K-1} q_i \; \sin^2\theta_i \, .$$

The angles θ_i are illustrated in Fig. 5.11 for $K = 2$.

Fig. 5.11 Angles between measurement vectors and states



To minimize P_e we must rotate the constellation of the vectors $|\gamma_i\rangle$ around the respective axes $|\mu_i\rangle$ until a minimum is reached.

This optimization technique has recently been used by the scientists of JPL because it makes it possible to obtain useful results even in the presence of thermal noise [9–11].

5.11.4 Generalization of Kennedy's Theorem

Recently [3], Kennedy's theorem has been partially extended to mixed states and precisely:

Theorem 5.4 *In a system specified by K density operators $\rho_0, \ldots, \rho_{K-1}$, the optimal measurement operators Q_i (maximizing the correct decision probability) have rank not higher than that of the corresponding density operators*

$$\operatorname{rank}(Q_i) \le \operatorname{rank}(\rho_i), \qquad i = 0, 1, \ldots, K-1. \tag{5.99}$$

The connection with the original theorem can be understood considering the consequences on the factors of the operators. If $h_i = \operatorname{rank}(\rho_i)$, the corresponding factor γ_i is an $n \times h_i$ matrix and the measurement factor μ_i has dimensions $n \times \tilde{h}_i$, with $\tilde{h}_i \le h_i = \operatorname{rank}(\rho_i)$, but it is not restrictive to suppose that it has the same dimensions $n \times h_i$ as γ_i (and so we will suppose in the following). In particular, if the ranks are unitary, the factors become kets, $\gamma_i = |\gamma_i\rangle$ and $\mu_i = |\mu_i\rangle$, as established by Kennedy's theorem.

Also the considerations made on the subspace $\mathcal{U}$ generated by the states (see (5.89) and Proposition 5.4) can be generalized. It must be remembered that the state factors are a collection of kets of $\mathcal{H}$ and the state matrix Γ collects these kets. Then the subspace $\mathcal{U}$ is generated according to

$$\mathcal{U} = \operatorname{span}\{\text{kets of } \Gamma\} = \operatorname{Im}\, \Gamma$$

and Proposition 5.4 is extended by saying that it is not restrictive to suppose that the kets of the measurement vectors $|\mu_i\rangle$ belong to the space generated by the states

$$\operatorname{Im} \mu_i \subseteq \mathcal{U}. \tag{5.100}$$

5.12 The Geometry of a Constellation of States

We continue with the study of decision, investigating the geometry generated by the states in the Hilbert space. The basic tools used herein are the eigendecomposition (EID) and the singular value decomposition (SVD). We will refer to pure states, and only at the end of this section the concepts will be extended to mixed states. Also, we refer to equal a priori probabilities, which imply that $q_i = 1/K$; to get general results the states should be replaced by weighted states.

5.12.1 State Matrix and Measurement Matrix

In Sect. 5.7.2 we introduced the state matrix $\varGamma$ and the measurement matrix M, which, with pure states, result in

$$\underset{n\times K}{\varGamma} = [|\gamma_0\rangle, |\gamma_1\rangle, \ldots, |\gamma_{K-1}\rangle], \qquad \underset{n\times K}{M} = [|\mu_0\rangle, |\mu_1\rangle, \ldots, |\mu_{K-1}\rangle].$$

With these matrices, the problem of decision becomes: given the state matrix $\varGamma$, find the measurement matrix M. We have seen that the measurement vectors are given by a linear combination of the states (see (5.92)), that is,

$$|\mu_i\rangle = \sum_{j=0}^{K-1} a_{ij}\,|\gamma_j\rangle\,; \tag{5.101}$$

this combination in matrix terms can be written as

$$\underset{n\times K}{M} = \varGamma\,A, \qquad \underset{K\times K}{A} = [a_{ij}]. \tag{5.102}$$

At this point, the problem is already simplified, because it is sufficient to search for the coefficient matrix A, which is $K \times K$ and therefore of smaller dimensions than the dimensions $n \times K$ of the measurement matrix (where n can become infinite).

It will be useful to compare the matrix expression (5.102) with the following:

$$\underset{n\times K}{M} = \underset{n\times n}{C}\;\underset{n\times K}{\varGamma}\,, \qquad \underset{n\times n}{C} = [c_{ij}] \tag{5.103}$$

which, differently from the linear combination (5.102), gives the relation

$$|\mu_i\rangle = C\,|\gamma_i\rangle, \tag{5.103a}$$

in which the single vector $|\gamma_i\rangle$ is transformed to the vector $|\mu_i\rangle$, with same index i.

Example 5.4 We write explicitly relations (5.102) and (5.103) in the binary case with the purpose of showing how to deal with composite matrices, whose entries are vectors instead of scalar elements.

The matrices $\varGamma$ and M in an n-dimensional Hilbert space, where the kets must be considered as column vectors of size n, are

$$\varGamma = [|\gamma_1\rangle, |\gamma_2\rangle] = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \vdots & \vdots \\ \gamma_{n1} & \gamma_{n2} \end{bmatrix}, \qquad M = [|\mu_1\rangle, |\mu_2\rangle] = \begin{bmatrix} \mu_{11} & \mu_{12} \\ \vdots & \vdots \\ \mu_{n1} & \mu_{n2} \end{bmatrix}.$$

For $K = 2$ relation (5.102) becomes

$$\underset{1\times 2}{M} = \underset{1\times 2}{\varGamma}\ \underset{2\times 2}{A} \quad \to \quad [|\mu_1\rangle, |\mu_2\rangle] = [|\gamma_1\rangle, |\gamma_2\rangle] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{5.104a}$$

and more explicitly

$$\underset{n\times 2}{M} = \underset{n\times 2}{\varGamma}\ \underset{2\times 2}{A} = \begin{bmatrix} \mu_{11} & \mu_{12} \\ \vdots & \vdots \\ \mu_{n1} & \mu_{n2} \end{bmatrix} = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \vdots & \vdots \\ \gamma_{n1} & \gamma_{n2} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}. \tag{5.104b}$$

The different dimensions, as appearing in the two writings above, are justified as follows: in (5.104a) the kets are regarded a single objects of dimensions 1 × 1, whereas in (5.104b) they become $1 \times n$ column vectors.

For $K = 2$ relation (5.103) becomes

$$\underset{1\times 2}{M} = \underset{1\times 1}{C}\ \underset{1\times 2}{\varGamma} \quad \to \quad [|\mu_1\rangle, |\mu_2\rangle] = C\ [|\gamma_1\rangle, |\gamma_2\rangle] \tag{5.105a}$$

and more explicitly

$$\underset{n\times 2}{M} = \underset{n\times n}{C}\ \underset{n\times 2}{\varGamma} \quad \to \quad \begin{bmatrix} \mu_{11} & \mu_{12} \\ \vdots & \vdots \\ \mu_{n1} & \mu_{n2} \end{bmatrix} = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{bmatrix} \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \vdots & \vdots \\ \gamma_{n1} & \gamma_{n2} \end{bmatrix} \tag{5.105b}$$

Now in (5.105a) the matrix C must be regarded as a single object of dimension 1 × 1, and in fact, using this interpretation, it gives explicitly the relation

$$|\mu_1\rangle = C\,|\gamma_1\rangle, \qquad |\mu_2\rangle = C\,|\gamma_2\rangle$$

in agreement with (5.103a).

Problem 5.12 ⋆ Write the relations of Example 5.4 using the results of Helstrom's theory.

5.12.2 Matrices of the Inner Products and of the Outer Products

From the state matrix $\Gamma = [|\gamma_0\rangle, |\gamma_1\rangle, \ldots, |\gamma_{K-1}\rangle]$ two matrices can be formed

$$\underset{K\times K}{G} = \Gamma^* \Gamma, \qquad \underset{n\times n}{T} = \Gamma \Gamma^*. \tag{5.106}$$

The matrix G, called **Gram's matrix**, is the *matrix of inner products* with elements

$$G_{ij} = \langle \gamma_i | \gamma_j \rangle \tag{5.107}$$

while the matrix T gives the sum of the K *outer products*

$$T = \sum_{i=0}^{K-1} |\gamma_i\rangle\langle\gamma_i|. \tag{5.108}$$

These statements can be verified indicating with γ_{ri} the rth element of the column vector $|\gamma_i\rangle$, and performing the operations indicated in (5.106). As T is the sum of elementary operators in the Hilbert space $\mathcal{H}$, also T can be considered an operator of $\mathcal{H}$, which is sometimes called **Gram's operator** (see [12]).

The matrices (5.106) have the following properties:

(1) they are Hermitian semidefinite positive,
(2) both have the same rank as the matrix Γ,
(3) they have the same eigenvalues different from zero (and positive).

Let us prove (3). If λ is an eigenvalue of G, it follows that $G|v\rangle = \lambda|v\rangle$, where $|v\rangle$ is the eigenvector. Then, multiplying this relation by Γ we have

$$\Gamma G|v\rangle = \Gamma\Gamma^*\Gamma|v\rangle = T\Gamma|v\rangle = \lambda\Gamma|v\rangle$$

hence $T|u\rangle = \lambda|u\rangle$ with $|u\rangle = \Gamma|v\rangle$. Then λ is also an eigenvalue of T with eigenvector $\Gamma|v\rangle$. Analogously, we can see that if $\lambda \neq 0$ is an eigenvalue of T with eigenvector $|u\rangle$, we have that λ is also an eigenvalue of G with eigenvector $\Gamma^*|u\rangle$.

The properties (1), (2), and (3) have obvious consequences on the EID of G and of T. Indicating with r the rank and with $\sigma_1^2, \ldots, \sigma_r^2$ the positive eigenvalues, we obtain

$$T = U\Lambda_T U^* = \sum_{i=1}^{r} \sigma_i^2\, |u_i\rangle\langle u_i| = U_r \Sigma_r^2 U_r^* \tag{5.109a}$$

$$G = V\Lambda_G V^* = \sum_{i=1}^{r} \sigma_i^2\, |v_i\rangle\langle v_i| = V_r \Sigma_r^2 V_r^* \tag{5.109b}$$

where

- U is a $n \times n$ unitary matrix,
- $\{|u_i\rangle\}$ is an orthonormal basis of $\mathcal{H}$ formed by the columns of the matrix U,
- Λ_T is an $n \times n$ diagonal matrix whose first r diagonal elements are the positive eigenvalues σ_i^2, and the other $n - r$ diagonal elements are null,
- V is an $K \times K$ unitary matrix,
- $\{|v_i\rangle\}$ is an orthonormal basis of $\mathbb{C}^K$ formed by the columns of V,
- Λ_G is an $K \times K$ diagonal matrix whose first r diagonal elements are the positive eigenvalues σ_i^2 and the other $n - r$ diagonal elements are null,
- U_r and V_r are formed by the first r columns of U and V, respectively,
- $\Sigma_r^2 = \mathrm{diag}[\sigma_1^2, \ldots, \sigma_r^2]$.

In (5.109) appear both the *full* form and the *reduced* form of the EIDs (see Sect. 2.11).

5.12.3 Singular Value Decomposition of Γ

Combining the EIDs of Gram's operator T and of Gram's matrix G we obtain the SVD of the state matrix Γ. The result is (see [13])

$$\Gamma = U\Sigma V^* = \sum_{i=1}^{r} \sigma_i |u_i\rangle\langle v_i| = U_r \Sigma_r V_r^* \tag{5.110}$$

where U, V, U_r, V_r, and Σ_r are the matrices that appear in the previous EIDs, Σ is an $n \times K$ diagonal matrix whose first r diagonal elements are given by the square root σ_i of the positive eigenvalues σ_i^2 of T and G and the other diagonal elements are null.

Before discussing and applying the above decompositions, let us develop a couple of examples.

Example 5.5 Consider a binary system ($K = 2$) on $\mathcal{H} = \mathbb{C}^4$, where the two states are specified by the matrix

$$\Gamma = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

The matrices G and T become respectively

$$G = \Gamma^*\Gamma = \begin{bmatrix} 1 & -\frac{\mathrm{i}}{2} \\ \frac{\mathrm{i}}{2} & 1 \end{bmatrix} \qquad T = \Gamma\Gamma^* = \frac{1}{4}\begin{bmatrix} 2 & -1-\mathrm{i} & 1+\mathrm{i} & 0 \\ -1+\mathrm{i} & 2 & -2 & 1+\mathrm{i} \\ 1-\mathrm{i} & -2 & 2 & -1-\mathrm{i} \\ 0 & 1-i & -1+\mathrm{i} & 2 \end{bmatrix}.$$

The eigenvalues of G are $\sigma_1^2 = 3/2$ and $\sigma_2^2 = 1/2$ and the corresponding EID is

$$G = V \Lambda_G V^* \quad \text{with} \quad V = \frac{1}{2}\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad \Lambda_G = \Sigma_r^2 = \begin{bmatrix} \frac{3}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$$

which coincides with the reduced EID, because $r = K = 2$. The eigenvalues of T are $\sigma_1^2 = 3/2$, $\sigma_2^2 = 1/2$, $\sigma_3 = \sigma_4 = 0$ and the corresponding reduced EID is

$$T = U_r \Sigma_r^2 U_r^* \quad \text{with} \quad U_r = \begin{bmatrix} 0 & 1 \\ \frac{1}{\sqrt{3}} & 0 \\ -\frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{3}} & 0 \end{bmatrix}, \qquad U_r^* = \begin{bmatrix} 0 & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

$$\Sigma_r^2 = \begin{bmatrix} \frac{3}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}.$$

The reduced SVD of Γ is: $\Gamma = U_r \Sigma_r V^*$, where the factors are specified above.

Example 5.6 Consider a constellation composed by two coherent states with real parameters $\pm\alpha$ (see Sect. 3.2.2)

$$|\gamma_1\rangle = |-\alpha\rangle, \qquad |\gamma_2\rangle = |\alpha\rangle, \qquad |\gamma_1\rangle, |\gamma_1\rangle \in \mathcal{G}, \alpha \in \mathbb{R}$$

which, as well known, must be defined on an infinite-dimensional Hilbert space. The purpose of the example is to show that, in spite of the infinite dimensions, eigenvalues and eigenvectors can be developed in finite terms (at least for the parts that are connected to the following applications).

The expressions of the two states are (see (3.4))

$$|\gamma_1\rangle = \sum_{n=0}^{\infty} \mathrm{e}^{-\alpha^2/2} \frac{(\alpha)^n}{\sqrt{n!}} |n\rangle, \qquad |\gamma_2\rangle = \sum_{n=0}^{\infty} \mathrm{e}^{-\alpha^2/2} \frac{(-\alpha)^n}{\sqrt{n!}} |n\rangle \tag{5.111}$$

an so the corresponding matrix becomes

$$\Gamma = [|\gamma_1\rangle, |\gamma_2\rangle] = \sum_{n=0}^{\infty} \frac{\mathrm{e}^{-\alpha^2/2}}{\sqrt{n!}} [\alpha^n, (-\alpha)^n] |n\rangle \tag{5.112}$$

and has dimensions $\infty \times 2$. We can easily see that these two vectors are linearly independent and therefore the rank of Γ is $r = K = 2$.

Gram's matrix is 2×2 and becomes

$$G = \begin{bmatrix} \langle\gamma_1|\gamma_1\rangle & \langle\gamma_1|\gamma_2\rangle \\ \langle\gamma_2|\gamma_1\rangle & \langle\gamma_2|\gamma_2\rangle \end{bmatrix} = \begin{bmatrix} 1 & \gamma_{12} \\ \gamma_{12} & 1 \end{bmatrix} \tag{5.113}$$

where (see (7.10)) $\gamma_{12} = \mathrm{e}^{-2\alpha^2}$, whereas Gram's operator T is infinite dimensional and has a rather complicated expression, that can be obtained from (5.111) developing the outer products as follows:

$$T = |\gamma_1\rangle\langle\gamma_1| + |\gamma_2\rangle\langle\gamma_2|.$$

The eigenvalues of G are given by the solution of the equation

$$\det(G - \lambda I) = (1 - \lambda)^2 - \gamma_{12}^2 = 0$$

and therefore we have, with the notation of (5.113)

$$\sigma_1^2 = 1 + \gamma_{12}, \qquad \sigma_2^2 = 1 - \gamma_{12} \tag{5.114}$$

and the normalized eigenvectors are

$$|v_1\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \qquad |v_2\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

In this way, we have performed the spectral decomposition of G in the form (5.109b) with

$$V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \qquad \Lambda_G = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}.$$

The spectral decomposition of T, given by (5.109a), requires the computation of the eigenvectors $|u_1\rangle, |u_2\rangle$ which are of infinite dimension. In principle, such computation can be done, but it is very complicated, and so the vectors, for now, are left indicated in a nonexplicit form.

The singular value decomposition of Γ results in

$$\Gamma = \sigma_1|u_1\rangle\langle v_1| + \sigma_2|u_2\rangle\langle v_2|$$

where the singular values are $\sigma_{1,2} = \sqrt{1 \pm \gamma_{12}}$.

5.12.4 Spaces, Subspaces, Bases, and Operators

In the above decompositions several spaces and subspaces come into play. The reference environment is the Hilbert space $\mathcal{H}$, which is assumed of dimension n. We then have the subspace generated by the states

$$\mathcal{U} = \mathrm{span}(|\gamma_0\rangle, |\gamma_1\rangle, \dots |\gamma_{K-1}\rangle)$$

of dimension r, which is also the subspace where the measurement vectors $|\mu_i\rangle$ operate (see Fig. 5.9). The unitary operator

$$\underset{n\times n}{U} = [\,|u_1\rangle, \dots, |u_n\rangle] \;:\; \mathcal{H} \to \mathcal{H}$$

provides with its n columns an orthonormal basis for $\mathcal{H}$, while its first r columns, corresponding to the non-null eigenvalues σ_i^2, form a basis for the subspace $\mathcal{U}$

$$\mathcal{U} = \operatorname{span}(|u_1\rangle, \dots, |u_r\rangle) \subseteq \mathcal{H}.$$

These r eigenvectors were collected in the matrix U_r that appears in the reduced EID of T (see (5.109a)); the remaining $n-r$ eigenvectors $|u_{r+1}\rangle, \dots, |u_n\rangle$ generate the complementary space $\mathcal{U}^\perp$. Then the following resolutions are found

$$\sum_{k=1}^{n} |u_k\rangle\langle u_k| = U\,U^* = I_{\mathcal{H}}, \qquad \sum_{k=1}^{r} |u_k\rangle\langle u_k| = U_r\,U_r^* = P_{\mathcal{U}} \tag{5.115}$$

where $P_{\mathcal{U}}$ is the *projector* on $\mathcal{U}$. Analogously, the unitary operator ($K \times K$ matrix)

$$V = [\,|v_1\rangle, \dots, |v_K\rangle] \;:\; \mathbb{C}^K \to \mathbb{C}^K$$

provides with its K columns a basis for $\mathbb{C}^K$, while its first r columns provide a basis for an r-dimensional subspace $\mathcal{V}$ of $\mathbb{C}^K$.

$$\operatorname{span}(|v_1\rangle, \dots, |v_r\rangle) = \mathcal{V} \subseteq \mathbb{C}^K.$$

We obtain the resolutions

$$\sum_{k=1}^{K} |v_k\rangle\langle v_k| = V\,V^* = I_K, \qquad \sum_{k=1}^{r} |v_k\rangle\langle v_k| = V_r\,V_r^* = P_{\mathcal{V}}. \tag{5.116}$$

The state matrix defines a *linear transformation*[8]

$$\Gamma \;:\; \mathbb{C}^K \to \mathcal{H}$$

because it “accepts” at the input a ket $|v\rangle \in \mathbb{C}^K$ and produces the ket $\Gamma\,|v\rangle \in \mathcal{H}$. The image of Γ is

$$\operatorname{im}\Gamma = \mathcal{U}.$$

[8] The term *operator*, in practice represented by a square matrix, is reserved to *linear transformations* from one space to the same space.

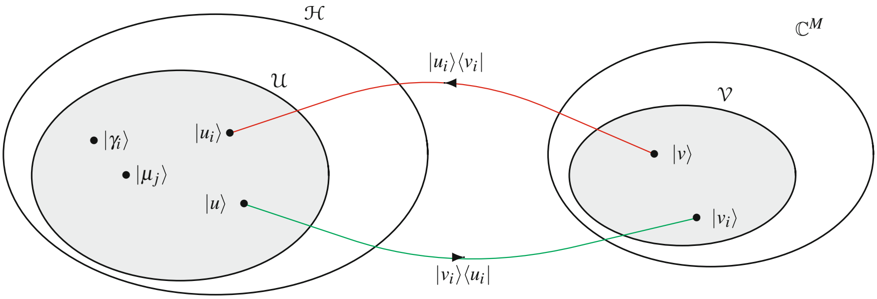


Fig. 5.12 Spaces and subspaces generated by a constellation of states. In *red* and *green* the transjectors

Analogously, the adjoint matrix $\Gamma^* : \mathcal{H} \to \mathbb{C}^K$ operates on a ket $|u\rangle \in \mathcal{H}$ and returns the ket $\Gamma^*|u\rangle \in \mathbb{C}^K$. The image of Γ^* is: im $\Gamma^* = \mathcal{V}$. The connection between $\mathbb{C}^K$ and $\mathcal{H}$ is made by the elementary operators $|u_i\rangle\langle v_i|$ appearing in the SVD (5.110). These operators transform a ket $|v\rangle$ of $\mathbb{C}^K$ to the ket

$$|u_i\rangle\langle v_i|v\rangle = k_i|u_i\rangle \in \mathcal{H}, \qquad \text{with} \quad k_i = \langle v_i|v\rangle$$

and, because they provide a transfer (from $\mathbb{C}^K$ to $\mathcal{H}$ and from $\mathcal{H}$ to $\mathbb{C}^K$), they are named "transjectors" in [14] (Fig. 5.12).

Analogously, the connection between $\mathcal{H}$ and $\mathbb{C}^K$ is done by the elementary operators $|v_i\rangle\langle u_i|$ of the SVD (5.110).

5.12.5 The Geometry with Mixed States

All the above considerations, referring to pure states, can be extended in a rather obvious way to mixed states with some dimensional changes. The starting point is the matrix of the states, which now collects the factors γ_i of the density operators ρ_i

$$\underset{n\times H}{\Gamma} = \left[\gamma_0, \gamma_1, \ldots, \gamma_{K-1}\right] \tag{5.117}$$

where the number of the columns $H = h_0 + h_1 + \cdots h_{K-1}$ is given by the total number of columns of the state factors γ_i. As we have seen in (5.68), this matrix can be considered as a collection of H kets of $\mathcal{H}$, which generate the subspace $\mathcal{U}$, whose dimension r is always given by the rank of Γ.

Gram's operator has the expressions

$$\underset{n\times n}{T} = \Gamma\Gamma^* = \sum_{i=0}^{K-1} \gamma_i\gamma_i^* = \sum_{i=0}^{K-1} \rho_i \tag{5.118}$$

and therefore can be directly evaluated from the ρ_i, without finding their factorizations. Its dimensions remain $n \times n$. Instead, Gram's matrix becomes $H \times H$ and has the structure

$$\underset{H\times H}{G} = \Gamma^* \Gamma = \begin{bmatrix} \gamma_0^* \gamma_0 & \cdots & \gamma_0^* \gamma_{K-1} \\ \vdots & \ddots & \vdots \\ \gamma_{K-1}^* \gamma_0 & \cdots & \gamma_{K-1}^* \gamma_{K-1} \end{bmatrix} \tag{5.119}$$

where the $\gamma_i^* \gamma_j$ are not ordinary inner products, but matrices of dimensions $h_i \times h_j$.

Finally, the subspace $\mathcal{V}$ becomes of dimensions $H \geq K$. This part concerning mixed states will be further developed in Chap. 8.

5.12.6 Conclusions

We have seen that a constellation of states (or of state factors) gathered in the matrix Γ, can be defined on the Hilbert space $\mathcal{H}$ and, more precisely, on its subspace $\mathcal{U}$, generating several operators.

It remains to evaluate the measurement matrix M identifying the measurement operators. To get specific results we must state the objective, which, in the context of quantum communications, is the *maximization of the correct decision probability*. An alternative objective, which brings to a suboptimal solution, is to *minimize the quadratic error between the states and the corresponding measurement vectors*. This technique, called *square root measurement* (SRM), will be seen in the next chapter.

5.13 The Geometrically Uniform Symmetry (GUS)

The set of the states (constellation) can have a symmetry that facilitates its study and its performance evaluation. The kind of symmetry that allows for these simplifications is called *geometrically uniform symmetry* (GUS) and is verified in several quantum communications systems, like the quantum systems obtained with the modulations PSK and PPM and all the binary systems.[9]

5.13.1 The Geometrically Uniform Symmetry with Pure States

A constellation of K pure states

$$\{|\gamma_0\rangle, |\gamma_1\rangle, \ldots, |\gamma_{K-1}\rangle\}$$

[9] The interest of the GUS is confined to the case in which the a priori probabilities are equal ($q_i = 1/K$).

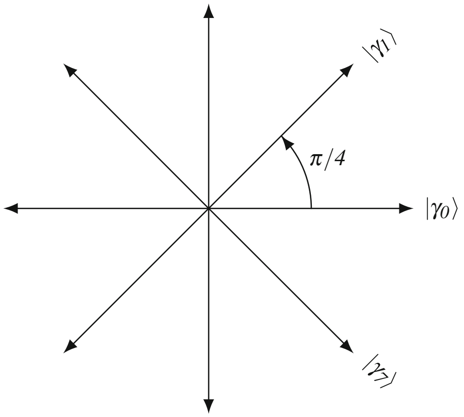


Fig. 5.13 Constellation of states with the geometrically uniform symmetry in the complex plane $\mathbb{C}$. The reference state is $|\gamma_0\rangle = 1$ (the complex number 1) and the symmetry operator is $S = \mathrm{e}^{\mathrm{i}\pi/4}$

has the *geometrically uniform symmetry* when the two properties are verified:

(1) the K states $|\gamma_i\rangle$ are obtained from a single reference state $|\gamma_0\rangle$ in the following way

$$|\gamma_i\rangle = S^i|\gamma_0\rangle, \qquad i = 0, 1, \ldots, K-1 \tag{5.120a}$$

where S is a unitary operator, called *symmetry operator*;

(2) the operator S is a Kth root of the identity operator in the sense that

$$S^K = I_{\mathcal{H}}. \tag{5.120b}$$

An elementary example of constellation that verifies the GUS is given by the K roots of unity in the complex plane, as shown in Fig. 5.13 for $K = 8$.

In the presence of the GUS, the specification of the constellation is limited to the reference state $|\gamma_0\rangle$ and to the symmetry operator S. In addition, it simplifies the decision, because, as we shall see for the optimal decision, we can choose the measurement vectors with the same symmetry as the states, that is,

$$|\mu_i\rangle = S^i|\mu_0\rangle, \qquad i = 0, 1, \ldots, K-1. \tag{5.121}$$

In the next chapter we will verify that the PSK and PPM systems have the GUS. Here we limit ourselves to the binary case.

5.13.2 All Binary Constellations Have the GUS

A constellation of two arbitrary states, $|\gamma_0\rangle$ and $|\gamma_1\rangle$, is always geometrically uniform, with symmetry operator S defined by [14]

$$S = I_{\mathcal{H}} - 2\frac{|w\rangle\langle w|}{\langle w|w\rangle} \tag{5.122}$$

where $|w\rangle = |\gamma_1\rangle - |\gamma_0\rangle$ if the two states have inner product $X := \langle\gamma_0|\gamma_1\rangle$ real. In this case S is a "reflector," which reflects a state with respect to the hyperplane (bisector) determined by the vectors $|\gamma_0\rangle$ and $|\gamma_1\rangle$. It can be verified from definition (5.122) that S is unitary and $S^2 = I_{\mathcal{H}}$ (see problems).

If the inner product X is complex, $X = |X|\mathrm{e}^{\mathrm{i}\phi}$, we modify $|\gamma_1\rangle$ as $|\tilde{\gamma}_1\rangle = \mathrm{e}^{-\mathrm{i}\phi}|\gamma_1\rangle$ and apply (5.122) to the states $|\gamma_0\rangle$ and $|\tilde{\gamma}_1\rangle$, which have a real inner product. This does not represent any restriction because $|\gamma_1\rangle$ and $|\tilde{\gamma}_1\rangle$ differ by a phase factor and therefore represent the same physical state.

5.13.3 The GUS with Mixed States

The definition of GUS is now extended to mixed states. A constellation of K density operators

$$\{\rho_0, \rho_1, \ldots, \rho_{K-1}\}$$

has the *geometrically uniform symmetry* when the following two properties are verified:

(1) the K operators ρ_i are obtained from a single reference operator ρ_0 as

$$\rho_i = S^i \rho_0 \, (S^i)^*, \qquad i = 0, 1, \ldots, K-1 \tag{5.123}$$

where S is a unitary operator called *symmetry operator*;

(2) the operator S is a Kth root of the identity operator

$$S^K = I_{\mathcal{H}}. \tag{5.123b}$$

This extension is in harmony with the fact that with pure states the density operators become $\rho_i = |\gamma_i\rangle\langle\gamma_i|$. In addition, with the factorization of the density operators, $\rho_i = \gamma_i\gamma_i^*$, relation (5.123) gives

$$\gamma_i = S^i \, \gamma_0, \qquad i = 0, 1, \ldots, K-1 \tag{5.124}$$

which generalizes (5.120a). In the context of optimal decision [3] we will prove that the same symmetry is transferred to the measurement operators, and also to the measurement factors, namely,

$$\mu_i = S^i \mu_0, \qquad i = 0, 1, \ldots, K-1. \tag{5.125}$$

5.13.4 Generalizations of the GUS

The GUS can be generalized in two ways. We limit ourselves to introducing the two generalizations in the case of pure states. In the first generalization [3], we have L reference states $|\gamma_{01}\rangle, \ldots, |\gamma_{0L}\rangle$, instead of a single state $|\gamma_0\rangle$, and the constellation is subdivided into L subconstellations generated by a single symmetry operator S in the form $|\gamma_{ik}\rangle = S^i|\gamma_{0k}\rangle$. An example of modulation that has this kind of *composite* GUS is the Quadrature Amplitude Modulation (QAM), which will be seen in Chap. 7.

In the second type of generalization [14], we have K distinct symmetry operators S_i, made up of K unitary matrices forming a multiplicative group, and each state of the constellation is generated in the form $|\gamma_i\rangle = S_i|\gamma_0\rangle$ from a single reference state $|\gamma_0\rangle$.[10]

5.13.5 Eigendecomposition of the Symmetry Operator

The EID of the symmetry operator S plays an important role in the analysis of Communications Systems having the GUS. We give the two equivalent forms of EIDs of S (see Sects. 2.10 and 2.11)

$$S = \sum_{i=1}^{k} \lambda_i \, P_i , \qquad S = Y \, \Lambda \, Y^* = \sum_{i=0}^{n-1} \overline{\lambda}_i |y_i\rangle\langle y_i| \tag{5.126}$$

where $\{\lambda_i, i = 1, \ldots, k\}$ are the distinct eigenvalues of S, $\{P_i, i = 1, \ldots, k\}$ form a projector system, that is, with $P_i P_j = \delta_{ij} P_i$, Y is an $n \times n$ unitary matrix, and $\Lambda = \text{diag}[\overline{\lambda}_1, \ldots, \overline{\lambda}_n]$ contains the nondistinct eigenvalues. In general, the distinct eigenvalues λ_i have a multiplicity $c_i \geq 1$.

Considering that S is a unitary operator, the $\overline{\lambda}_i$ have unitary amplitude and, because $S^K = I_{\mathcal{H}}$, the eigenvalues have the form

$$\lambda_i = W_K^{r_i}, \qquad 0 \leq r_i < K \tag{5.127}$$

where $W_K := \mathrm{e}^{\mathrm{i}2\pi/K}$ and r_i are integers. Now, in the second EID, collecting the elementary projectors $|y_j\rangle\langle y_j|$ with a common eigenvalue, we arrive at the form

$$S = \sum_{i=0}^{K-1} W_K^i \, Y_i \, Y_i^* \tag{5.128}$$

[10] In the literature [3] the set of the states that satisfy (5.120) is called *cyclic state set*, whereas the term *geometrically uniform symmetry* indicates the general case, which is obtained with a multiplicative group of unitary matrices.

where Y_i are $n \times c_i$ matrices, with c_i the multiplicity of $\lambda_i = W_K^i$. Note that the projectors are given by $P_i = Y_i Y_i^*$.

Example 5.7 In the PSK the symmetry operator is given by

$$S = \mathrm{diag}[W_K^k, k = 0, 1, \ldots, K-1]. \tag{5.129}$$

As S is diagonal, its EID is immediately found as $S = I_n\, S\, I_n^*$, with I_n the identity matrix. For example, for $K = 3$ and $n = 6$, we have tree distinct eigenvalues

$$\Lambda = \mathrm{diag}[1, W_3, W_3^2, 1, W_3, W_3^2]$$

and the EID results in

$$S = \begin{bmatrix} 1&0&0&0&0&0\\ 0&1&0&0&0&0\\ 0&0&1&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&1&0\\ 0&0&0&0&0&1 \end{bmatrix} \begin{bmatrix} 1&0&0&0&0&0\\ 0&W_3&0&0&0&0\\ 0&0&W_3^2&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&W_3&0\\ 0&0&0&0&0&W_3^2 \end{bmatrix} \begin{bmatrix} 1&0&0&0&0&0\\ 0&1&0&0&0&0\\ 0&0&1&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&1&0\\ 0&0&0&0&0&1 \end{bmatrix}$$

Now, to obtain the form (5.128), we must collect in the matrices Y_i the eigenvectors corresponding to the eigenvalues W_3^i. Thus

$$Y_0 = \begin{bmatrix} 1&0\\ 0&0\\ 0&0\\ 0&1\\ 0&0\\ 0&0 \end{bmatrix}, \qquad Y_1 = \begin{bmatrix} 0&0\\ 1&0\\ 0&0\\ 0&0\\ 0&1\\ 0&0 \end{bmatrix}, \qquad Y_2 = \begin{bmatrix} 0&0\\ 0&0\\ 1&0\\ 0&0\\ 0&0\\ 0&1 \end{bmatrix}.$$

5.13.6 Commutativity of S with T

An important property with GUS, proved in Appendix section "Commutativity of the Operators T and S", is given by:

Proposition 5.8 *Gram's operator and the symmetry operator of the GUS commute*

$$TS = ST. \tag{5.130}$$

This leads to the simultaneous diagonalization (see Theorem 2.4) of T and S, stated by

$$T = U\, \Sigma^2\, U^*, \qquad S = U \Lambda U^*. \tag{5.131}$$

Note that, in general, the eigenvalues of the symmetry operator are multiple and then the diagonalization of S is not unique (see Theorem 2.3 of Sect. 2.11). This multiplicity will be used to find useful simultaneous decompositions, as will be seen at the end of Chap. 8.

Problem 5.13 ⋆⋆ Prove that the quantum states of $\mathcal{H} = \mathbb{C}^4$

$$|\gamma_0\rangle = \frac{1}{2}[1, -1, 1, -1]^{\mathrm{T}}, \qquad |\gamma_1\rangle = \frac{1}{2}[1, 1, -1, 1]^{\mathrm{T}}$$

verify the GUS for a binary transmission. Find the symmetry operator S, verify that S has the properties of a symmetry operator and that $|\gamma_1\rangle$ is obtained from $|\gamma_0\rangle$ as $|\gamma_1\rangle = S\,|\gamma_0\rangle$.

Problem 5.14 ⋆ Find the EID of the symmetry operator S of the previous problem.

Problem 5.15 ⋆⋆ Prove that the two quantum states of $\mathcal{H} = \mathbb{C}^4$

$$|\gamma_0\rangle = \frac{1}{2}[1, -1, 1, -1]^{\mathrm{T}}, \qquad |\gamma_1\rangle = \frac{1}{2}[1, 1, -\mathrm{i}, 1]^{\mathrm{T}}$$

verify the GUS for a binary transmission, and find the corresponding symmetry operator S. Note that in this case the inner product $X := \langle\gamma_0|\gamma_1\rangle$ is complex.

5.14 Optimization with Geometrically Uniform Symmetry

In the general case of *weighted* density operators the geometrically uniform symmetry (GUS) is established by the condition

$$\widehat{\rho}_i = S^i \widehat{\rho}_0\,(S^i)^*, \qquad i = 0, 1, \ldots, K-1. \tag{5.132}$$

In such case, the search for the optimal measurement operators is simplified because the data are restricted to the reference operator $\widehat{\rho}_0$ and to the symmetry operator S, and in addition the search can be restricted to the measurement operator Q_0 only.

5.14.1 *Symmetry of the Measurement Operators*

The GUS is transferred also to the measurement operators, according to

Proposition 5.9 *If the weighted density operators have the GUS, established by (5.132), it is not restrictive to suppose that also the optimal measurement operators have the GUS, with the same symmetry operator, namely*

$$Q_i = S^i\, Q_0\, S^{-\mathrm{i}}, \qquad i = 0, 1, \ldots, K-1. \tag{5.133}$$

Proof Holevo's theorem ensures that there exists a system of optimal measurement operators $\mathbf{Q} = \mathbf{Q}_{\text{opt}} \in \mathcal{M}_0$ that maximizes the functional $J(\mathbf{Q})$ defined by (5.80). The point here is to prove that from this system, which does not necessarily enjoy the GUS, another system can be obtained $\tilde{\mathbf{Q}} \in \mathcal{M}_0$ that enjoys the GUS and has the same properties as the original system. To this end, we define

$$\tilde{Q}_0 = \frac{1}{K}\sum_{i=0}^{K-1} S^{-\mathrm{i}} Q_i S^i, \qquad \tilde{Q}_i = S^i\, \tilde{Q}_0 S^{-i} \quad , \quad i = 1, \ldots, K-1.$$

We soon verify that the new operators are PSD. In addition

$$\sum_{i=0}^{K-1} \tilde{Q}_i = \frac{1}{K}\sum_{i=0}^{K-1}\sum_{j=0}^{K-1} S^{i-j} Q_j S^{-(i-j)} = \frac{1}{K}\sum_{j=0}^{K-1}\sum_{k=0}^{K-1} S^k Q_j S^{-k}$$

where the periodicity of the symmetry operator S is used. Then

$$\sum_{i=0}^{K-1} \tilde{Q}_i = \frac{1}{K}\sum_{k=0}^{K-1} S^k \sum_{j=0}^{K-1} Q_j S^{-k} = \frac{1}{K}\sum_{k=0}^{K-1} S^k S^{-k} = I_{\mathcal{H}}.$$

We conclude that the new operators $\tilde{Q}_i$ are legitimate measurement operators. We have also

$$\begin{aligned} J(\tilde{\mathbf{Q}}) &= \sum_{i=0}^{K-1} \mathrm{Tr}[\widehat{\rho}_i\, \tilde{Q}_i] = \sum_{i=0}^{K-1} \mathrm{Tr}[S^i \widehat{\rho}_0\, \tilde{Q}_0 S^{-i}] \\ &= \sum_{i=0}^{K-1} \mathrm{Tr}[\widehat{\rho}_0\, \tilde{Q}_0] = K\,\mathrm{Tr}[\widehat{\rho}_0)\, \tilde{Q}_0] \\ &= \mathrm{Tr}\left[\widehat{\rho}_0 \sum_{i=0}^{K-1} S^{-i} Q_i S^i\right] = \sum_{i=0}^{K-1} \mathrm{Tr}[S^i \widehat{\rho}_0 S^{-i} Q_i] = J(\mathbf{Q}) \end{aligned}$$

so that even the new measurement operators are optimal. □

We must observe also that, choosing measurement operators that enjoy the GUS, for the maximum correct decision probability we simply have

$$P_{c\max} = J(\mathbf{Q}_{\text{opt}}) = K\,\mathrm{Tr}[\widehat{\rho}_0\, Q_{0,\text{opt}}] \tag{5.134}$$

where $Q_{0,\text{opt}}$ (to be found) identifies the optimal measurement operator system.

5.14.2 Holevo's Theorem with GUS

From the previous results, Holevo's theorem becomes:

Theorem 5.5 (Holevo's theorem with GUS) *In a K-ary system characterized by the weighted density operators $\widehat{\rho}_i = q_i \rho_i$, that enjoy the GUS according to (5.132), the optimal measurement operators Q_i can be chosen with the same GUS, according to (5.133). Then the reference operator Q_0 produces a system of optimal operators if and only if, having defined the operator*

$$L = \sum_{i=0}^{K-1} S^i Q_0 \widehat{\rho}_0 S^{-i}, \tag{5.135}$$

we have that the operator $L - \widehat{\rho}_0$ is PSD and verifies the condition $(L - \widehat{\rho}_0) Q_0 = 0_{\mathcal{H}}$. We also have that S commutes with L.

In fact, the operator L is obtained by (5.83) substituting the symmetry expressions (5.132) and (5.133). We can also verify that

$$L = S^i L S^{-i} \quad \text{for every integer} \quad i \tag{5.136}$$

from which we obtain, in particular, that S and L commute. From (5.136) we can prove that, if $L - \widehat{\rho}_0$ is PSD, so are $L - \widehat{\rho}_i$, and that, if $(L - \widehat{\rho}_0) Q_0 = 0_{\mathcal{H}}$, also $(L - \widehat{\rho}_i) Q_i = 0_{\mathcal{H}}$, so that all the conditions of Holevo's theorem are verified.

Even the dual theorem is simplified taking the following form [15]:

Theorem 5.6 (Dual theorem with GUS) *In a K-ary system characterized by the weighted density operators $\widehat{\rho}_i = q_i \rho_i$ that enjoy the GUS with symmetry operators S, a measurement operator system $\{Q_i\}$ that enjoy the GUS is optimal if there exists a PSD operator X with the properties: (1) $X \geq \rho_0$, (2) $X\,S = S\,X$, and (3)* Tr$[X]$ *is minimal. The operator Q_0 that generates the optimal operators satisfies the condition $(X - \widehat{\rho}_0) Q_0 = 0_{\mathcal{H}}$ and the minimum obtained for* Tr$[X]$ *coincides with the requested maximum of $J(\mathbf{Q})$.*

In the assumed conditions we have in fact that $X = S^i X S^{-i}$ for every i. Thus $X - \widehat{\rho}_i = S^i (X - \widehat{\rho}_0) S^{-i}$ is PSD and $(X - \widehat{\rho}_i) Q_i = S^i (X - \widehat{\rho}_0) S^{-i} = 0_{\mathcal{H}}$, in such a way that the conditions of the theorem dual to Holevo's theorem are satisfied.

Note that, in the presence of GUS, the quantum source and the optimal decision become completely specified by the symmetry operator S and by the reference operators ρ_0 and Q_0 (or by their factors γ_0 and μ_0). This has a consequence also in the simplification of convex linear programming (CSP).

5.14.3 Numerical Optimization with MatLab©

In the presence of GUS, referring to Theorem 5.6, the input data are reduced to the weighted density $\widehat{\rho}_0$ and to the symmetry operator S. The constraints to be applied are

$$X - \widehat{\rho}_0 \geq 0, \qquad XS = SX$$

and the requested output is the operator X of minimal trace.

In MatLab the use of the cvx procedure seen in Sect. 5.10 becomes

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% cvx procedure applied to the dual problem with GUS
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 cvx_begin
variables X(dim)
minimize(trace(X))
subject to
 X>rho0;
 X*S==S*X;
  cvx_end
```

Applications of this simplified procedure to Quantum Communications systems will be seen in Chaps. 7 and 8.

5.15 State Compression in Quantum Detection

Quantum detection is formulated in an n-dimensional (possibly infinite) Hilbert space $\mathcal{H}$, but in general, the quantum states and the corresponding measurement operators span an r-dimensional subspace $\mathcal{U}$ of $\mathcal{H}$, with $r \leq n$. Quantum detection could be restricted to this subspace, but the operations involved are redundant for $r < n$, since the kets in $\mathcal{U}$ have n components, as the other kets of $\mathcal{H}$. It is possible and convenient to perform a compression from the subspace $\mathcal{U}$ onto a “compressed” space $\overline{\mathcal{U}}$, where the redundancy is removed (kets are represented by r components). We will show that in the “compressed” space the quantum detection can be perfectly reformulated without loss of information, and some properties become simpler than in the original (uncompressed) Hilbert space $\mathcal{H}$ [16].

State compression has some similarity with **quantum compression**, which will be developed in Chap. 12 in the framework of Quantum Information Theory. Both techniques have the target of representing quantum states more efficiently, but state compression does not consider the information content (entropy) of the states and is based only on geometrical properties.

Before proceeding it is convenient to recall the dimensions, which have a fundamental role in this topic:

- n: dimension of the Hilbert space $\mathcal{H}$,
- K: size of the alphabet,
- r: common rank of Γ, G, and T and dimension of the compressed space $\overline{\mathcal{H}}$.

$$\bullet \quad \text{pure states} \qquad \underset{n\times K}{\Gamma}\;, \qquad \underset{K\times K}{G}\;, \qquad \underset{n\times n}{T}\;, \tag{5.137a}$$

$$\bullet \quad \text{mixed states} \qquad \underset{H\times K}{\Gamma}\;, \qquad \underset{H\times H}{G}\;, \qquad \underset{n\times n}{T}\;, \tag{5.137b}$$

where $H = h_0 + \cdots h_{K-1}$ with h_i the number of columns of the factors γ_i and μ_i.

We will refer to mixed states since they represent the general case and the most interesting one with compression.

5.15.1 State Compression and Expansion

To find the compression operation (and also the expansion) we rewrite the SVD of the state matrix Γ, given by (5.110)

$$\underset{n\times H}{\Gamma} = U\,\Sigma\,V_r^* = U_r\,\Sigma_r\,V_r^* = \sum_{i=1}^{r} \sigma_i |u_i\rangle\langle v_i| \tag{5.138}$$

where $U = [|u_1\rangle, \ldots, |u_n\rangle]$ is an $n \times n$ unitary matrix, $V_r = [|v_1\rangle, \ldots, |v_r\rangle]$ is an $r \times r$ unitary matrix, Σ is an $n \times r$ diagonal matrix whose first r diagonal entries $\sigma_1, \ldots \sigma_r$ are the (positive) singular values, and the other diagonal entries are zero, $\Sigma_r = \text{diag}\{\sigma_1, \ldots \sigma_r\}$ is $r \times r$ diagonal, $U_r = [|u_1\rangle, \ldots, |u_r\rangle]$ is formed by the first r columns of U. We also recall that U_r gives the *projector* operator onto $\mathcal{U}$ as (see (5.115))

$$\sum_{i=1}^{r} |u_i\rangle\langle u_i| = U_r\,U_r^* = P_{\mathcal{U}}. \tag{5.139}$$

In the r-dimensional subspace $\mathcal{U}$ the kets $|u\rangle$ have n components, as in the rest of $\mathcal{H}$, but it is possible to compress each $|u\rangle \in \mathcal{U}$ into a ket $|\overline{u}\rangle$, with $r \le n$ components, without loss of information. The key remark is that for a ket $|u\rangle$ of $\mathcal{U}$ the projection coincides with the ket $|u\rangle$ itself

$$\boxed{P_{\mathcal{U}}\,|u\rangle = |u\rangle, \qquad \forall\,|u\rangle \in \mathcal{U}.} \tag{5.140}$$

Considering (5.139) we can split the identity (5.140) into the pair

$$|\overline{u}\rangle = U_r^*\,|u\rangle, \qquad |u\rangle = U_r\,|\overline{u}\rangle \qquad \forall |u\rangle \in \mathcal{U}$$

where the first relation represents a **compression**, with *compressor* U_r^*, and the second an **expansion**, with *expander* U_r. The compressor U_r^* generates the r-dimensional subspace

$$\overline{\mathcal{U}} := U_r^*\,\mathcal{U} = \{|\overline{u}\rangle = U_r^*|u\rangle,\ |u\rangle \in \mathcal{U}\}$$

and the expander U_r restores the original subspace as $\mathcal{U} = U_r\,\overline{\mathcal{U}}$. In particular the compressed Hilbert space is given by

$$\boxed{\overline{\mathcal{H}} := \overline{\mathcal{U}} = U_r^*\,\mathcal{U}.} \tag{5.141}$$

Compression and expansion are schematically depicted in Fig. 5.14.

Now, all the detection operations in the original Hilbert space $\mathcal{H}$ can be transferred into the compressed space $\overline{\mathcal{U}}$ (here we mark compressed objects with an overline, as $\overline{\mathcal{U}}$). In the transition from $\mathcal{U}$ onto $\overline{\mathcal{U}}$ the geometry of kets is preserved (isometry). In fact, if $|u\rangle, |v\rangle \in \mathcal{U}$ and $|\overline{u}\rangle, |\overline{v}\rangle \in \overline{\mathcal{U}}$ are the corresponding compressed kets, we find for the inner products: $\langle\overline{u}|\overline{v}\rangle = \langle u|P_{\mathcal{U}}|v\rangle = \langle u|v\rangle$. In $\overline{\mathcal{U}}$ the state matrix becomes

$$\underset{r\times r}{\overline{\Gamma}} = \underset{r\times n}{U_r^*}\ \underset{n\times r}{\Gamma} \tag{5.142}$$

and collects the compressed states $\overline{\gamma}_i = U_r^*\gamma_i$. From $\overline{\Gamma}$ we can restore Γ by expansion, as $\Gamma = U_r\,\overline{\Gamma}$. Analogously, for the measurement matrix we find $\overline{M} = U_r^* M$

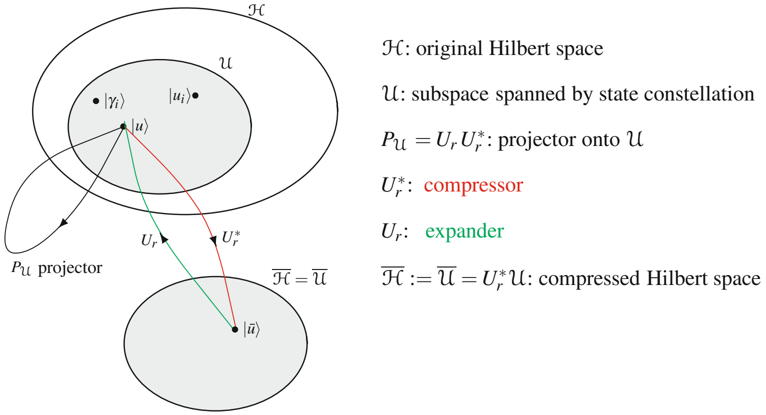

$\mathcal{H}$
$\mathcal{U}$
$|\gamma_i\rangle$
$|u_i\rangle$
$|u\rangle$
$P_{\mathcal{U}}$ projector
U_r
U_r^*
$\overline{\mathcal{H}} = \overline{\mathcal{U}}$
$|\bar{u}\rangle$
$\mathcal{H}$: original Hilbert space
$\mathcal{U}$: subspace spanned by state constellation
$P_{\mathcal{U}} = U_r U_r^*$: projector onto $\mathcal{U}$
U_r^*: compressor
U_r: expander
$\overline{\mathcal{H}} := \overline{\mathcal{U}} = U_r^*\mathcal{U}$: compressed Hilbert space

Fig. 5.14 The geometry for quantum compression: passage from the subspace $\mathcal{U}$ to the "compressed" space $\overline{\mathcal{U}} = U_r^*\mathcal{U}$, where U_r^* is the compressor. $\overline{\mathcal{U}}$ gives the compressed Hilbert space $\overline{\mathcal{H}}$

and $M = U_r \overline{M}$. For the density operators $\rho_i = \gamma_i \gamma_i^*$ and the measurement operators $\Pi_i = \mu_i \mu_i^*$ the compression/expansion give

$$\begin{aligned} \underset{r\times r}{\overline{\rho}_i} &= U_r^* \, \rho_i \, U_r, & \underset{n\times n}{\rho_i} &= U_r \, \overline{\rho}_i \, U_r^* \\ \underset{r\times r}{\overline{\Pi}_i} &= U_r^* \, \Pi_i \, U_r, & \underset{n\times n}{\Pi_i} &= U_r \, \overline{\Pi}_i \, U_r^*. \end{aligned} \tag{5.143}$$

Note that, while $U_r \, U_r^*$ gives the projector $P_{\mathcal{U}}$, $U_r^* \, U_r$ gives the identities

$$\boxed{U_r \, U_r^* = P_{\mathcal{U}}, \qquad U_r^* \, U_r = I_r.} \tag{5.144}$$

In fact $U_r^* \, U_r = \sum_{i=1}^r |u_i\rangle\langle u_i|$, where $|u_i\rangle$ are orthonormal.

5.15.2 Properties in the Compressed Space

We review some properties in the compressed space, starting from the corresponding properties in the original Hilbert space.

Gram operator. The Gram operator $T := \Gamma \Gamma^*$ acting on the original Hilbert space $\mathcal{H}$ has dimension $n \times n$. In the compressed Hilbert space $\overline{\mathcal{H}}$ it becomes

$$\underset{r\times r}{\overline{T}} = U_r^* \, U_r \, \Sigma_r^2 \, U_r^* \, U_r = \Sigma_r^2, \tag{5.145}$$

and therefore the *compressed Gram operator is always diagonal*. On the other hand, the Gram matrix $G := \Gamma^* \, \Gamma$ does not change: $\overline{G} = G$. In fact, compression preserves inner products (see Property (1) in the next subsection).

Probabilities. The relation giving the transition probabilities is exactly preserved in the transition to the compressed space, namely (see Problem 5.16)

$$\boxed{p(j|i) = \text{Tr}[\Pi_j \, \rho_i] = \text{Tr}[\overline{\Pi}_j \, \overline{\rho}_i].} \tag{5.146}$$

Hence the relation for the probability of a correct detection

$$P_c = \sum_{i=0}^{K-1} q_i \text{Tr}[\Pi_i \, \rho_i] = \sum_{i=0}^{K-1} q_i \text{Tr}[\overline{\Pi}_i \, \overline{\rho}_i]. \tag{5.147}$$

This result is very important: it states that, once obtained the compressed operators, for the evaluation of the system performance, **it is not required to return back**

to the original uncompressed space. This conclusion is particularly important in the optimization with convex semidefinite programming (CSP), where the numerical evaluations can be completely carried out in the compressed space.

5.15.3 Compression as a Linear Mapping

Relation (5.144) defines a linear mapping connecting the subspace $\mathcal{U}$ to the compressed space $\overline{\mathcal{H}}$

$$U_r^* : \ \rho \in \mathcal{U} \ \rightarrow \ \overline{\rho} \in \overline{\mathcal{H}}. \tag{5.148}$$

This mapping has several interesting properties:

(1) the compressor U_r^* preserves inner products[11]: $\langle \overline{x}|\overline{y}\rangle = \langle x|y\rangle, \ |x\rangle, |y\rangle \in \mathcal{U}$,
(2) the compression preserves the PSD condition: $\rho \geq 0 \ \rightarrow \ \overline{\rho} \geq 0$,
(3) the compression is *trace preserving*: $\mathrm{Tr}[\overline{\rho}] = \mathrm{Tr}[\rho]$.
(4) the compression preserves the quantum entropy: $S(\overline{\rho}) = S(\rho)$ (see Chap. 12).

We prove statement (1). If $|x\rangle, |y\rangle \in \mathcal{U}$, we get $\langle \overline{x}|\overline{y}\rangle = \langle x|U_r U_r^*|y\rangle = \langle x|P_{\mathcal{U}}|y\rangle$, where $P_{\mathcal{U}}|y\rangle = |y\rangle$ by the fundamental property (5.140). Hence $\langle \overline{x}|\overline{y}\rangle = \langle x|y\rangle$. Similar is the proof of statement (2). The proof of (3) and (4) will be seen in Sect. 12.6.

A final comment. In the context of quantum channels, which will be seen in Sect. 12.8, a compression mapping may be classified as a *noiseless quantum channel*. This is essentially due to the fact that compression is a reversible transformation.

5.15.4 State Compression with GUS

The GUS is preserved in the compressed space (see Problem 5.17 for the proof).

Proposition 5.10 *If the states γ_i have the GUS with generating state γ_0 and symmetry operator S, then the compressed states $\overline{\gamma}_i$ have the GUS with generating state $\overline{\gamma}_0 = U_r^* \gamma_0$ and symmetry operator $\overline{S} = U_r^* \, S \, U_r$.*

The simultaneous diagonalization of T and S seen in Proposition 5.10 is also useful to establish other properties related to the GUS. In fact, by choosing the compressor U_r^* from the common eigenvector matrices U as in Eq. (5.131), we find the further properties:

Proposition 5.11 *With the simultaneous diagonalization the compressed symmetry operator becomes diagonal, with diagonal entries formed by the first r diagonal entries of the matrix Λ.*

[11] An operator from one space to another space is called *isometric* if it preserves norms and inner products [17].

In fact, decomposing Λ in the form diag$[\Lambda_r, \Lambda_c]$, where Λ_r is $r \times r$ and Λ_c is $(n-r) \times (n-r)$, we get

$$\overline{S} = U_r^* \, U \, \Lambda \, U^* \, U_r = [I_r \; 0] \begin{bmatrix} \Lambda_r & 0 \\ 0 & \Lambda_c \end{bmatrix} \begin{bmatrix} I_r \\ 0 \end{bmatrix} = \Lambda_r .$$

Proposition 5.12 *With the simultaneous diagonalization the compressed Gram operator is simply given by*

$$\overline{T} = \text{diag}[K \, \overline{\rho}_0(i,i), i = 1, \dots, r]$$

where $\overline{\rho}_0(i,i)$ are the diagonal entries of the compressed generating density operator $\overline{\rho}_0$.

In fact, $\overline{T} = \sum_{i=0}^{K-1} \overline{S}^i \, \overline{\rho}_0 \, \overline{S}^{-i}$, where $\overline{S}$ is diagonal. Then, the i, j entry is given by

$$\overline{T}(i,j) = \sum_{k=0}^{K-1} \overline{S}^k(i,i) \, \overline{\rho}_0(i,j) \, \overline{S}^{-k}(j,j).$$

In particular, considering that $\overline{S}$ is unitary diagonal, the diagonal entries are

$$\overline{T}(i,i) = \sum_{k=0}^{K-1} \overline{S}^k(i,i) \, \overline{\rho}_0(i,i) \, \overline{S}^{-k}(i,i) = K \, \overline{\rho}_0(i,i)$$

and the evaluation can be limited to these diagonal entries, since $\overline{T}$ is diagonal (see (5.145)) (in general $\overline{\rho}_0$ is not diagonal).

5.15.5 Compressor Evaluation

The leading parameter in compression is the **dimension of the compressed space** r, which is given by the rank of the state matrix Γ, but also by the rank of the Gram matrix G and of the Gram operator T. For the evaluation of the compressor we can use the reduced SVD of Γ, or the reduced EID of G and of T. In any case, for the choice, it is important to have in mind the dimensions of these matrices shown in (5.27).

With pure states, where often the dimension n of the Hilbert space is greater than the alphabet size K and the kets of Γ are linearly independent, r **is determined by the alphabet size** K and the EID of the Gram matrix, of dimension $K \times K$, becomes the natural choice.

With mixed states the choice depends on the specific application. In several cases of practical interest, n may be very large, so that the decompositions represent a very hard numerical task. But, in the presence of GUS, the computational complexity can be reduced, using the commutativity of the Gram operator T with the symmetry operator S. This will be seen in detail in the next chapters in correspondence with the specific applications (see the last two sections of Chap. 8).

Problem 5.16 ⋆⋆ Prove that the evaluation of the transition probabilities in the compressed space is based on the same formula as in the uncompressed space, that is,

$$p(j|i) = \mathrm{Tr}[\Pi_j\, \rho_i] = \mathrm{Tr}[\overline{\Pi}_j\, \overline{\rho}_i].$$

Hint: Use orthonormality relationship $U_r^*\, U_r = I_r$, where I_r is the $r \times r$ identity matrix.

Problem 5.17 ⋆⋆⋆ Prove Proposition 5.10, which states that the GUS is preserved after a compression. *Hint:* Use orthonormality relationship $U_r^*\, U_r = I_r$, where I_r is the $r \times r$ identity matrix.

Problem 5.18 ⋆⋆ Consider the state matrix of $\mathcal{H} = \mathbb{C}^4$

$$\Gamma = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 1 & -1 \\ -1 & 1 \end{bmatrix}$$

Find the compressor U_r^* and the compressed versions of the state matrix Γ and of the Gram operator T.

Problem 5.19 ⋆⋆ Consider a binary transmission where the quantum states are specified by the state matrix of the previous problem. Apply Helstrom's theory with $q_0 = 1/3$ to find the probability of a correct decision P_c. Then apply the compression and evaluate P_c from the compressed states.

Problem 5.20 ⋆⋆ Consider the binary constellation of Problem 5.13, where we determined the symmetry operator S. Find the compressor U_r^* showing, in particular, that the compressed symmetry operator $\overline{S}$ is diagonal.

Appendix

Proof of Holevo's Theorem

We refer to the classes introduced at the beginning of Sect. 5.8 and illustrated in Fig. 5.8. We start by proving Proposition 5.3. The set $\mathcal{M}$ of the K-tuples of Hermitian operators is closed with respect to addition and multiplication by a real number (the sum of two Hermitian operators and the product of a Hermitian operator by a real scalar are Hermitian operators), so it is in fact a real (Kn^2)-dimensional vector space. In such a space, the operation

$$(\mathbf{P},\mathbf{Q})=\sum_{i=0}^{K-1}\mathrm{Tr}[P_iQ_i],\qquad \mathbf{P},\mathbf{Q}\in\mathcal{M},$$

enjoys the property $(\mathbf{P},\mathbf{Q})=(\mathbf{Q},\mathbf{P})$, a consequence of the cyclic property of the trace, as well as of the property $(\mathbf{P},\mathbf{P})\ge 0$ with $(\mathbf{P},\mathbf{P})=0$, only if $\mathbf{P}$ is formed by null operators. Therefore, we are dealing with an operation of **inner product** and so $\mathcal{M}$ is a Hilbert space. In this space, the subset $\mathcal{M}_0$, formed by the K-tuples of PSD operators and resolving the identity, is closed and bounded, and therefore compact. From the classical Weierstrass theorem, in such a set, the continuous functional $J(\mathbf{Q})$ admits a maximum.

We now move on to Holevo's theorem, proving that the conditions indicated are sufficient conditions for maximization. Let $\mathbf{Q}=[Q_0,\dots,Q_{K-1}]\in\mathcal{M}_0$, where the Q_i satisfy the conditions (5.83) and (5.82), and let $\mathbf{P}=[P_0,\dots,P_{K-1}]$ be an arbitrary K-tuple of $\mathcal{M}_0$. Then, recalling the definition of L given by (5.81)

$$\begin{aligned}\sum_{i=0}^{K-1}\mathrm{Tr}[P_i\widehat{\rho}_i]&=\mathrm{Tr}[L]+\sum_{i=0}^{K-1}\mathrm{Tr}[P_i(\widehat{\rho}_i-L)]\\&=\sum_{i=0}^{K-1}\mathrm{Tr}[Q_i\widehat{\rho}_i]-\sum_{i=0}^{K-1}\mathrm{Tr}[P_i(L-\widehat{\rho}_i)].\end{aligned}$$

On the other hand, because the trace of the product of PSD operators is nonnegative, for every i we have $\mathrm{Tr}[P_i(L-\widehat{\rho}_i)]\ge 0$ and $J(\mathbf{P})\le J(\mathbf{Q})$. Therefore, the system $\mathbf{Q}$ is optimal and the sufficiency of the hypothesis of Holevo's theorem is proved.

We can also prove that the definition of L and the condition (5.82) imply the condition (5.83). In fact, we can write

$$0=\mathrm{Tr}[L]-\sum_{i=0}^{K-1}\mathrm{Tr}[Q_i\widehat{\rho}_i]=\sum_{i=0}^{K-1}\mathrm{Tr}[Q_i(L-\widehat{\rho}_i)].$$

As all the terms of the last sum are nonnegative, it must be $\mathrm{Tr}[(L-\widehat{\rho}_i)Q_i]=0$ for every i, then $(L-\widehat{\rho}_i)Q_i=0_{\mathcal{H}}$.

The necessity of the conditions of Holevo's theorem is based on continuity considerations. Let $\mathbf{Q}\in\mathcal{M}_0$ be an optimal system and let U_{jk}, $j,k=0,\dots,K-1$ be operators such that

$$\sum_{j=0}^{K-1}U_{jm}^*U_{jn}=\delta_{mn}I_{\mathcal{H}}.$$

Then, having defined the operators $P_j = S_j^* S_j$, with

$$S_j = \sum_{k=0}^{K-1} U_{jk} Q_k^{1/2},$$

it is easy to verify that $\mathbf{P} = [P_0, \dots, P_{K-1}] \in \mathcal{M}_0$ and, from the optimality of $\mathbf{Q}$, it must be $J(\mathbf{P}) \le J(\mathbf{Q})$.

We now appropriately particularize the operators U_{jk}, $j, k = 0, \dots, K-1$, imposing that $U_{jj} = I_{\mathcal{H}}$ for $j = 2, \dots, K-1$, and

$$\begin{bmatrix} U_{00} & U_{01} \\ U_{10} & U_{11} \end{bmatrix} = \exp\left(\varepsilon \begin{bmatrix} 0_{\mathcal{H}} & -A^* \\ A & 0_{\mathcal{H}} \end{bmatrix} \right)$$

with $\varepsilon > 0$ arbitrarily small and A arbitrary linear operator. Finally, we suppose that all the other operators U_{jk} be null. We then verify that

$$\begin{bmatrix} U_{00} & U_{01} \\ U_{10} & U_{11} \end{bmatrix}^* \begin{bmatrix} U_{00} & U_{01} \\ U_{10} & U_{11} \end{bmatrix} = \begin{bmatrix} I_{\mathcal{H}} & 0_{\mathcal{H}} \\ 0_{\mathcal{H}} & I_{\mathcal{H}} \end{bmatrix}$$

so that the operators U_{jk} satisfy the above conditions. The operators P_j, $j = 2, \dots, K-1$ coincide with the operators Q_j, while, neglecting the infinitesimals ε^2 and those of higher order, we obtain

$$U_{00} = U_{11} = I_{\mathcal{H}}, \qquad U_{01} = -\varepsilon A^*, \qquad U_{10} = \varepsilon A$$

$$S_0 = Q_0^{1/2} - \varepsilon A^* Q_1^{1/2}, \qquad S_1 = Q_1^{1/2} + \varepsilon A Q_0^{1/2}$$

and eventually

$$\begin{aligned} P_0 &= Q_0 - \varepsilon(Q_1^{1/2} A Q_0^{1/2} + Q_0^{1/2} A^* Q_1^{1/2}) \\ P_1 &= Q_1 + \varepsilon(Q_0^{1/2} A^* Q_1^{1/2} + Q_1^{1/2} A Q_0^{1/2}). \end{aligned}$$

It follows that

$$\begin{aligned} J(\mathbf{P}) - J(\mathbf{Q}) &= \sum_{j=0}^{K-1} \mathrm{Tr}[\widehat{\rho}_j (P_j - Q_j] \\ &= \mathrm{Tr}[\widehat{\rho}_0 (P_0 - Q_0)] + \mathrm{Tr}[\widehat{\rho}_1 (P_1 - Q_1)] \\ &= \varepsilon\, \mathrm{Tr}[(\widehat{\rho}_1 - \widehat{\rho}_0)\, (Q_1^{1/2} A Q_0^{1/2} + Q_0^{1/2} A^* Q_1^{1/2})] \\ &= \varepsilon\, \mathrm{Tr}[Q_0^{1/2} (\widehat{\rho}_1 - \widehat{\rho}_0) Q_1^{1/2} A + Q_1^{1/2} (\widehat{\rho}_1 - \widehat{\rho}_0) Q_0^{1/2} A^*]. \end{aligned}$$

As the coefficient of ε must be null to ensure that the difference be non positive for every value of the arbitrary operator A, it must be $Q_0^{1/2}(\widehat{\rho}_1 - \widehat{\rho}_0)Q_1^{1/2} = 0_{\mathcal{H}}$ or, equivalently, $Q_0(\widehat{\rho}_1 - \widehat{\rho}_0)Q_1 = 0_{\mathcal{H}}$. As the reasoning can be repeated for every couple of indexes i and j, it follows that it must be $Q_i(\widehat{\rho}_j - \widehat{\rho}_i)Q_j = 0_{\mathcal{H}}$, that is, $Q_i\widehat{\rho}_j Q_j = Q_i\widehat{\rho}_i Q_j$. Summing both sides with respect to i, we obtain for every j, $\widehat{\rho}_j Q_j = LQ_j$, coinciding with (5.83). At this point, it should be proved that the operators $L - \widehat{\rho}_j$ are PSD. For a rigorous (and very technical) proof of the result, please refer to [3].

Proof of Kennedy's Theorem

Kennedy's theorem (Theorem 5.3) can be derived in a generalized form from Holevo's theorem. We recall that this requires in the first place that the operators $L - \widehat{\rho}_i$ be PSD for every i. If we assume that the eigenvalues of the operators $\widehat{\rho}_i$ span over the entire Hilbert space $\mathcal{H}$, the operator L is positive definite and has rank n. From the optimality conditions of Holevo's theorem

$$(L - \widehat{\rho}_i)Q_i = 0_{\mathcal{H}},$$

we have first of all that, if $|y\rangle$ belongs to the image of the operator Q_i, that is, if there exists $|x\rangle \in \mathcal{H}$ such that $|y\rangle = Q_i|x\rangle$, then $(L - \widehat{\rho}_i)|y\rangle = 0$, and $|y\rangle$ belongs to the null space of the operator $L - \widehat{\rho}_i$. We then have that the image of Q_i is a subspace contained in the null space $\mathcal{N}(L - \widehat{\rho}_i)$ of $L - \widehat{\rho}_i$, therefore its dimension, coinciding with the rank of Q_i, is not greater than the dimension of the null space $\mathcal{N}(L - \widehat{\rho}_i)$ and this yields the inequality

$$\mathrm{rank}(Q_i) \le \dim(\mathcal{N}(L - \widehat{\rho}_i)) = n - \mathrm{rank}(L - \widehat{\rho}_i).$$

From the subadditivity of the rank, i.e., from $\mathrm{rank}(A + B) \le \mathrm{rank}(A) + \mathrm{rank}(B)$, letting $A = L - \widehat{\rho}_i$ and $B = \widehat{\rho}_i$, we obtain $n = \mathrm{rank}(L) \le \mathrm{rank}(L - \widehat{\rho}_i) + \mathrm{rank}(\widehat{\rho}_i)$, which, substituted in the above inequality, yields $\mathrm{rank}(Q_i) \le \mathrm{rank}(\widehat{\rho}_i)$.

Let us now consider the special case in which we have n pure states $|\gamma_i\rangle$, linearly independent, generating the n-dimensional space $\mathcal{H}$, so that the operators $\widehat{\rho}_i$ have rank 1. Then the optimal measurement operators Q_i must have rank not greater than 1, and therefore either be null, or have the form $Q_i = |\mu_i\rangle\langle\mu_i|$. As it must be

$$\sum_{i=0}^{n-1} Q_i = \sum_{i=0}^{n-1} |\mu_i\rangle\langle\mu_i| = I_{\mathcal{H}},$$

the vectors $|\mu_i\rangle$ cannot be null and must be linearly independent. Furthermore, as, for every j,

$$|\mu_j\rangle = \sum_{i=0}^{n-1} |\mu_i\rangle\langle\mu_i|\mu_j\rangle$$

from the comparison of the two sides, we obtain $\langle\mu_i|\mu_j\rangle = \delta_{ij}$ and the measurement vectors are orthonormal.

Commutativity of the Operators T and S

Let us prove Proposition 5.10. Using (5.120a) in the definition of Gram's operator (5.108) and remembering that S is a unitary operator, so that $S^* = S^{-1}$, we obtain

$$T = \sum_{i=0}^{K-1} |\gamma_i\rangle\langle\gamma_i| = \sum_{i=0}^{K-1} S^i|\gamma_0\rangle\langle\gamma_0|S^{-\mathrm{i}}$$

hence

$$\begin{aligned} T\,S &= \sum_{i=0}^{K-1} S^i|\gamma_0\rangle\langle\gamma_0|S^{-\mathrm{i}+1} = S S^{-1}\sum_{i=0}^{K-1} S^i|\gamma_0\rangle\langle\gamma_0|S^{-\mathrm{i}+1} \\ &= S\sum_{i=0}^{K-1} S^{i-1}|\gamma_0\rangle\langle\gamma_0|S^{-\mathrm{i}+1} = S\sum_{k=0}^{K-1} S^k|\gamma_0\rangle\langle\gamma_0|S^{-k} = S\,T \end{aligned}$$

where in the last step we exploited the periodicity of S^i with respect to i.

References

1. C.W. Helstrom, J.W.S. Liu, J.P. Gordon, Quantum-mechanical communication theory. Proc. IEEE **58**(10), 1578–1598 (1970)
2. K. Kraus, *States, Effect and Operations: Fundamental Notions of Quantum Theory*, Lecture Notes in Physics, vol. 190 (Springer, New York, 1983)
3. Y.C. Eldar, A. Megretski, G.C. Verghese, Optimal detection of symmetric mixed quantum states. IEEE Trans. Inf. Theory **50**(6), 1198–1207 (2004)
4. A.S. Holevo, Statistical decision theory for quantum systems. J. Multivar. Anal. **3**(4), 337–394 (1973)
5. H.P. Yuen, R. Kennedy, M. Lax, Optimum testing of multiple hypotheses in quantum detection theory. IEEE Trans. Inf. Theory **21**(2), 125–134 (1975)
6. M. Grant, S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.1. March 2014, http://cvxr.com/cvx
7. M. Grant, S. Boyd, Graph implementations for nonsmooth convex programs, in: *Recent Advances in Learning and Control*, ed. by V. Blondel, S. Boyd, H. Kimura, (eds). Lecture Notes in Control and Information Sciences. (Springer, 2008), pp. 95–110, http://stanford.edu/~boyd/graph_dcp.html

8. R.S. Kennedy, A near-optimum receiver for the binary coherent state quantum channel. Massachusetts Institute of Technology, Cambridge (MA), Technical Report, January 1973. MIT Research Laboratory of Electronics Quarterly Progress Report 108
9. V. Vilnrotter and C.W. Lau, Quantum detection theory for the free-space channel. NASA, Technical Report, August 2001. Interplanetary Network Progress (IPN) Progress Report 42–146
10. V. Vilnrotter, C.W. Lau, Quantum detection and channel capacity using state-space optimization. NASA, Technical Report, February 2002. Interplanetary Network Progress (IPN) Progress Report 42–148
11. V. Vilnrotter, C.W. Lau, Binary quantum receiver concept demonstration. NASA, Technical Report, Interplanetary Network Progress (IPN) Progress Report 42–165, May 2006
12. K. Kato, M. Osaki, M. Sasaki, O. Hirota, Quantum detection and mutual information for QAM and PSK signals. IEEE Trans. Commun. **47**(2), 248–254 (1999)
13. R.A. Horn, C.R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1998)
14. Y.C. Eldar, G.D. Forney, On quantum detection and the square-root measurement. IEEE Trans. Inf. Theory **47**(3), 858–872 (2001)
15. A. Assalini, G. Cariolaro, G. Pierobon, Efficient optimal minimum error discrimination of symmetric quantum states. Phys. Rev. A **81**, 012315 (2010)
16. G. Cariolaro, R. Corvaja, G. Pierobon, Compression of pure and mixed states in quantum detection. in Global Telecommunications Conference, vol. 2011 (GLOBECOM, IEEE, 2011), pp. 1–5
17. A.S. Holevo, V. Giovannetti, Quantum channels and their entropic characteristics. Rep. Prog. Phys. **75**(4), 046001 (2012)