Chapter 6
Quantum Decision Theory: Suboptimization

6.1 Introduction

In the previous chapter we have seen that **optimization** is very difficult, and exact solutions are only known in few cases (binary systems and systems where the state constellation have the *geometrically uniform symmetry*, GUS). To overcome the difficulty, **suboptimization** is considered.

In quantum communications the most important suboptimal criterion is based on the minimization of the quadratic error between the states and the measurement vectors, known by the acronym LSM (least squares measurements), and also SRM (square root measurements), because its solution is based on the square root of an operator.

From a historical point of view, we must start from quantum SRM (square root measurement), introduced by Hausladen and other authors in 1996 [1], who proposed as measurement matrix $M = T^{-1/2}\Gamma$, where T is Gram's operator and $T^{-1/2}$ is its inverse square root. With this choice, the quantum decision is not in general optimal, but it gives a good approximation of the performance ("pretty good" is the judgment given by the authors and very often echoed in the literature).

The quantum least squares measurements (LSM) were subsequently developed by Eldar and Forney, in two articles [2, 3], deserving particular attention, because they formalize the whole problem in a very clear and general way, establishing a connection between the LSM and other types of measurements. In particular, they proved that the LSM technique produces the same results as the SRM technique, and precisely that the optimal measurement matrix (which minimizes the quadratic error) can be obtained both from Gram's operator and from Gram's matrix in the following way.

$$M_0 = T^{-\frac{1}{2}}\Gamma = \Gamma G^{-\frac{1}{2}}. \tag{6.1}$$

An important result is concerned with the SRM in the presence of GUS, which gives the **optimal decision for pure states**, allowing the exact evaluation of the error probability. Recently [4, 5], the SRM technique has been systematically applied to

G. Cariolaro, *Quantum Communications*, Signals and Communication Technology,
DOI 10.1007/978-3-319-15600-2_6

the performance evaluation of most popular quantum communications systems. From a computational viewpoint, the SRM can be improved with the technique of quantum compression [6], which has been introduced at the end of the previous chapter.

Organization of the Chapter

The SRM technique is mainly based on the SVD of the state matrix Γ and on the EID of the Gram matrix G and of the Gram operator T, developed in Sect. 5.12 of the previous chapter. For this reason these decompositions are recalled before developing the SRM.

The SRM for pure states is developed in Sects. 6.2 and 6.3 and extended to mixed states in Sect. 6.4. In Sects. 6.5 and 6.6 the SRM is developed assuming that the state constellations have the GUS.

In Sect. 6.7 the SRM technique is combined with the compression technique, showing the advantage of working in a compressed space, mainly in the presence of GUS. Finally, in Sect. 6.8 the quantum Chernoff bound is introduced as a further technique of suboptimization in quantum detection.

Recall from the Previous Chapter

For convenience we reconsider the main matrices and the related decompositions seen in Sect. 5.12 of the previous chapter, which are useful to SRM.

- *State and measurement matrices*

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

mixed states $\underset{n\times H}{\Gamma} = [\gamma_0, \gamma_1, \ldots, \gamma_{K-1}]\,, \qquad \underset{n\times H}{M} = [\mu_0, \mu_1, \ldots, \mu_{K-1}].$

Relations

$$M = \Gamma\, A\,, \qquad M = C\,\Gamma. \tag{6.2}$$

In particular, the second relation gives

$$|\mu_i\rangle = C\,|\gamma_i\rangle \quad \text{or} \quad \mu_i = C\,\gamma_i\,. \tag{6.2a}$$

Singular value decomposition of Γ

$$\Gamma = U\,\Sigma\,V^* = \sum_{i=1}^{r} \sigma_i\,|u_i\rangle\langle v_i| = U_r\,\Sigma_r\,V_r^* \tag{6.3}$$

- *Gram's matrix and Gram's operator*

$$\text{pure states} \qquad \underset{K\times K}{G} = \Gamma^* \Gamma\,, \qquad \underset{n\times n}{T} = \Gamma\,\Gamma^* \tag{6.4a}$$

$$\text{mixed states} \qquad \underset{H\times H}{G} = \Gamma^* \Gamma\,, \qquad \underset{n\times n}{T} = \Gamma\,\Gamma^* \tag{6.4b}$$

Relations

$$\text{pure states} \qquad G_{ij} = \langle \gamma_i | \gamma_j \rangle\,, \qquad T = \sum_{i=0}^{K-1} |\gamma_i\rangle\langle\gamma_i| \tag{6.5a}$$

$$\text{mixed states} \qquad G_{ij} = \gamma_i^*\, \gamma_j\,, \qquad T = \sum_{i=0}^{K-1} \rho_i\,. \tag{6.5b}$$

Eigendecompositions

$$T = U\,\Lambda_T U^* = \sum_{i=1}^{r} \sigma_i^2\, |u_i\rangle\langle u_i| = U_r\,\Sigma_r^2 U_r^* \tag{6.6a}$$

$$G = V\,\Lambda_G V^* = \sum_{i=1}^{r} \sigma_i^2\, |v_i\rangle\langle v_i| = V_r\,\Sigma_r^2 V_r^*. \tag{6.6b}$$

6.2 Square Root Measurements (SRM)

6.2.1 Formulation

Considering the equivalence between LSM and SRM, we find it convenient to introduce the topic in the sense of LSM, but we will use the more consolidated acronym SRM.

In the case of pure states, the measurement vectors $|\mu_i\rangle$ are chosen with the criterion of making the differences between the states and the measurement vectors, $|e_i\rangle = |\gamma_i\rangle - |\mu_i\rangle$, as "small" as possible (Fig. 6.1), and more specifically we look

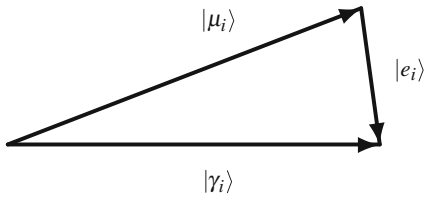

Fig. 6.1 In the LSM method the quadratic average of the "errors" $|e_i\rangle = |\gamma_i\rangle - |\mu_i\rangle$ is minimized

for **the measurement vectors $|\mu_i\rangle$ that minimize the quadratic error**

$$\mathcal{E} = \sum_{i=0}^{K-1} \langle e_i | e_i \rangle = \sum_{i=0}^{K-1} (\langle \gamma_i | - \langle \mu_i |)(|\gamma_i\rangle - \mu_i\rangle)$$

with the constraint of resolution of the identity

$$M M^* = \sum_{i=0}^{K-1} |\mu_i\rangle\langle\mu_i| = I_{\mathcal{H}} \quad \rightarrow \quad P_{\mathcal{U}} \tag{6.7}$$

where $I_{\mathcal{H}}$ can be replaced by $P_{\mathcal{U}}$ (see Proposition 5.5).

Introducing the difference between the state matrix and the measurement matrix: $E = [|e_0\rangle, |e_1\rangle, \ldots, |e_{K-1}\rangle] = \Gamma - M$, the quadratic error can be written in the form

$$\mathcal{E} = \mathrm{Tr}[E^* E] = \mathrm{Tr}[E E^*] \, . \tag{6.8}$$

We observe that if the vectors $|\gamma_i\rangle$ were orthonormal, the minimum of $\mathcal{E}$, satisfying the constraint (6.7), would be trivially $|\mu_i\rangle = |\gamma_i\rangle$, $1 \le i \le K$, which yields $\mathcal{E} = 0$.

The above can be extended to mixed states, for which the error is considered between the state factors and the measurement factors, $e_i = \gamma_i - \mu_i$, and the quadratic error is still given by (6.8). In any case we assume **equiprobable symbols**, that is, with equal a priori probabilities, $q_i = 1/K$, but the SRM method can be extended to generic a priori probabilities q_i substituting the states $|\gamma_i\rangle$ with the weighted states $\sqrt{q_i}|\gamma_i\rangle$ (see [2]).

As we will see, the SRM method always leads to explicit results and, in general, provides a good overestimation of the error probability.

6.2.2 Computation of the Optimal Measurement Matrix

Now we search for the optimal measurement matrix, $M = M_0$, that minimizes the quadratic error $\mathcal{E}$. Even though in quantum communications the states are always independent and so the rank of the state matrix is $r = K$, for greater generality (and for the interest that the general case will have with the extension of the method to mixed states), we suppose that Γ has a generic rank r. We obtain:

Theorem 6.1 *The measurement matrix M that minimizes the quadratic error $\mathcal{E}$ with the constraint (6.7), is given by*

$$M_0 = \sum_{i=1}^{r} |u_i\rangle\langle v_i| = U_r \, V_r^* \, , \tag{6.9}$$

that is, by the sum of the r transjectors $|u_i\rangle\langle v_i|$ seen in the (reduced) SVD of the state matrix Γ. The minimum quadratic error is

$$\mathcal{E}_{\min} = \sum_{i=0}^{K-1} (1-\sigma_i)^2 ,$$

where σ_i are the square roots of the eigenvalues (i.e., the singular values) of the Gram operator and of the Gram matrix (see (6.6)).

Note that here we indicate as *optimum* the measurement matrix giving the minimum square error (representing the "best" solution in this context). In general this does not provide the optimum decision, which minimizes the error probability.

Proof We follow the demonstration by [2] with some simplification. In the expression (6.8) of the quadratic error we take explicitly the trace with respect to the orthonormal basis $|u_i\rangle$, seen in the EID of Gram's operator in the previous chapter (see (5.109a)). In this way we find

$$\mathcal{E} = \mathrm{Tr}[E E^*] = \sum_{i=1}^{n} \langle u_i | E E^* | u_i \rangle = \sum_{i=1}^{n} \langle d_i | d_i \rangle \tag{6.10}$$

where

$$|d_i\rangle := E^* |u_i\rangle = (\Gamma - M)^* |u_i\rangle . \tag{6.10a}$$

Let us now consider the reduced SVD of Γ^* (see (6.3))

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

which gives $\Gamma^* |u_i\rangle = \sigma_i |v_i\rangle$. Now, letting

$$|a_i\rangle = M^* |u_i\rangle , \qquad i = 1, \ldots, r \tag{6.11}$$

(6.10a) becomes $|d_i\rangle = (\Gamma - M)^* |u_i\rangle = \sigma_i |v_i\rangle - |a_i\rangle$ and the ith component of the quadratic error results in

$$\begin{aligned} \mathcal{E}_i &= \langle d_i | d_i \rangle = \sigma_i^2 \langle v_i | v_i \rangle + \langle a_i | a_i \rangle - \sigma_i \langle v_i | a_i \rangle - \sigma_i \langle a_i | v_i \rangle \\ &= \sigma_i^2 + 1 - \sigma_i \langle v_i | a_i \rangle - \sigma_i \langle a_i | v_i \rangle . \end{aligned}$$

The minimum of $\mathcal{E}_i$ is reached when the quantity $\sigma_i \langle v_i | a_i \rangle + \sigma_i \langle a_i | v_i \rangle$ is maximum. Because of the constraint $|\langle v_i | a_i \rangle| \le 1$, this quantity is maximum when $|a_i\rangle = |v_i\rangle$, that is, when (6.11) becomes $|v_i\rangle = M^* |u_i\rangle$ for an appropriate $M = M_0$. From this relation, (6.9) follows.

6.2.3 Consequences of the Result

From the expression (6.9) of the optimal measurement matrix, it can be soon verified:

Corollary 6.1 *If the states $|\gamma_i\rangle$ are linearly independent, that is, if the rank of Γ is $r = K$, the optimal measurement vectors result orthonormal, $\langle\mu_i|\mu_j\rangle = \delta_{ij}$, and therefore the corresponding measurement operators $Q_i = |\mu_i\rangle\langle\mu_i|$ constitute a projector system.*

In fact, let us consider the (optimal) Gram's matrix of the measurement vectors

$$M_0^* M_0 = \sum_{i=0}^{K-1} |v_i\rangle\langle u_i| \sum_{j=0}^{K-1} |u_j\rangle\langle v_j| = \sum_{i=0}^{K-1} |v_i\rangle\langle v_i| = I_K$$

where (5.116) has been used. To conclude, it suffices to observe that *Gram's matrix of the measurement vectors M* has as elements the inner products $\langle\mu_i|\mu_i\rangle$ (see (5.107)).

In addition, we find what was anticipated by (6.1):

Corollary 6.2 *The optimal measurement matrix $M_0 = U_r V_r^*$ can be calculated also from the expressions*

$$M_0 = \Gamma(\Gamma^*\Gamma)^{-1/2} = \Gamma G^{-1/2} \tag{6.12a}$$

$$M_0 = (\Gamma\Gamma^*)^{-1/2}\Gamma = T^{-1/2}\Gamma \tag{6.12b}$$

where $G^{-1/2}$ and $T^{-1/2}$ are the inverse square roots of G and T that are obtained from the corresponding reduced EIDs in the following way

$$G^{-1/2} = V_r\, \Sigma_r^{-1}\, V_r^*\,, \qquad T^{-1/2} = U_r\, \Sigma_r^{-1}\, U_r^*\,. \tag{6.13}$$

For example, the proof of (6.12a) is carried out using the expression $G^{-1/2}$ introduced above, and the reduced SVD of Γ. We obtain

$$\Gamma G^{-1/2} = U_r \Sigma_r V_r^* V_r \Sigma_r^{-1} V^* = U_r V_r^* = M_0$$

where we took into account that $V_r^* V_r = I_r$ and that $\Sigma\Sigma_r^{-1} = I_r$.

On the inverse square roots. In (6.13) we have formally introduced the inverse square roots $G^{-1/2}$ and $T^{-1/2}$ of G and of T. To obtain these roots we start from the corresponding reduced EIDs and we operate on the common diagonal matrix $\Sigma_r^2 = \mathrm{diag}[\sigma_1^2, \ldots, \sigma_r^2]$, taking its inverse square root $\Sigma_r^{-1} = \mathrm{diag}[1/\sigma_1, \ldots, 1/\sigma_r]$, where the σ_i^2 are all positive, and therefore there are no indeterminacy problems. In general, $G^{-1/2}$ and $T^{-1/2}$ should be intended as *pseudoinverses* (according to Moore–Penrose formula [7, 8]). The Moore–Penrose pseudoinverse is based on the full EID (not reduced) $U\Lambda U^*$, by taking the reciprocal of each nonzero element on

the diagonal, leaving the zeros in place. Here we prefer to use the **reduced EID**, where the diagonal matrix is regular and has no problem for its inversion. In any case note that the inversion can bring about some surprising result. For example, it can be verified that the relation $T^{-1/2}T^{1/2} = U_r U_r^*$ does not yield, in general, the identity $I_{\mathcal{H}}$, but it gives the projector $P_{\mathcal{U}} = U_r U_r^*$ and only if $r = K$ one actually produces the identity $I_{\mathcal{H}}$.

The path followed so far to introduce the inverse square roots, based on the *reduced* EIDs, is slightly unusual; in fact, the EIDs are normally considered full, and this entails the complication of having to introduce several diagonal matrices [2].

Problem 6.1 Prove that $T^{-1/2}T^{1/2}$ does not yield, in general, the identity $I_{\mathcal{H}}$, but the projector $P_{\mathcal{U}} = U_r U_r^*$. Only if $r = K$ one actually produces the identity $I_{\mathcal{H}}$.

Problem 6.2 ⋆⋆ Consider the following state matrix of $\mathcal{H} = \mathbb{C}^4$

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

Find the inverse square root $G^{-1/2}$ and $T^{-1/2}$ based on the two approaches: (1) the Moore–Penrose pseudoinverse and (2) the reduced EID.

6.3 Performance Evaluation with the SRM Decision

With the SRM method, we have seen that the optimal measurement matrix has three distinct expressions

$$M_0 = U_r\, V_r^* = T^{-1/2}\Gamma = \Gamma G^{-1/2}\,. \tag{6.14}$$

The first expression is bound to the reduced SVD of the measurement matrix Γ, while the other two are obtained from the reduced EIDs of T and G, respectively.

From the measurement matrix, which collects the measurement vectors $|\mu_i\rangle$, the measurement operators can be computed as $Q_i = |\mu_i\rangle\langle\mu_i|$ and from these the performance of the quantum system. The transition probabilities result in

$$p_c(j\,|\,i) = \mathrm{Tr}[\rho_i\, Q_j] = |\langle\mu_j|\gamma_i\rangle|^2 \tag{6.15}$$

and the correct decision probability (with equiprobable symbols)

$$P_c = \frac{1}{K}\sum_{i=0}^{K-1} |\langle\mu_i|\gamma_i\rangle|^2. \tag{6.16}$$

We now discuss the three possible methods, expressing them in a form suitable for computation.

6.3.1 Method Based on the SVD of the State Matrix

The optimal measurement matrix, evaluated according to the expression

$$M_0 = U_r \, V_r^* \,, \tag{6.17}$$

can be obtained directly from the reduced SVD of the state matrix, which has the form (see (5.110)): $\Gamma = U_r \, \Sigma_r \, V_r^*$. Therefore, in this expression it suffices to suppress the diagonal matrix to obtain the optimal measurement matrix. This is the most direct method as it does not require to calculate the inverse root square of a matrix.

6.3.2 Method Based on Gram's Operator

Let us start from *Gram's operator*,

$$T = \Gamma \, \Gamma^* = \sum_{i=0}^{K-1} |\gamma_i\rangle\langle\gamma_i| \,, \tag{6.18}$$

which is a positive semidefinite Hermitian operator (see Sect. 2.10.4). Then it is possible to define its square root, using the EID (5.109a), which is, in the reduced form, $T = U_r \, \Sigma_r^2 \, U_r^*$ and gives

$$T^{-\frac{1}{2}} = U_r \, \Sigma_r^{-1} \, U_r^* \tag{6.19}$$

from which we obtain the optimal measurement matrix as

$$M_0 = T^{-\frac{1}{2}} \Gamma. \tag{6.20}$$

At this point we observe that (6.20) falls into the form (5.103), that is, $M = C \, \Gamma$, and then the measurement vectors are simply obtained according to

$$\boxed{|\mu_i\rangle = T^{-\frac{1}{2}} |\gamma_i\rangle} \tag{6.21}$$

from which we get the elementary measurement operators as

$$Q_i = |\mu_i\rangle\langle\mu_i| = T^{-\frac{1}{2}}|\gamma_i\rangle\langle\gamma_i|T^{-\frac{1}{2}}. \tag{6.22}$$

6.3.3 Method Based on Gram's Matrix

We start from *Gram's matrix*

$$\underset{K\times K}{G} = \Gamma^*\,\Gamma = \begin{bmatrix} \langle\gamma_0|\gamma_0\rangle & \cdots & \langle\gamma_0|\gamma_{K-1}\rangle \\ \vdots & \ddots & \vdots \\ \langle\gamma_{K-1}|\gamma_0\rangle & \cdots & \langle\gamma_{K-1}|\gamma_{K-1}\rangle \end{bmatrix} \tag{6.23}$$

which is obtained by computing the inner products $\langle\gamma_i|\gamma_j\rangle$. We then evaluate the reduced EID, which has the form

$$G = V_r\,\Sigma_r^2\,V_r^* \tag{6.24}$$

where the diagonal matrix Σ_r^2 is the same as the one appearing in the previous EID. From this we compute the inverse square root

$$G^{-\frac{1}{2}} = V_r\,\Sigma_r^{-1}\,V_r^* \tag{6.25}$$

and then we obtain the optimal measurement matrix as

$$M_0 = \Gamma\,G^{-\frac{1}{2}}\,. \tag{6.26}$$

This form is of the type (5.102), that is, $M = \Gamma\,A$, which expresses in a compact form the fact that the (optimal) measurement vectors are given by a linear combination of the states, that is,

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

Now, as $A = G^{-1/2}$ and therefore $a_{ij} = (G^{-1/2})_{ij}$, the measurement vectors result explicitly in

$$|\mu_i\rangle = \sum_{j=0}^{K-1} (G^{-1/2})_{ij}|\gamma_j\rangle\,. \tag{6.27}$$

The transition probabilities are computed from the *mixed inner products* $b_{ij} = \langle\mu_i|\gamma_j\rangle$, which define the $K \times K$ matrix $B = M^*\Gamma$ (see (5.74)). Now, from (6.24) and (6.25) we have

$$B := M^* \Gamma = G^{-1/2} \Gamma^* \Gamma = G^{-1/2} G = G^{1/2} \, . \tag{6.28}$$

Then the matrix of the mixed inner products becomes simply $G^{1/2}$ and since $p_c(j|i) = |b_{ij}|^2$ we have

$$\boxed{p_c(j|\, i) = \left| (G^{\frac{1}{2}})_{ij} \right|^2} \tag{6.29}$$

from which we obtain the correct decision probability with equiprobable symbols

$$\boxed{P_c = \frac{1}{K} \sum_{i=0}^{K-1} \left| (G^{\frac{1}{2}})_{ii} \right|^2 .} \tag{6.30}$$

Example 6.1 Consider a binary system ($K = 2$) on $\mathcal{H} = \mathbb{C}^4$, in which the two states are specified by the matrix

$$\Gamma = \left[|\gamma_1\rangle, |\gamma_2\rangle \right] = \frac{1}{2\sqrt{13}} \begin{bmatrix} 5 & 1 \\ 3 - 2\,\mathrm{i} & 3 + 2\,\mathrm{i} \\ 1 & 5 \\ 3 + 2\,\mathrm{i} & 3 - 2\,\mathrm{i} \end{bmatrix} \tag{6.31}$$

which has rank $r = K = 2$. The reduced SVD of Γ becomes: $\Gamma = U_r \Sigma V_r^*$, where

$$U_r = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -\mathrm{i} \\ 1 & -1 \\ 1 & \mathrm{i} \end{bmatrix} , \qquad \Sigma_r = \begin{bmatrix} \sqrt{\frac{18}{13}} & 0 \\ 0 & \sqrt{\frac{8}{13}} \end{bmatrix} , \qquad V = V_r = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} .$$

Then, from (6.17) we get the optimal measurement matrix

$$M_0 = U_r V^* = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -\mathrm{i} \\ 1 & -1 \\ 1 & \mathrm{i} \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{2\sqrt{2}} \begin{bmatrix} 2 & 0 \\ 1 - \mathrm{i} & 1 + \mathrm{i} \\ 0 & 2 \\ 1 + \mathrm{i} & 1 - \mathrm{i} \end{bmatrix} . \tag{6.32}$$

The measurement vectors become then

$$|\mu_1\rangle = \frac{1}{2\sqrt{2}} \begin{bmatrix} 2 \\ 1 - \mathrm{i} \\ 0 \\ 1 + \mathrm{i} \end{bmatrix} , \qquad |\mu_2\rangle = \frac{1}{2\sqrt{2}} \begin{bmatrix} 0 \\ 1 + \mathrm{i} \\ 2 \\ 1 - \mathrm{i} \end{bmatrix}$$

and their orthonormality can be verified, in particular $\langle \mu_1 | \mu_2 \rangle = 0$, in agreement with Corollary 6.1. We can now compute the transition probabilities from (6.15), that is, $p_c(j\,|\,i) = |\langle \mu_j | \gamma_i \rangle|^2$. We obtain the matrix

$$p_c = \begin{bmatrix} \frac{25}{26} & \frac{1}{26} \\ \frac{1}{26} & \frac{25}{26} \end{bmatrix} \tag{6.33}$$

from which we have that the error probability with equiprobable symbols results in $P_e = 1 - P_c = \frac{1}{26}$.

We leave it to the reader to verify that the other two performance evaluation methods, based on the reduced EIDs of G and of T, lead to the same results found with the SVD of Γ.

Problem 6.3 ⋆⋆ Consider the state matrix Γ given by (6.31) of Example 6.1. Check that the methods based on the EIDs of G and T give the same transition probabilities as obtained with the SVD of Γ.

Problem 6.4 ⋆⋆ With the data of the previous problem, find the relations

$$\mu_1 = C\,\gamma_1\,, \qquad \mu_2 = C\,\gamma_2.$$

These relations are somewhat intriguing since they lead to think that μ_1 depends only on γ_1 and not on γ_2 and μ_2 only on γ_2. Explain why not.

6.3.4 Properties of the SRM

We have seen that the operators of the SRM can always be calculated in a rather simple manner for any constellation of states $|\gamma_i\rangle$ and therefore for any quantum communications system. This is already a first advantage. It remains to understand whether the SRM are optimal or close to optimal. It has been proved by Holevo in 1979 [9] that the SRM are *asymptotically optimal*, in the sense that they become optimal in practice when the average number of photons is large enough. Furthermore, these measurements become optimal when the constellation of the states enjoys the geometrically uniform symmetry (GUS) (see below).

Another advantage of the SRM regards their practical implementation. In fact, receivers based on the SRM have already been implemented (in 1999), using QED cavities [10].

6.4 SRM with Mixed States

The SRM method is normally used for decision in the presence of pure states, but recently [11] this method was extended to systems described by density operators (mixed states), allowing for the evaluation of the performance of quantum communications systems even *in the presence of thermal noise*.

The technique behind this generalization, consisting in passing from pure states to density operators, is the usual factorization of the density operators

$$\rho_i = \gamma_i \gamma_i^*$$

which allows us to proceed in basically the same way as seen with pure states, using the following correspondence

state $\|\gamma_i\rangle$	$\rightarrow$	state factor γ_i
measurement vector $\|\mu_i\rangle$	$\rightarrow$	measurement factor μ_i

whose consequences are summarized in Table 6.1. As done with pure states, we keep the hypothesis of equiprobable symbols.

Table 6.1 The SRM method with pure states and with mixed states

Operation	Pure states	Mixed states
Density operators		$\rho_0, \ldots, \rho_{K-1}$
States/factor states	$\|\gamma_0\rangle, \ldots, \|\gamma_{K-1}\rangle$	$\gamma_0, \ldots, \gamma_{K-1}$
State matrix	$\Gamma = [\|\gamma_0\rangle, \ldots, \|\gamma_{K-1}\rangle]$	$\Gamma = [\gamma_0, \ldots, \gamma_{K-1}]$
Gram's matrix	$G = \Gamma^* \Gamma = \left[\langle\gamma_i\|\gamma_j\rangle\right]$	$G = \Gamma^* \Gamma = \left[\gamma_i^* \gamma_j\right]$
Gram's operator	$T = \Gamma \Gamma^* = \sum_{i=0}^{K-1} \|\gamma_i\rangle\langle\gamma_i\|$	$T = \Gamma \Gamma^* = \sum_{i=0}^{K-1} \gamma_i \gamma_i^*$
Measurement vectors/factors	$\|\mu_i\rangle = T^{-\frac{1}{2}} \|\gamma_i\rangle$	$\mu_i = T^{-\frac{1}{2}} \gamma_i$
Measurement matrix	$M = T^{-\frac{1}{2}} \Gamma = \Gamma G^{-\frac{1}{2}}$	$M = T^{-\frac{1}{2}} \Gamma = \Gamma G^{-\frac{1}{2}}$
Mixed product matrix	$B = M^* \Gamma = \left[\langle\mu_i\|\gamma_j\rangle\right] = G^{1/2}$	$B = M^* \Gamma = \left[\mu_i^* \gamma_j\right] = G^{1/2}$
Measurement operators	$Q_i = \|\mu_i\rangle\langle\mu_i\|$	$Q_i = \mu_i \mu_i^*$
Transition probabilities $p(j\|i)$	$\|\langle\mu_j\|\gamma_i\rangle\|^2 = \|b_{ji}\|^2$	$\mathrm{Tr}[\mu_j \mu_j^* \gamma_i \gamma_i^*] = \mathrm{Tr}[b_{ji}^* b_{ji}]$
Correct decision probability P_c	$\frac{1}{K} \sum_{i=0}^{K-1} \|b_{ii}\|^2$	$\frac{1}{K} \sum_{i=0}^{K-1} \mathrm{Tr}[b_{ii}^* b_{ii}]$

6.4.1 Discussion of the Method

Having obtained the factors γ_i of the density operators ρ_i, we form the state matrix

$$\underset{n\times H}{\Gamma} = \left[\gamma_0, \gamma_1, \ldots, \gamma_{K-1}\right] \tag{6.34}$$

where H is the total number of columns, and from this we obtain Gram's operator (see (5.118)) and Gram's matrix (see (5.119))

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

Theorem 6.1 and the subsequent corollaries still hold, so the optimal measurement matrix can be calculated from three distinct expressions

$$M_0 = U_r\, V_r^* = T^{-1/2}\Gamma = \Gamma G^{-1/2}. \tag{6.35}$$

The first expression is bound to the reduced SVD of the measurement matrix Γ, while the other two are obtained from the reduced EIDs of T and G.

From $M_0 = \left[\mu_0, \mu_1, \ldots, \mu_{K-1}\right]$ we get the measurement factors μ_i and, from these, the measurement operators $Q_i = \mu_i \mu_i^*$. The relation giving the mixed product matrix $B = [b_{ij}] = [\mu_i^*\, \gamma_j]$ still holds

$$B = M^*\,\Gamma = G^{1/2}. \tag{6.36}$$

Finally, we obtain the transition probabilities and the correct decision probability from (5.75)

$$p_c(j|i) = \mathrm{Tr}[b_{ji}^*\, b_{ji}], \qquad P_c = \sum_{i\in\mathcal{A}} q_i \mathrm{Tr}[b_{ii}^*\, b_{ii}]. \tag{6.37}$$

where now b_{ij} is the i, j block of the matrix $G^{1/2}$.

Example 6.2 Let us consider Problem 5.9 of the previous chapter, where starting from two density operators ρ_0 and ρ_1 of $\mathcal{H} = \mathbb{C}^4$, we found the factors γ_0 of dimensions 4×2 and γ_1 of dimensions 4×3. From these factors, the 4×5 state matrix is formed

$$\Gamma = \left[\gamma_0, \gamma_1\right] = \begin{bmatrix} -0.54117 & -0.02018 & -0.47937 & -0.06934 & 0.03124 \\ -0.54117 & -0.02018 & \mathrm{i}\,0.51339 & 0.0 & \mathrm{i}\,0.02917 \\ -0.54117 & -0.02018 & 0.479370 & -0.06934 & -0.03124 \\ -0.33238 & 0.09857 & -\mathrm{i}\,0.51339 & 0.0 & -\mathrm{i}\,0.02917 \end{bmatrix}.$$

From Γ we obtain the 4×4 Gram's operator

$$T = \Gamma\ \Gamma^* = \begin{bmatrix} 0.52885 & 0.29327 + \mathrm{i}\,0.24519 & 0.06731 & 0.17788 - \mathrm{i}\,0.24519 \\ 0.29327 - \mathrm{i}\,0.24519 & 0.55769 & 0.29327 + \mathrm{i}\,0.24519 & -0.08654 \\ 0.06731 & 0.29327 - \mathrm{i}\,0.24519 & 0.52885 & 0.17788 + \mathrm{i}\,0.24519 \\ 0.17788 + \mathrm{i}\,0.24519 & -0.08654 & 0.17788 - \mathrm{i}\,0.24519 & 0.38462 \end{bmatrix}$$

and the 5×5 Gram's matrix

$$G = \Gamma^*\ \Gamma = \begin{bmatrix} 0.98906 & 0.0 & -\mathrm{i}\,0.10719 & 0.07505 & -\mathrm{i}\,0.00609 \\ 0.0 & 0.01094 & -\mathrm{i}\,0.06096 & 0.00280 & -\mathrm{i}\,0.00346 \\ \mathrm{i}\,0.10719 & \mathrm{i}\,0.06096 & 0.98673 & 0.0 & \\ 0.07505 & 0.00280 & & 0.00962 & \\ \mathrm{i}\,0.00609 & \mathrm{i}\,0.00346 & 0.0 & 0.0 & 0.00365 \end{bmatrix}.$$

The four eigenvalues of T are all positive and precisely

$$1.04854 \quad 0.941288 \quad 0.101754 \quad 0.0647906$$

and it can be verified that G has the same positive eigenvalues (the fifth eigenvalue of G is null).

As T has full rank, its inverse square root must be intended in the ordinary sense, and results in

$$T^{-1/2} = \begin{bmatrix} 6.55880 & -3.71728 - \mathrm{i}\,3.44423 & -0.73775 + \mathrm{i}\,1.90902 & -1.85711 + \mathrm{i}\,2.39197 \\ -3.71728 + \mathrm{i}\,3.44423 & 8.04954 & -3.71728 - \mathrm{i}\,3.44423 & 0.66455 \\ -0.73775 - \mathrm{i}\,1.90902 & -3.71728 + \mathrm{i}\,3.44423 & 6.55880 & -1.85711 - \mathrm{i}\,2.39197 \\ -1.85711 - \mathrm{i}\,2.39197 & 0.66455 & -1.85711 + \mathrm{i}\,2.39197 & 6.11094 \end{bmatrix}.$$

From $T^{-1/2}$ we obtain the measurement factors

$$\mu_0 = T^{-1/2}\gamma_0 = \begin{bmatrix} 0.48977 + \mathrm{i}\,0.01653 & -0.25565 - \mathrm{i}\,0.26963 \\ 0.54077 & -0.00921 \\ 0.48977 - \mathrm{i}\,0.01653 & -0.25565 + \mathrm{i}\,0.26963 \\ 0.47493 & 0.60823 \end{bmatrix}$$

$$\mu_1 = T^{-1/2}\gamma_1 = \begin{bmatrix} -0.02278 - \mathrm{i}\,0.49923 & -0.40600 + \mathrm{i}\,0.18138 & 0.09726 - \mathrm{i}\,0.40602 \\ 0.50687 & 0.48486 - \mathrm{i}\,0.07894 & -0.45743 \\ -0.02278 + \mathrm{i}\,0.49923 & -0.44257 - \mathrm{i}\,0.04322 & 0.09726 + \mathrm{i}\,0.40602 \\ -0.49203 & 0.26908 - \mathrm{i}\,0.04381 & 0.29679 \end{bmatrix}.$$

Finally, from (6.37) we obtain the transition probabilities $p_c(j|i)$, whose matrix is

$$p_c = \begin{bmatrix} 0.986242 & 0.013758 \\ 0.01084 & 0.013758 \end{bmatrix}$$

hence, with equiprobable symbols, we have

$$P_c = 0.986242 \qquad P_e = 0.013758.$$

Computation of P_e from the eigenvalues (Helstrom). Having considered a binary case, we know how to compute the optimal projectors according to Helstrom's theory, which is based on the eigenvalues of the decision operator $D = \frac{1}{2}(\rho_1 - \rho_0)$. The eigenvalues of D become

$$\{-0.977483, 0.977145, 0.00422318, -0.00388438\}$$

So, applying (5.22), we obtain

$$P_c = 0.981368 \qquad P_e = 0.018632$$

and we realize that the SRM gives an *underestimate* of the error probability.

6.5 SRM with Geometrically Uniform States (GUS)

The geometrically uniform symmetry (GUS) has been introduced in Sect. 5.13. Now, it is evident that the GUS on a constellation of states leads to a symmetry also on the measurement vectors, with remarkable simplifications, but the most important consequence is that, with pure states, the SRM method in the presence of GUS provides the *optimal* decision (maximizing the correct decision probability), as will be seen toward the end of this section.

6.5.1 Symmetry of Measurement Operators Obtained with the GUS

In Proposition 5.9 we have seen that if the state constellation has the GUS, also the *optimal* measurement operators have the same symmetry. We now prove that this property also holds for the measurement operators obtained with the SRM, which are not optimal in general. [1]

[1] It is useful to recall that we call *optimal* the measurement operators obtained with the maximization of the correct decision probability, while the measurement operators obtained with the SRM minimize the quadratic error between the measurement vectors and the state vectors.

In Proposition 5.8 we have seen that the Gram operator T and the symmetry operator S commute and therefore they are simultaneously diagonalizable

$$T\,S = S\,T \quad \Longleftrightarrow \quad T = U\,\Lambda_T\,U^*\,, \qquad S = U\,\Lambda_S U^*. \tag{6.38a}$$

Then, also the powers of T and of S are simultaneously diagonalizable and therefore commute

$$T^\alpha = U\,\Lambda_T^\alpha U^*\,, \qquad S^\beta = U\,\Lambda_S^\beta U^* \quad \Longleftrightarrow \quad T^\alpha\,S^\beta = S^\beta\,T^\alpha. \tag{6.38b}$$

In particular $T^{-\frac{1}{2}}$ commutes with S^i for every i

$$T^{-\frac{1}{2}}\,S^i = S^i\,T^{-\frac{1}{2}}\,, \qquad i = 0, 1, \ldots, K-1\,. \tag{6.39}$$

Then, combining (6.21) with (6.39) we obtain

$$|\mu_i\rangle = T^{-\frac{1}{2}}|\gamma_i\rangle = T^{-\frac{1}{2}}S^i|\gamma_0\rangle = S^i T^{-\frac{1}{2}}|\gamma_0\rangle\,.$$

The above result can be formulated as follows:

Theorem 6.2 *If a constellation of states $|\gamma_i\rangle$ has the GUS with symmetry operator S, also the measurement vectors obtained with the SRM have the GUS with the same symmetry operator, namely*

$$|\mu_i\rangle = S^i|\mu_0\rangle\,, \qquad i = 0, 1, \ldots, K-1 \tag{6.40}$$

where

$$|\mu_0\rangle = T^{-\frac{1}{2}}|\gamma_0\rangle\,. \tag{6.40a}$$

Thus, from (6.40), all the measurement vectors can be obtained from the reference vector $|\mu_0\rangle$. This property is then transferred to the measurement operators Q_i with the usual rules.

6.5.2 Consequences of the GUS on Gram's Matrix

When the states have the GUS, Gram's matrix becomes circulant and the SRM methodology can be developed to arrive at explicit results.

We recall that a matrix $G = [G_{ij}]$ of dimensions $K \times K$ is called *circulant* if its elements depend only on the difference of the indexes, modulo K, that is, they are of the type

$$G_{ij} = r_{i-j\ (\mathrm{mod}\,K)}. \tag{6.41}$$

For example for $K = 4$ we have the structure

$$G = \begin{bmatrix} r_0 & r_1 & r_2 & r_3 \\ r_3 & r_0 & r_1 & r_2 \\ r_2 & r_3 & r_0 & r_1 \\ r_1 & r_2 & r_3 & r_0 \end{bmatrix}$$

and the elements of the rows are obtained as permutations of those of the first row. Therefore, a circulant matrix is completely specified by its first row, which for convenience we will call *circulant vector*.

Now, from (5.120) it results that the inner products

$$\begin{aligned} G_{ij} &= \langle \gamma_i | \gamma_j \rangle = \langle \gamma_0 | (S^*)^i S^j | \gamma_0 \rangle \\ &= \langle \gamma_0 | S^{j-i} | \gamma_0 \rangle = r_{i-j \pmod K} \end{aligned}$$

depend upon the difference $i - j \pmod K$, so ensuring that Gram's matrix is (Hermitian) circulant with circulant vector

$$[r_0, r_1, \ldots, r_{K-1}] = [1, \langle \gamma_0 | S | \gamma_0 \rangle, \ldots, \langle \gamma_0 | S^{K-1} | \gamma_0 \rangle].$$

The EID of a circulant Gram's matrix is expressed through the matrix of the DFT (Discrete Fourier Transform), given by

$$W_{[K]} = \frac{1}{\sqrt{K}} \begin{bmatrix} 1 & 1^0 & 1^{-1} & \ldots & 1^{-(K-1)} \\ 1 & W_K^{-1} & W_K^{-2} & \ldots & W_K^{-2(K-1)} \\ \vdots & & \vdots & & \\ 1 & W_K^{-(K-1)} & W_K^{-2(K-1)} & \ldots & W_K^{-(K-1)(K-1)} \end{bmatrix} \tag{6.42}$$

where $W_K = \mathrm{e}^{\mathrm{i}2\pi/K}$. From the orthonormality condition

$$\sum_{s=0}^{K-1} \frac{1}{K} W_K^{rs} = \delta_{r0} \tag{6.43}$$

it can be verified that the columns of $W_{[K]}$

$$|w_p\rangle = \frac{1}{\sqrt{K}} \left[W_K^{-p}, W_K^{-2p}, \ldots, W_K^{-p(K-1)} \right]^{\mathrm{T}}, \qquad p = 0, 1, \ldots, K-1 \tag{6.44}$$

form an orthonormal basis of $\mathbb{C}^K$, i.e., $\langle w_p | w_q \rangle = \delta_{pq}$.

Theorem 6.3 *A circulant Gram's matrix* $G = [G_{ij}] = [r_{i-j \,(mod\, K)}]$ *has the following EID*

$$G = W^* \Lambda W = \sum_{p=0}^{K-1} \lambda_p |w_p\rangle\langle w_p| \quad \text{with} \quad W := W_{[K]} \tag{6.45}$$

where the eigenvalues are given by the DFT of the circulant vector

$$\lambda_p = \sum_{q=0}^{K-1} G_{0q} W_K^{-pq} = \sum_{q=0}^{K-1} r_q W_K^{-pq} \tag{6.45a}$$

and $\Lambda = \text{diag}\,[\lambda_0, \lambda_1, \ldots, \lambda_{K-1}]$.

The theorem is proved in Appendix section "On the EID of a Circulant Matrix".

6.5.3 Performance Evaluation

From Theorem 6.3 we soon find the square roots

$$G^{\pm\frac{1}{2}} = \sum_{p=0}^{K-1} \lambda_p^{\pm\frac{1}{2}} |w_p\rangle\langle w_p| = W^* \Lambda^{\pm\frac{1}{2}} W \tag{6.46}$$

whose elements are given by

$$(G^{\pm\frac{1}{2}})_{ij} = \frac{1}{K} \sum_{p=0}^{K-1} \lambda_p^{\pm\frac{1}{2}} W_K^{-p(i-j)}. \tag{6.46a}$$

We can then evaluate the transition probabilities from (6.29), where the element ij is computed from (6.46a); thus

$$p_c(j|i) = \left| \frac{1}{K} \sum_{p=0}^{K-1} \lambda_p^{\frac{1}{2}} W_K^{-p(i-j)} \right|^2, \qquad i, j = 0, 1, \ldots, K-1 \tag{6.47}$$

in particular, the diagonal transition probabilities are found to be all equal

$$p_c(i|i) = \left[\frac{1}{K} \sum_{p=0}^{K-1} \lambda_p^{\frac{1}{2}} \right]^2 \qquad \text{(independent of } i\text{)} \tag{6.47a}$$

and therefore the correct decision probability (6.30) becomes explicitly

$$P_c = \left[\frac{1}{K} \sum_{p=0}^{K-1} \lambda_p^{\frac{1}{2}} \right]^2. \tag{6.48}$$

The measurement vectors $|\mu_i\rangle$ are obtained as linear combination of the states (see (6.27)), but considering that they have the GUS, their evaluation can be limited to the reference vector, given by

$$|\mu_0\rangle = \sum_{j=0}^{K-1} (G^{-1/2})_{ij} |\gamma_j\rangle. \tag{6.49}$$

In conclusion, when Gram's matrix G is circulant, to evaluate the measurement vectors and their performance, it suffices to compute their eigenvalues, given simply by the DFT of the first row of G. This methodology will be applied to PSK and PPM modulations in the next chapter.

6.5.4 *Optimality of SRM Decision with Pure States Having the GUS*

We now prove that the SRM decision, when the states have the GUS, realizes the minimum error probability.

Proposition 6.1 *When the constellation of* **pure states** *verifies the GUS, the SRM becomes* **optimum**, *achieving the minimum error probability.*

For the poof we use Holevo's theorem, in the version given by Corollary 5.1. With equiprobable symbols, the first conditions of Holevo's theorem result from (5.97)

$$b_{ij}\, b_{jj}^* - b_{ii}\, b_{ji}^* = 0\,, \qquad \forall i, \forall j \tag{6.50}$$

where $b_{ij} = \langle \mu_i | \gamma_j \rangle$ are the mixed products. Their matrix is given by (see (6.28))

$$B = G^{1/2} = W^* \Lambda^{1/2} W$$

and it is symmetric. Its elements b_{ij} depend only upon the difference $i - j$, as indicated also by (6.46a), and then they can be expressed in the form

$$b_{ij} = f(j-i) \quad \text{with} \quad b_{ij}^* = f(i-j)\,.$$

Therefore, from (6.50) it follows $f(j-i) f^*(0) - f(0) f(j-i) = 0$, which is verified because $f(0)$ is real.

For a proof of the second condition, we address the reader to [2].

6.5.5 Helstrom's Bound with SRM

In Sect. 5.13 we have seen that a binary constellation always satisfies the GUS, and hence the SRM gives the optimal decision. Then with the SRM approach we have to obtain the Helstrom bound.

In the binary case the state matrix is $\Gamma = [|\gamma_0\rangle, |\gamma_1\rangle]$ and Gram's matrix is

$$G = \Gamma^* \Gamma = \begin{bmatrix} 1 & X \\ X^* & 1 \end{bmatrix}, \qquad X := \langle \gamma_0 | \gamma_1 \rangle.$$

and in general is not circulant because $X^* \neq X$ and hence we cannot apply the approach based on the DFT.

We evaluate the square roots of G by hand. Assuming that $G^{1/2}$ has the form[2]

$$G^{1/2} = \begin{bmatrix} a & b \\ b^* & a \end{bmatrix}$$

we find the conditions

$$a^2 + |b|^2 = 1\,, \qquad 2ab = X$$

which give

$$a^2 + \frac{|X|^2}{4a^2} = 1 \quad \rightarrow \quad a^4 - \frac{1}{4}|X|^2 = 0.$$

The solution is

$$a = \frac{1}{\sqrt{2}} \sqrt{1 + \sqrt{1 - |X|^2}}$$

and

$$b = \frac{X}{2a} = \frac{X}{\sqrt{2}|X|} \sqrt{1 - \sqrt{1 - |X|^2}} = \frac{\mathrm{e}^{\mathrm{i}\beta}}{\sqrt{2}} \sqrt{1 - \sqrt{1 - |X|^2}}$$

where $\beta = \arg X$. As a check

$$G^{1/2} G^{1/2} = \begin{bmatrix} 1 & X \\ X^* & 1 \end{bmatrix} = G.$$

From $G^{1/2}$we have the correct decision probability from (6.30) as

$$P_c = \frac{1}{2} \left[|(G^{1/2})_{00}|^2 + |(G^{1/2})_{00}|^2 \right] = \frac{1}{2} \left[1 + \sqrt{1 - |X|^2} \right]$$

that is, the Helstrom bound.

[2] The assumption that the diagonal elements are equal is in agreement with a Sasaki's et al. [12] theorem, which states that in a optimal decision the square root of the Gram matrix must have all the diagonal elements equal.

Next we evaluate the optimal measurement vectors. Considering that $\det G^{1/2} = \sqrt{1-|X|^2}$, the inverse of $G^{1/2}$ is

$$G^{-1/2} = \frac{1}{\sqrt{1-|X|^2}} \begin{bmatrix} a & -b \\ -b^* & a \end{bmatrix}.$$

Using the identities

$$\frac{1}{1-|X|} \pm \frac{1}{1+|X|} = \sqrt{2}\frac{\sqrt{1 \pm \sqrt{1-|X|^2}}}{\sqrt{1-|X|^2}}$$

we find (with $\beta = \arg X$)

$$G^{-1/2} = \frac{1}{2} \begin{bmatrix} \frac{1}{1-|X|} + \frac{1}{1+|X|} & \mathrm{e}^{\mathrm{i}\beta}(\frac{1}{1-|X|} - \frac{1}{1+|X|}) \\ \mathrm{e}^{-\mathrm{i}\beta}(\frac{1}{1-|X|} - \frac{1}{1+|X|}) & \frac{1}{1-|X|} + \frac{1}{1+|X|} \end{bmatrix}$$

which gives the measurement matrix as $M = \Gamma G^{-1/2}$.

When the inner product is real the Gram matrix turns out to be circulant and the approach based on the DFT can be applied to get the Helstrom bound (see Problem 6.5).

Table 6.2 The SRM method in general, and with geometrically uniform symmetry (GUS)

Operation	General case	With GUS
Constellation of states	$\|\gamma_0\rangle, \ldots, \|\gamma_{K-1}\rangle$	$\|\gamma_i\rangle = S^i\|\gamma_0\rangle$
State matrix Γ	$[\|\gamma_0\rangle, \ldots, \|\gamma_{K-1}\rangle]$	$[\|\gamma_0\rangle, \ldots, S^{K-1}\|\gamma_0\rangle]$
Gram's matrix $G = \Gamma^* \Gamma$	$\left[\langle\gamma_i\|\gamma_j\rangle\right]$	$\left[\langle\gamma_0\|S^{j-i}\|\gamma_0\rangle\right] = W^* \Lambda W$
Gram's operator $T = \Gamma \Gamma^*$	$\sum_{i=0}^{K-1} \|\gamma_i\rangle\langle\gamma_j\|$	$\sum_{i=0}^{K-1} S^i\|\gamma_0\rangle\langle\gamma_0\|S^{-i}$
Measurement vectors	$\|\mu_i\rangle = T^{-\frac{1}{2}}\|\gamma_i\rangle$	$\|\mu_0\rangle = T^{-\frac{1}{2}}\|\gamma_0\rangle$, $\|\mu_i\rangle = S^i\|\mu_0\rangle$
Measurement matrix M	$T^{-\frac{1}{2}}\Gamma = \Gamma G^{-\frac{1}{2}}$	$T^{-\frac{1}{2}}\Gamma = \Gamma W^* \Lambda^{-\frac{1}{2}} W$
Mixed product matrix B	$M^*\Gamma = G^{\frac{1}{2}}$	$G^{\frac{1}{2}} = W^* \Lambda^{\frac{1}{2}} W$
Transition probabilities $p_c(j\|i)$	$\|\langle\mu_j\|\gamma_i\rangle\|^2 = \|b_{ji}\|^2$	$\left\|\frac{1}{K}\sum_{p=0}^{K-1} \lambda_p^{\frac{1}{2}} W_K^{-p(i-j)}\right\|^2$
Correct decision probability	$P_c = \frac{1}{K}\sum_{i=0}^{K-1} \|b_{ii}\|^2$	$P_c = \left[\frac{1}{K}\sum_{p=0}^{K-1} \lambda_p^{\frac{1}{2}}\right]^2$

Summary of the SRM Method

Table 6.2 summarizes the computation procedure of the SRM method, on the left in the general case, and on the right in the presence of GUS.

Problem 6.5 ⋆⋆ Apply the SRM approach to find the optimal decision in a binary system with equiprobable symbols and with a real inner product X.

6.6 SRM with Mixed States Having the GUS

We have seen that the (GUS) can be extended from pure states to density operators, with the condition

$$\rho_i = S^i \rho_0 \, (S^i)^* \,, \qquad i = 0, 1, \ldots, K-1. \tag{6.51}$$

This extension entails, for the factors, the relation

$$\gamma_i = S^i \, \gamma_0 \,, \qquad i = 0, 1, \ldots, K-1$$

and the same symmetry is transferred to the measurement operators,

$$Q_i = S^i \, Q_0 \, (S^i)^* \,, \qquad i = 0, 1, \ldots, K-1 \tag{6.52}$$

as well as to the measurement factors

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

With the SRM method in the presence of GUS, the performance evaluation becomes simpler, as already seen with the pure states, but some complication arises, due to the fact that Gram's matrix is not circulant, but *block circulant* [4]. However, we can still manage to formulate the computation based on the DFT, arriving at results explicit enough.

Relation (6.35) still holds, in particular

$$M_0 = T^{-1/2} \Gamma = \Gamma G^{-1/2}$$

so that we have two possible approaches.

6.6.1 Gram Operator Approach

This approach is based on the evaluation of the inverse square root $T^{-1/2}$ of the Gram operator T, and the reference measurement operator is given by (see (6.40a))

$$Q_0 = T^{-1/2}\,\rho_0\,T^{-1/2}. \tag{6.53}$$

Proposition 6.2 *The transition probabilities with mixed states having the GUS can be obtained from the reference density operator and the inverse square root of the Gram operator as*

$$p_c(j|i) = \mathrm{Tr}\left[S^{i-j}\,\rho_0\,S^{-(i-j)}\,Q_0\right] \tag{6.54}$$

with Q_0 given by (6.53). The correct decision probability is given by the synthetic formula

$$\boxed{P_c = \mathrm{Tr}\left[(\rho_0\,T^{-1/2})^2\right].} \tag{6.55}$$

In fact,

$$\begin{aligned} p_c(j|i) = \mathrm{Tr}[\rho_i\,Q_j] &= \mathrm{Tr}\left[S^i\,\rho_0\,S^{-i}\,S^j\,Q_0\,S^{-j}\right] \\ &= \mathrm{Tr}\left[S^{i-j}\,\rho_0\,S^{-(i-j)}\,Q_0\right]. \end{aligned}$$

Then, using (6.53), we obtain

$$P_c = \mathrm{Tr}\left[\rho_0 T^{-1/2}\rho_0\,T^{-1/2}\right] = \mathrm{Tr}\left[T^{-1/2}\rho_0 T^{-1/2}\rho_0\right]$$

and (6.55) follows at once.

6.6.2 Gram Matrix Approach

With the Gram matrix it is less trivial to get useful results, because they need the EID of the symmetry operator, given by (5.128)

$$S = \sum_{i=0}^{K-1} W_k^i\,P_i\ ,$$

where P_i are projectors. The Gram matrix is formed by the blocks of order h_0

$$G_{ij} = \gamma_i^*\gamma_j = \gamma_0^* S^{j-i}\gamma_0 = \sum_{k=0}^{K-1} W_K^{k(j-i)}\gamma_0^* P_k\gamma_0 = \frac{1}{K}\sum_{k=0}^{K-1} W_K^{k(j-i)} D_k \tag{6.56}$$

where

$$D_k := K\gamma_0^*\,P_k\,\gamma_0. \tag{6.57}$$

Then we find that the (i, j) block of G has the structure

$$G_{ij} = r_{i-j \,(\mathrm{mod}\, K)}. \tag{6.58}$$

Since G_{ij} depends only on the difference $(j - i)$ mod K, the matrix G turns out to be *block circulant*, with blocks of the same order $h_k = h_0$. Then one can extend what seen with pure states in Sect. 6.5.3, operating on the blocks, instead of on the scalar elements, to get the explicit factorization of G, namely[3]

$$G = W_{(h_0)}\, D\, W^*_{(h_0)}$$

where $D = \mathrm{diag}[D_0, \ldots, D_{K-1}]$ and $W_{(h_0)}$ is the $Kh_0 \times Kh_0$ block DFT matrix

$$W_{(h_0)} = \frac{1}{\sqrt{K}} \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & W_K^{-1} & W_K^{-2} & \ldots & W_K^{-2(K-1)} \\ \vdots & & \vdots & & \\ 1 & W_K^{-(K-1)} & W_K^{-2(K-1)} & \ldots & W_K^{-(K-1)(K-1)} \end{bmatrix} \otimes I_{h_0}. \tag{6.59}$$

As a consequence, the diagonal blocks are given as the DFT of the first block row of G, namely

$$D_k = \sum_{s=0}^{K-1} W_K^{-ks}\, G_{0s}. \tag{6.60}$$

Now we have to find the square root of G and this can be done as seen with pure states, but acting on blocks instead of on scalars. We find

$$G^{1/2} = W_{(h_0)}\, D^{1/2}\, W^*_{(h_0)}$$

where

$$D^{1/2} = \mathrm{diag}[D_0^{1/2}, \ldots, D_{K-1}^{1/2}].$$

In particular, the (i, j) block is given by

$$\boxed{(G^{1/2})_{ij} = \frac{1}{K} \sum_{k=0}^{K-1} W_K^{k(j-i)} D_k^{1/2}.} \tag{6.61}$$

Note that we have found the alternative expressions (6.57) and (6.60) for the diagonal blocks D_k, where the first is based on the EID of the symmetry operator.

[3] This is not a standard EID, because the diagonal blocks D_i are not diagonal matrices.

This EID is used in the proof, but we can use the alternative expression (6.57) to avoid its evaluation (which may be difficult).

To summarize:

Proposition 6.3 *With the GUS the* (i, j) *block of the Gram matrix can be written in the forms*

$$G_{ij} = \gamma_i^* \gamma_j = \gamma_0^* S^{j-i} \gamma_0 = \frac{1}{K} \sum_{k=0}^{K-1} W_K^{k(j-i)} D_k \tag{6.62}$$

where the matrices D_k *of order* h_0 *are Hermitian PSD given by*

$$D_k = \sum_{i=0}^{K-1} \gamma_0^* \gamma_i W_K^{-ki}. \tag{6.63}$$

The (i, j) *block of the matrix* $G^{1/2}$ *is given by relation (6.61).*

Now, using expression (6.61), one can obtain the transition probabilities from (6.37) with $b_{ij} = (G^{1/2})_{ij}$. For the correct decision probability one finds

$$\boxed{P_c = \text{Tr}\left[\frac{1}{K} \sum_{k=0}^{K-1} D_k^{1/2}\right]^2}. \tag{6.64}$$

The reference measurement factor μ_0 can be obtained as in (6.29), that is,

$$\mu_0 = \sum_{j=0}^{K-1} (G^{-1/2})_{ij} \gamma_j. \tag{6.65}$$

Remark on optimality. Differently from the case of pure states, the SRM method with GUS is not optimal in general with mixed states. In fact, for optimality, the further condition is required for the reference factors [11]

$$b_{00} = \mu_0^* \gamma_0 = \alpha\, I \tag{6.66}$$

where I is the identity matrix, and α a proportionality constant. Note that $b_{00} = (G^{1/2})_{00}$. As we will see, the PSK and PPM systems verify the GUS even in the presence of noise, but do not verify the further condition (6.66), hence the SRM method is not optimal.

Application to generalized GUS. The above theory of SRM for mixed states with GUS can be used for pure states having the first form of generalized GUS introduced in Sect. 5.13.4. This possibility will be applied in Sect. 7.11 to Quantum Communications systems using the QAM modulation.

Problem 6.6 ⋆ Write explicitly the block DFT matrix, defined by (6.59), for $K = 4$ and $h_0 = 2$ and prove that it is a unitary matrix.

Problem 6.7 ⋆⋆ Prove in general that the block DFT matrix, defined by (6.59), is a unitary matrix.

Problem 6.8 ⋆⋆⋆ Extend Theorem 6.3 on circulant matrices to block circulant matrices.

Problem 6.9 ⋆⋆ To check the fundamental formulas of the SRM with mixed states having the GUS, consider the following degenerate case of reference state factor in a quaternary system

$$\gamma_0 = \frac{1}{\sqrt{3}}[|\beta_0\rangle, |\beta_0\rangle, |\beta_0\rangle]$$

where $|\beta_0\rangle$ is an arbitrary pure state, and the symmetry operator S generates the other state factor in the form $\gamma_i = S^i \gamma_0$, $i = 1, 2, 3$. Find the correct decision probability P_c.

6.7 Quantum Compression with SRM

The technique of compression seen at the end of the previous chapter, for the reduction of redundancy in quantum states, can be applied to the detection based on the SRM. In practice, quantum compression is useful in numerical computations because it reduces the size of the matrices. For instance in quantum communications using the PPM format the computational complexity may become huge and the compression allows us to get results otherwise not reachable.

We recall that compression preserves the GUS and therefore we can apply the very efficient technique that combines the SRM with the GUS, *after the state compression*.

We now review the main simplifications achieved with the application of the compression to the SRM.

6.7.1 Simplification with Compression in the General Case

We first recall that all the detection probabilities can be evaluated in the compressed space exactly as in the uncompressed space, as stated by relations (5.146) and (5.147). Also, in the compressed space, the Gram operator is always diagonal (see (5.145)).

It is also convenient to recall the dimensions of the quantities involved in the compression. We refer to **mixed states**, from which we have the case of pure states as a particularization. Before compression the dimensions are

$$\underset{n\times h_i}{\gamma_i} \quad \underset{n\times h_i}{\mu_i} \quad \underset{n\times H}{\Gamma} \quad \underset{n\times H}{M} \quad \underset{H\times H}{G} \quad \underset{n\times n}{T}$$

where $H = h_0 + \cdots + h_{K-1}$ with h_i the rank ρ_i. After compression we have

$$\underset{r\times h_i}{\overline{\gamma}_i} \quad \underset{r\times h_i}{\overline{\mu}_i} \quad \underset{r\times K}{\overline{\Gamma}} \quad \underset{r\times K}{\overline{M}} \quad \underset{H\times H}{\overline{G}} \quad \underset{r\times r}{\overline{T}} \ .$$

The case of **pure states** is obtained by setting

$$h_i = 1\,, \qquad H = K\,.$$

6.7.2 Simplification of Compression in the Presence of GUS

The GUS is preserved in the compressed space, as stated by Proposition 5.10.

The main property in the presence of GUS is that the Gram operator T commutes with the symmetry operator S, that is, T and S becomes simultaneously diagonalizable, as stated by

$$T = U\,\Sigma^2\,U^*\,, \qquad S = U\,\Lambda\,U^*. \tag{6.67}$$

This allows us to establish simple formulas for both T and S, as in Propositions 5.11 and 5.12. A sophisticated technique to find a very useful simultaneous diagonalization is descried at the end of Chap. 8.

Problem 6.10 ⋆⋆ Solve Problem 6.3 introducing compression.

6.8 Quantum Chernoff Bound

We have seen that the SRM is a suboptimal method that gives an upper bound of the error probability in a quantum communications system. Another suboptimal method is given by the *quantum Chernoff bound*, which recently received a great attention, especially for Gaussian quantum states [13], as a simple mean to estimate the performance of quantum discrimination [14–16].

The Chernoff bound is usually employed in Telecommunications and Probability Theory to establish an upper bound to the error probability [17] or more in general to bound the probability that a random variable exceeds a certain quantity, based on the knowledge of the characteristic function or of the moments of the random variable. The extension of the Chernoff bound to quantum systems, leading to the *quantum Chernoff bound*, is considered in several works [13–16, 18], employing the bound as a tool to estimate the error probability in the discrimination of quantum states, both for single-mode and for multi-mode states. The Chernoff bound can be

seen also as a distance measure between operators. The Chernoff distance has been investigated, for example, in [14–16] and related to other distiguishability measures, such as fidelity.

In a recent paper [19] Corvaja has compared the Chernoff bound, and other bounds, with the SRM bound in terms of both performance and complexity.

6.8.1 Formulation

The quantum Chernoff bound has the limitation that it can be applied only to binary quantum systems. For the binary case, where the states are described by the density operators ρ_0 and ρ_1, the quantum Chernoff bound states that error probability can be bounded by the expression

$$P_e \leq \frac{1}{2} \inf_{0 \leq s \leq 1} \mathrm{Tr}\left[\rho_0^s \rho_1^{1-s}\right] \tag{6.68}$$

where s is a real parameter. Therefore, the bound requires the evaluation of the fractional power of operators (in practice of a square matrix) for all the values of the minimization parameter s. This is obtained with an eigendecomposition of the kind

$$\rho_i = U_i \, \Lambda_i U_i^* \quad \longrightarrow \quad \rho_i^s = U_i \, \Lambda_i^s U_i^*. \tag{6.69}$$

Although in general the bound requires the minimization with respect to the real value s, when the Gaussian states have the same covariance matrix or the same thermal noise component and no relative displacement (see Chap. 11), the optimum is attained for $s = 1/2$. In this case the square root of the density operators must be evaluated and the bound becomes

$$P_e \leq \frac{1}{2} \mathrm{Tr}\left[\sqrt{\rho_0}\sqrt{\rho_1}\right]. \tag{6.70}$$

In the comparison reported in [19] it is shown that for mixed states the SRM solution provides a tighter bound than the Chernoff bound in the binary case, with a comparable numerical complexity. Moreover, the SRM has the advantage that it can be applied also to the general K-ary case.

Problem 6.11 ⋆⋆ Consider the binary system specified by the pure states

$$|\gamma_0\rangle = \frac{1}{\sqrt{13}}[5, 3 - 2\mathrm{i}, 1, 3 + 2\mathrm{i}]^{\mathrm{T}}, \qquad |\gamma_1\rangle = \frac{1}{2\sqrt{13}}[1, 3 + 2\mathrm{i}, 5, 3 - 2\mathrm{i}]^{\mathrm{T}}.$$

Check that: (1) Hestrom's theory gives $P_e = 1/26$, (2) the Chernoff bound gives $P_e = 25/338$.

Appendix

On the EID of a Circulant Matrix

Let us prove Theorem 6.3. To this end, consider the matrix

$$Z := W^* G \quad \text{with} \quad W = W_{[K]}. \tag{6.71}$$

From inspection of the structure of the element Z_{ij} of Z and bearing in mind the condition (6.41), we have

$$\begin{aligned} Z_{ij} &= \frac{1}{\sqrt{K}} \sum_{t=0}^{K-1} W_K^{it} G_{tj} = \frac{1}{\sqrt{K}} \sum_{t=0}^{K-1} W_K^{it} r_{j-t \,(\mathrm{mod}\, K)} \\ &= \frac{1}{\sqrt{K}} \Big(\sum_{t=0}^{j} W_K^{it} r_{j-t} + \sum_{t=j+1}^{K-1} W_K^{it} r_{K+j-t} \Big). \end{aligned}$$

Letting $k = j - t$ in the first summation, and $k = K + j - t$ in the second, we have

$$\begin{aligned} Z_{ij} &= \frac{1}{\sqrt{K}} \Big(\sum_{k=0}^{j} W_K^{i(j-k)} r_k + \sum_{k=j+1}^{K-1} W_K^{i(K+j-k)} r_k \Big) \\ &= \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} W_K^{i(j-k)} r_k = \frac{1}{\sqrt{K}} W_K^{ij} \sum_{k=0}^{K-1} W_K^{-ik} r_k \\ &= \frac{1}{\sqrt{K}} W_K^{ij} \lambda_i \end{aligned}$$

where (see (6.45a))

$$\lambda_i := \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} W_K^{-ik} r_k.$$

From the above result we infer that the matrix Z can be written in the form

$$Z = \Lambda W^* . \tag{6.72}$$

Then, to obtain (6.45) from (6.71) and (6.72), it suffices to recall that W is unitary, then $W^* = W^{-1}$.

References

1. P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, W.K. Wootters, Classical information capacity of a quantum channel. Phys. Rev. A **54**, 1869–1876 (1996)
2. Y.C. Eldar, G.D. Forney, On quantum detection and the square-root measurement. IEEE Trans. Inf. Theory **47**(3), 858–872 (2001)
3. Y.C. Eldar, G.D. Forney, Optimal tight frames and quantum measurement. IEEE Trans. Inf. Theory **48**(3), 599–610 (2002)
4. G. Cariolaro, G. Pierobon, Performance of quantum data transmission systems in the presence of thermal noise. IEEE Trans. Commun. **58**(2), 623–630 (2010)
5. G. Cariolaro, G. Pierobon, Theory of quantum pulse position modulation and related numerical problems. IEEE Trans. Commun. **58**(4), 1213–1222 (2010)
6. G. Cariolaro, R. Corvaja, G. Pierobon, Compression of pure and mixed states in quantum detection, in *2011 IEEE Global Telecommunications Conference (GLOBECOM, 2011)*, pp. 1–5 (2011)
7. R. Penrose, A generalized inverse for matrices. Math. Proc. Camb. Philos. Soc. **51**, 406–413 (1955)
8. R.A. Horn, C.R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1998)
9. A.S. Holevo, Statistical decision theory for quantum systems. J. Multivar. Anal. **3**(4), 337–394 (1973)
10. M. Sasaki, T. Sasaki-Usuda, M. Izutsu, O. Hirota, Realization of a collective decoding of code-word states. Phys. Rev. A **58**, 159–164 (1998)
11. Y.C. Eldar, A. Megretski, G.C. Verghese, Optimal detection of symmetric mixed quantum states. IEEE Trans. Inf. Theory **50**(6), 1198–1207 (2004)
12. M. Sasaki, K. Kato, M. Izutsu, O. Hirota, Quantum channels showing superadditivity in classical capacity. Phys. Rev. A **58**, 146–158 (1998)
13. C. Weedbrook, S. Pirandola, R. Garcí-Patrón, N.J. Cerf, T.C. Ralph, J.H. Shapiro, S. Lloyd, Gaussian quantum information. Rev. Mod. Phys. **84**, 621–669 (2012)
14. K.M.R. Audenaert, M. Nussbaum, A. Szkoła, F. Verstraete, Asymptotic error rates in quantum hypothesis testing. Commun. Math. Phys. **279**(1), 251–283 (2008)
15. K.M.R. Audenaert, J. Calsamiglia, R. Muñoz Tapia, E. Bagan, L. Masanes, A. Acin, F. Verstraete, Discriminating states: the quantum Chernoff bound. Phys. Rev. Lett. **98**, paper no. 160501 (2007)
16. J. Calsamiglia, R. Muñoz Tapia, L. Masanes, A. Acin, E. Bagan, Quantum chernoff bound as a measure of distinguishability between density matrices: application to qubit and Gaussian states. Phys. Rev. A **77**, 032311 (2008)
17. J.M. Wozencraft, *Principles of Communication Engineering* (Wiley, New York, 1965)
18. M. Nussbaum, A. Szkola, The Chernoff lower bound for symmetric quantum hypothesis testing. Ann. Stat. **37**(2), 1040–1057 (2009)
19. R. Corvaja, Comparison of error probability bounds in quantum state discrimination. Phys. Rev. A **87**, paper no. 042329 (2013)