Helstrom's Theory on Quantum Binary Decision Revisited

Gianfranco Cariolaro and Alberto Vigato
Department of Information Engineering (DEI), University of Padova, via G. Gradenigo 6/B - Padova, PD I–35131, Italy
Email: {cariolar,alvigato}@dei.unipd.it

***Abstract*—For a binary system specified by the density operators ρ_0 and ρ_1 and by the prior probabilities q_0 and q_1, Helstrom's theory permits the evaluation of the optimal measurement operators and of the corresponding maximum correct detection probability. The theory is based on the eigendecomposition (EID) of the operator, given by the difference of the weighted density operators, namely $D = q_1\rho_1 - q_0\rho_0$. In general, this EID is obtained explicitly only with pure states, whereas with mixed states it must be carried out numerically. In this letter we show that the same evaluation can be performed on the basis of a modified version of the Gram matrix. The advantage is due to the fact that the outer products of density operators are replaced by inner product, with a considerable dimensionality reduction. At the limit, in quantum optical communications the density operators have infinite dimensions, whereas the inner products are simply scalar quantities. The Gram matrix approach permits the explicit (not numerical) evaluation of a binary system performance in cases not previously developed.**

I. Introduction

In 1970 Helstrom *et. al.* [1] evaluated the optimal performance (minimum error probability) of a binary quantum system, both with pure and mixed states. Till now, this result represents the only explicit solution of quantum optimization. In fact, for multilevel quantum systems the optimal solution is known in very few particular cases.

The formulation of Helstrom's theory is the following. Let be ρ_0 and ρ_1 the density operators of the given binary quantum system and let q_0 and q_1 be the corresponding prior probabilities. Then, form the *decision operator* as $D = q_1\rho_1 - q_0\rho_0$ and consider its eigendecompositon (EID), which has the general form

$$D = q_1\rho_1 - q_0\rho_0 = \sum_k \eta_k |\eta_k\rangle\langle\eta_k| \tag{1}$$

where η_k are the eigenvalues and $|\eta_k\rangle$ the corresponding eigenvectors. Finally, the correct detection probability is given by

$$P_c = q_0 + \sum_{\eta_k>0} \eta_k \tag{2}$$

and the optimal measurement operators by

$$\Pi_1 = \sum_{\eta_k>0} |\eta_k\rangle\langle\eta_k|\,, \qquad \Pi_0 = \sum_{\eta_k<0} |\eta_k\rangle\langle\eta_k|\,. \tag{3}$$

However, the explicit evaluation poses severe difficulties when the dimensions of the underlying Hilbert space is larger than two [2].

Here, we reconsider Helstrom's theory presenting a method based on a "modified" version of the Gram matrix, which we call *skew Gram matrix*. The net difference is that, while the operator D is expressed in terms of *outer products* (see (1)) and has the dimension of the Hilbert space, the Gram matrix is defined in terms of *inner products*. The advantage of the replacement of outer products by inner product is a considerable dimensionality reduction. At the limit, in quantum optical communications the density operators have infinite dimensions, as stated by Glauber's theory on the representation of quantum laser radiation field [3], whereas the inner products are simply scalar quantities.

This new approach allows to find explicit results for the optimal detection, not only for the case of pure states (rank 1 density operators), as found in [1], [4], but also with mixed states. The explicit results are particularly simple when the quantum states have the *geometrical uniform symmetry* (GUS).

II. Binary detection based on the skew Gram matrix

Let $k_i = \mathrm{rank}(\rho_i)$ and assume that $k_0 + k_1 \le n$, where n is the dimension of the underlying Hilbert space. Let γ_i be the $n \times k_i$ factors of $\rho_i' = q_i\rho_i$, so that $\gamma_0\gamma_0^* = \rho_0'$ and $\gamma_1\gamma_1^* = \rho_1'$ (* denotes the conjugate transpose). Then, the state matrix is given by $\Gamma = [\gamma_0, \gamma_1]$ and the Gram matrix by

$$G = \Gamma^*\Gamma = \begin{bmatrix} \gamma_0^*\gamma_0 & \gamma_0^*\gamma_1 \\ \gamma_1^*\gamma_0 & \gamma_1^*\gamma_1 \end{bmatrix} = \begin{bmatrix} G_{00} & G_{01} \\ G_{10} & G_{11} \end{bmatrix} \tag{4}$$

where the blocks $G_{ij} = \gamma_i^*\gamma_j$ has dimension $k_i \times k_j$. The skew Gram matrix (SGM) is obtained by changing the sign of the first block row, namely

$$G_s = \begin{bmatrix} -G_{00} & -G_{01} \\ G_{10} & G_{11} \end{bmatrix}. \tag{5}$$

Note that the ordinary Gram matrix G is Hermitian and positive semidefinite (PSD), and hence G_{00}, G_{11} are Hermitian and PSD, while $G_{10} = G_{01}^*$. The SGM is neither Hermitian nor PSD. Note also that, given the density operator ρ_i', the factors γ_i are not unique, but the consequence of their multiplicity is irrelevant for quantum detection [5].

Let us first recall an importat statement from matrix analysis

Lemma 1 *For each pair of complex matrices $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times m}$ with $m \le n$, the square matrix $BA \in \mathbb{C}^{n\times n}$ has the same eigenvalues as $AB \in \mathbb{C}^{m\times m}$, counting both algebraic and geometrical multiplicity, together with an additional $n - m$ eigenvalues equal to 0.*

For the proof of this Lemma, we remand to [6]. Now, we prove the following

THEOREM 1 *The $k_0 + k_1$ non–zero eigenvalues η_k of the decision operator $D = q_1\rho_1 - q_0\rho_0$ can be evaluated as the eigenvalues of the SGM G_s defined by (5). Moreover, the related eigenvectors $|\eta_k\rangle$ of D then the optimum measurement operators from (3), can be evaluated from the eigenvectors θ_k of the SGM for the same eigenvalue η_k as $|\eta_k\rangle = c_k\Gamma\theta_k$ where the factor $c_k = 1/\sqrt{\theta_k^* G \theta_k}$ assures $\langle\eta_k|\eta_k\rangle = 1$.*

In words, the eigenvalues required to evaluate the correct detection probability according to (2), can be obtained by solving the eigenvalue equation

$$\det\begin{bmatrix} -G_{00} - \eta I_{k_0} & -G_{01} \\ G_{10} & G_{11} - \eta I_{k_1} \end{bmatrix} = 0 \qquad (6)$$

where I_k denotes the identity matrix of order k.

Proof: Let $\tilde{I} = \begin{bmatrix} -I_{k_0} & 0 \\ 0 & I_{k_1} \end{bmatrix}$. By defining $\Delta = \tilde{I}\Gamma^*$, it is easy to prove that $D = \Gamma\Delta$ and $G_s = \Delta\Gamma$. Now, from the Lemma 1, η_k, counted with their multiplicity, are eigenvalues of G_s.

To prove the second part note that $|\alpha_k\rangle = \Gamma\theta_k$ is an eigenvector of D. In fact $D|\alpha_k\rangle = D\Gamma\theta_k = \Gamma G_s\theta_k = \eta_k\Gamma\theta_k = \eta_k|\alpha_k\rangle$. Now we have to show that if $|\eta_k\rangle$ is a non–zero eigenvector of D it can be obtained by an appropriated θ_k. Let $\theta_k = \Delta|\eta_k\rangle$, we have that $G_s\theta_k = G_s\Delta|\eta_k\rangle = \Delta D|\eta_k\rangle = \eta_k\Delta|\eta_k\rangle = \eta_k\theta_k$ then it is an eigenvector for G_s and it cannot be zero, because otherwise $\Delta|\eta_k\rangle = 0 \Rightarrow \Gamma\Delta|\eta_k\rangle = D|\eta_k\rangle = \eta_k|\eta_k\rangle = 0$ with contradiction. ■

Remark. We observe that if Γ is full–rank, then SGM is full–rank from $\mathrm{rank}(\Gamma) = \mathrm{rank}(\Gamma^*\Gamma) = \mathrm{rank}(G_s)$. In this case, $\mathrm{rank}(D) = \mathrm{rank}(G_s)$ follows from Lemma 1. But it is not always true: in general we have $\mathrm{Im}(D) \subseteq \mathrm{Im}(\Gamma)$, then $\mathrm{rank}(D) \leq \mathrm{rank}(\Gamma) = \mathrm{rank}(G_s)$. A simple example which leads to $\mathrm{rank}(D) < \mathrm{rank}(G_s)$ is provided by

$$\Gamma = \left[\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] \quad \text{with } k_0 = k_1 = 2. \qquad (7)$$

Here Γ is not full–rank matrix and $\mathrm{rank}(G_s) = 3$ whereas $\mathrm{rank}(D) = 2$.

III. Check of the result with pure states

In the case of rank 1 (pure states) the weighted density operators have the form

$$\rho_0' = q_0\,|a\rangle\langle a|\,, \qquad \rho_1' = q_1\,|b\rangle\langle b| \qquad (8)$$

where $|a\rangle$ and $|b\rangle$ are normalized pure states.

A. Helstrom's approach

The decision operator D is given by

$$D = \rho_1' - \rho_0' = q_1\,|b\rangle\langle b| - q_0\,|a\rangle\langle a| \qquad (9)$$

and has rank 2 (if $|b\rangle\langle b|$ and $|a\rangle\langle a|$ are linearly independent). For the EID of D we have to know the components of the kets $|a\rangle$ and $|b\rangle$, namely

$$|a\rangle = [a_0\ a_1\ \cdots\ a_{n-1}]^{\mathsf{T}}\,,\ |b\rangle = [b_0\ b_1\ \cdots\ b_{n-1}]^{\mathsf{T}}\,. \qquad (10)$$

Hence $D_{ij} = q_1 b_i^* b_j - q_0 a_i^* a_j$, $i, j = 0, 1, \ldots, n-1$, where the components are linked by the normalization conditions $\langle a|a\rangle = \langle b|b\rangle = 1$ and $q_0 + q_1 = 1$. But the explicit evaluation of the eigenvalue equation $\det(D - \eta I)$ is not immediate, since it involves a determinant of an $n \times n$ matrix. So, another approach is followed [4], based on the fact that D has rank 2 and the eigenvectors $|\eta_i\rangle$ must belong to the two dimensional subspace spanned by D, so they are a linear combination of the states, namely $|\eta_i\rangle = A_i|a\rangle + B_i|b\rangle$. Introducing this linear combination into the eigenvalue/eigenvector definition $D|\eta_i\rangle = \eta_i|\eta_i\rangle$ and considering the normalization condition, one gets $q_1(A_i\,X^* + B_i)|b\rangle - q_0(A_i + B_i\,X)|a\rangle = \eta_i(A_i|a\rangle + B_i|b\rangle)$ where $X = \langle a|b\rangle$. Finally, considering that the states are linearly independent, so that their coefficients must be equal, we obtain the equations

$$\begin{array}{l} q_1(A_i\,X^* + B_i) = \eta_i\,B_i \\ q_0(A_i + B_i\,X) = \eta_i\,A_i \end{array}\,, \qquad i = 1, 2 \qquad (11)$$

whose solutions give the non–zero eigenvalues: $\eta_\pm = \frac{1}{2}(q_1 - q_0 \mp\sqrt{1 - 4q_0q_1|X|^2}\,)$, where $\eta_+ > 0$, $\eta_- < 0$. Hence, from (2) the correct decision probability is $P_c = q_0 + \eta_+$, that is

$$P_c = \tfrac{1}{2}\left(1 + \sqrt{1 - 4q_0q_1|X|^2}\,\right) \qquad (12)$$

which represents the so called *Helstom's bound*.

B. Skew Gram matrix (SGM) approach

Considering the normalization, $\langle a|a\rangle = \langle b|b\rangle = 1$, and the notation $X = \langle a|b\rangle$, the skew Gram matrix is

$$G_s = \begin{bmatrix} -q_0 & -\sqrt{q_0q_1}\,X \\ \sqrt{q_0q_1}\,X^* & q_1 \end{bmatrix}. \qquad (13)$$

Then, equation (6) becomes

$$\det\begin{bmatrix} -q_0 - \eta & -\sqrt{q_0q_1}X \\ \sqrt{q_0q_1}X^* & q_1 - \eta \end{bmatrix} = 0 \qquad (14)$$

which has exaclty the solutions $\eta_\pm$ obtained above.

Comparison. We have seen that Helstrom's approach is very articulated, because it starts with data expressed in terms of outer products, whereas the final result is expressed in terms of inner products (the parameter $X = \langle a|b\rangle$ in the case of pure states). The SGM approach is straightforward because it starts directly with inner products, the same we find in the final result. Note also that in the Helstrom's approach the original data are redundant (states of size n), whereas the system performance are completely determined by the geometry of inner products.

IV. An application to quantum state comparison

A problem of particular interest is quantum state comparison, where one wants to determine whether the states of a quantum system are identical or not [2]. It can be formulated as a binary problem of discrimination between a pure state $|a\rangle$ and a *uniformly* mixed state, namely

$$\rho_0 = |a\rangle\langle a|\,, \qquad \rho_1 = \frac{1}{h}\sum_{i=1}^{h}|b_i\rangle\langle b_i| \qquad (15)$$

where the h kets are supposed orthonormal, i. e. $\langle b_i|b_j\rangle = \delta_{ij}$.

We evaluate the probability of correct discrimination using the SGM approach. The factors of the above density operators are $\gamma_0 = \sqrt{q_0}\ |a\rangle$ and $\gamma_1 = \sqrt{q_1/h}\ [|b_1\rangle, \ldots, |b_h\rangle]$ and the corresponding SGM is

$$G_s = \begin{bmatrix} -q_0 & -V \\ V^* & (q_1/h)I_h \end{bmatrix} \tag{16}$$

where $V = \sqrt{q_0 q_1/h}\,[X_1, \ldots, X_h]$ with $X_i = \langle a|b_i\rangle$ and I_h is the $h \times h$ identity matrix. The eigenvalue equation (6) gives

$$\left(\frac{q_1}{h} - \eta\right)^{h-1} \left[\left(\eta - \frac{q_1}{h}\right)(\eta + q_0) + \frac{q_0 q_1}{h}\,\|X\|^2\right] = 0 \tag{17}$$

where $\|X\|^2 = \sum_{i=1}^{h} |X_i|^2$. Its $h+1$ solutions are

$$\begin{aligned} \eta_i &= \frac{q_1}{h}\ ,\ i = 1, \ldots, h-1 \\ \eta_\pm &= \frac{1}{2}\left[\frac{q_1}{h} - q_0 \pm \sqrt{\left(\frac{q_1}{h} + q_0\right)^2 - \frac{4 q_0 q_1}{h}\,\|X\|^2}\right] \end{aligned} \tag{18}$$

and are all positive, η_- excepted. Hence, (2) gives

$$P_c = \frac{1}{2}\left[1 + q_1\frac{h-1}{h} + \sqrt{\left(\frac{q_1}{h} + q_0\right)^2 - \frac{4 q_0 q_1}{h}\|X\|^2}\right]. \tag{19}$$

The case of interest is when all the $h+1$ kets in (15) have the same probability, that is when $q_1 = h/(h+1)$. In this case the error probability is $P_e = 1/(h+1)\left(1 - \sqrt{1 - |X\|^2}\right)$.

Comparison with [7], where the direct Helstrom approach is used, shows that the SGM approach is much more simpler.

V. General case of rank–2 density operators

We develop the performance evaluation of the binary quantum system when the density operators ρ_0 and ρ_1 have rank 2. A particular attention is paid for the reduction of the number of parameters to get compact and readable results.

A. Helstrom's approach

The weighted density operators have the form $\rho_0' = q_0\,(p_a\,|a\rangle\langle a| + p_c\,|c\rangle\langle c|)$, $\rho_1' = q_1\,(p_b\,|b\rangle\langle b| + p_d\,|d\rangle\langle d|)$ where p_a, p_b, $p_c = 1 - p_a$, and $p_d = 1 - p_b$ are the probabilities of the normalized states $|a\rangle$, $|b\rangle$, $|c\rangle$, and $|d\rangle$, respectively. In an n-dimensional Hilbert space the states have n components: a_i, b_i, c_i, d_i, $i = 0, 1, \ldots, n-1$. Without loss of generality we can assume that the states forming the same density operator are orthogonal [8], that is

$$\langle a|c\rangle = 0\,, \qquad \langle b|d\rangle = 0\,. \tag{20}$$

With the above specifications the decision operator becomes

$$\begin{aligned} D &= q_1\rho_1 - q_0\rho_0 \\ &= q_1\,[p_b\,|b\rangle\langle b| + p_d\,|d\rangle\langle d|] - q_0\,[p_a\,|a\rangle\langle a| + p_c\,|c\rangle\langle c|] \end{aligned} \tag{21}$$

and its entries are $D_{ij} = q_1[p_b\,b_i b_j^* + p_d\,d_i d_j^*] - q_0[p_a\,a_i a_j^* + p_c\,c_i c_j^*]$. Considering that $q_0 + q_1 = 1$, $p_a + p_c = 1$ and $p_b + p_d = 1$, the specification of the decision operator is given by q_0, p_a, p_b, $|a\rangle$, $|b\rangle$, $|c\rangle$, $|d\rangle$, where the kets are normalized and verify conditions (20).

For the evaluation of the eigenvalues of D we follow the procedure seen with pure states. Now, D has rank 4 and the eigenvectors $|\eta_i\rangle$ must belong to the four dimensional subspace spanned by D, so they are a linear combination of the states, namely $|\eta_i\rangle = A_i|a\rangle + B_i|b\rangle + C_i|c\rangle + D_i|d\rangle$. Hence, introducing this linear combination into the eigenvalue/eigenvector definition $D|\eta_i\rangle = \eta_i|\eta_i\rangle$ and considering the normalization condition and orthogonality (20), one gets

$$\begin{aligned} &A_i\,q_1\,p_a\,|a\rangle - A_i[\,q_0\,p_b X^*\,|b\rangle + \,p_d Y^*\,|d\rangle] + \\ &+B_i\,q_1[\,p_a X\,|a\rangle + \,p_c W\,|c\rangle] - B_i\,q_0\,p_b\,|b\rangle + \\ &+C_i\,q_1\,p_c\,|c\rangle - C_i\,q_0[\,p_b W^*\,|b\rangle + \,p_d Z^*\,|d\rangle] + \\ &+D_i\,q_1[\,p_a Y\,|a\rangle + \,p_c Z\,|c\rangle] - D_i\,q_0 p_d\,|d\rangle \\ &= \eta_i[A_i\,|a\rangle + B_i\,|b\rangle + C_i\,|c\rangle + D_i\,|d\rangle \end{aligned} \tag{22}$$

where we have introduced the inner products

$$X = \langle a|b\rangle,\ Y = \langle a|d\rangle,\ W = \langle c|b\rangle,\ Z = \langle c|d\rangle\,. \tag{23}$$

Finally, considering that the 4 states are linearly independent, so that their coefficients must be equal, we obtain 4 equations in the unknowns A_i, B_i, C_i, D_i and η_i. By solving with respect to η_i we obtain an algebraic equation of degree 4 whose solutions give the non zero eigenvalues. The procedure is cumbersome and we do not write explicitly the passages, since they are immediate from the SGM procedure.

B. Skew Gram matrix approach

We have to find the factors of the density operators. Considering the orthogonality condition (20), they are given by

$$\gamma_0 = q_0\,[\sqrt{p_a}\,|a\rangle, \sqrt{p_c}\,|c\rangle]\,,\ \gamma_1 = q_1\,[\sqrt{p_b}\,|b\rangle, \sqrt{p_d}\,|d\rangle]\ . \tag{24}$$

Considering again orthogonality (20) and normalization, the blocks of the SGM are given by

$$\begin{aligned} &G_{00} = \gamma_0^*\gamma_0 = q_0 \begin{bmatrix} p_a & 0 \\ 0 & p_c \end{bmatrix}, \quad G_{01} = \gamma_0^*\gamma_1 = \begin{bmatrix} x & y \\ w & z \end{bmatrix}, \\ &G_{11} = \gamma_1^*\gamma_1 = q_1 \begin{bmatrix} p_b & 0 \\ 0 & p_d \end{bmatrix}, \quad G_{10} = \gamma_1^*\gamma_0 = G_{01}^*, \end{aligned} \tag{25}$$

where

$$\begin{aligned} x &= \sqrt{q_0 q_1 p_a p_b}X, \quad y = \sqrt{q_0 q_1 p_a p_d}Y, \\ w &= \sqrt{q_0 q_1 p_c p_b}W, \quad z = \sqrt{q_0 q_1 p_c p_d}Z, \end{aligned} \tag{26}$$

are weighted inner products. We see that in general the SGM G_s depends on the seven parameters q_0, p_a, p_b, x, y, w, z. The eigenvalue equation of G_s is

$$\eta^4 + B\eta^3 + C\eta^2 + D\eta + E = 0 \tag{27}$$

where

$$\begin{aligned} E &= |xz|^2 + |wy|^2 - |x|^2 qs - |w|^2 qr - |y|^2 ps - |z|^2 pr + \\ &\quad - xw^* zy^* - x^* wz^* y + pqrs \\ D &= |x|^2(q-s) + |w|^2(q-r) + |y|^2(p-s) + \\ &\quad + |z|^2(p-r) - prq + prs \\ C &= |x|^2 + |w|^2 + |y|^2 + |z|^2 + pq - pr - qr - ps - qs + rs \\ B &= p + q - r - s = q_0 - q_1 \end{aligned} \tag{28}$$

with

$$p = q_0 p_a,\ q = q_0 p_c,\ r = q_1 p_b,\ s = q_1 p_c\,. \tag{29}$$

Equation (27) is a quartic equation, whose analytical solutions are known from the times of Cardano, but their expression are very long. When the states have equal prior probabilities we find $B = 0$ and (27) becomes the *depressed quartic equation* for the absence the third degree term. The depressed quartic equation played an important role in the history of mathematics. However, we need the sum $\eta_1 + \eta_2$ of the two positive solutions, which is more readable than the individual solutions. We find

$$\eta_1 + \eta_2 = \tag{30}$$
$$\sqrt{\frac{B^2}{4} - \frac{2C}{3} + \frac{\sqrt[3]{R+\sqrt{R^2-4S^3}}}{3\sqrt[3]{2}} + \frac{\sqrt[3]{2}S}{3\sqrt[3]{R+\sqrt{R^2-4S^3}}}} - \frac{B}{2}$$

where $R = 2C^3 - 9BDC - 72EC + 27D^2 + 27B^2E$ and $S = C^2 - 3BD + 12E$.

C. Interpretation of parameters and numerical example

We can take as reference the case of *pure states* $|a\rangle$ and $|b\rangle$ and their inner product $X = \langle a|b\rangle$, which determines the error probability through $|X|^2$ according to (12). In the ideal case the states are orthogonal, $X = 0$, and give $P_c = 1$. When $|X| > 0$ we find $P_c < 1$ and the degradation is to ascribe to the presence of the *shot noise* (in the classical interpretation).

Now, it is difficult to choose the data listed in (29) to carry out a numerical example, owing to the several conditions therein, which ultimately assured that the correspondent Gram matrix is PSD. To be sure that the data are feasible, we have chosen the density operators from Glauber's theory on *coherent states*, modeling the monochromatic electromagnetic radiations produced by a laser. In these theory a density operator $\rho(\alpha)$, represented by a matrix of infinite dimension, depends only on two parameters α and $\mathcal{N}$, with $N_\alpha = |\alpha|^2$ giving the average number of *signal photons*, and $\mathcal{N}$ the average number of *thermal noise photons*. For $\mathcal{N} = 0$ (absence of thermal noise) we find rank 1, that is pure states. Thus, to find rank 2 we have to consider a small amount of thermal noise. Note that for practical calculations, we need a finite n–dimensional approximation, of the infinite dimensional representation. As discussed in detail in [9], to get a moderately small value of the size n we have to choose a small value of $N_\alpha = |\alpha|^2$ and a rank 2 is assured with a very small value of $\mathcal{N}$. In practice, a good choice for a non symmetric case may be $n = 10$, $q_0 = 0.4$, $\alpha_0 = -1.2247 \to N_{\alpha_0} = 1.5$, $\alpha_1 = 1.3038 \to N_{\alpha_1} = 1.7$, $\mathcal{N} = 0.05$.

With these parameters we have evaluated the 10×10 density operators $\rho_0 = \rho(\alpha_0)$, $\rho_1 = \rho(\alpha_1)$, and the 10×2 factors γ_0, γ_1. For reason of space we omit the explicit numerical matrices and we write directly the blocks of the SGM, which result

$$G_{00} = \begin{bmatrix} 0.381 & 0 \\ 0 & 0.018 \end{bmatrix},\quad G_{01} = \begin{bmatrix} 0.019 & 0.011 \\ 0.011 & 0.005 \end{bmatrix},$$
$$G_{11} = \begin{bmatrix} 0.571 & 0 \\ 0 & 0.027 \end{bmatrix},\quad G_{10} = G_{01}^*\,. \tag{31}$$

From (31) we obtain (see (25)) the probabilities $p_a = 0.95240$, $p_c = 0.04534$, $p_b = 0.95243$, $p_d = 0.04533$ and the inner products $X = 0.04089$, $Y = 0.10346$, $W = 0.10345$, $Z = 0.22043$. We have evaluated the 10 eigenvalues of D, which result $\{-0.380307,\ -0.0174054,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0.0263829,\ 0.570885\}$ and give $P_c = 0.997268$, $P_e = 1 - P_c = 0.00273197$. We have checked that the 4 eigenvalues of G_s are exactly the non zero eigenvalues of D, in agreement with the theory. The coefficients of the eigenvalue equation (27) are $E = 0.000105191$, $D = 0.00183749$, $C = -0.216064$, $B = -0.198617$ and the solutions are $\{-0.380307,\ -0.0183563,\ 0.0263943,\ 0.570885\}$. To get the sum of the positive solutions (30) from the formulas, we have evaluated $R = -0.0190434$, $S = 0.0490408$. Then $\eta_1 + \eta_2 = 0.59728$, which gives again tha above probabilities.

The error probability with pure states, obtained with the same inner product $X = 0.04089$, is $P_e = 0.000401349$, that is one order of magnitude better than with rank 2. The degradation with mixed states is due to the presence of thermal noise.

VI. Rank 2 density operators with symmetry

An m–ary state constellation $\{\gamma_0, \gamma_1, \ldots, \gamma_{m-1}\}$ exhibits the *geometrically uniform symmetry* (GUS), see [5], if the states are related as $\gamma_i = S^i\gamma_0$, where S is a unitary operator S, such that $S^m = I_{\mathcal{H}}$ is the identity operator in the Hilbert space $\mathcal{H}$. With the GUS equal prior probability are assumed, that is $q_i = 1/m$.

Helstrom's approach is not very much simplified by the GUS, so we pass directly to the alternative equivalent approach.

A. SGM approach with GUS

We investigate the simplifications on the Gram matrices due to the GUS and equal prior probabilities. We know that, in the presence of GUS, the Gram matrix becomes *block–circulant*, that is its blocks G_{ij} depends only on the index differences $j - i$ [9]. In the binary case the block–circulant conditions are $G_{11} = G_{00}$, $G_{01} = G_{10} = G_{01}^*$. Inspection on (25) shows that this condition leads to the following simplifications: 1) $p_b = p_a$ and $p_d = p_c$, 2) the inner product $X = \langle a|b\rangle$, $Z = \langle c|d\rangle$ are real, 3) the inner products Y and W are conjugate. Hence, the leading blocks take the form

$$G_{00} = \frac{1}{2}\begin{bmatrix} p_a & 0 \\ 0 & p_c \end{bmatrix}, G_{01} = \frac{1}{2}\begin{bmatrix} p_a\,X & \sqrt{p_a p_c}\,Y \\ \sqrt{p_a p_c}\,Y^* & p_c\,Z \end{bmatrix} \tag{32}$$

where $p_a + p_c = 1$. In conclusion with the GUS the SGM depends only on the four parameters p_a, X, Y, Z.

The eigenvalue equation of G_s, given by (14), written in a convenient form, is

$$\eta^4 - \tfrac{1}{4}\,H\,\eta^2 + \tfrac{1}{16}\,L = 0 \tag{33}$$

where

$$\begin{aligned} H &= p_a^2(1 - X^2) + p_c^2(1 - Z^2) - 2p_a p_c|Y|^2 \\ L &= (p_a p_c)^2\left[|Y|^4 - 2(1 + XZ)|Y|^2 + (1 - X^2)(1 - Z^2)\right] \\ &= (p_a p_c)^2\left[(|Y|^2 - (1 + XZ))^2 - (X + Z)^2\right]\,. \end{aligned} \tag{34}$$

Hence, the quartic equation (27) degenerates into a bi–quadratic equation, whose solutions are straightforward. In particular, the two positive solutions are given by $\eta_{1,2} = \frac{1}{2}\sqrt{\frac{1}{2}(H \pm \sqrt{H^2-4L})}$, hence

$$\begin{aligned} P_c &= \tfrac{1}{2} + \eta_1 + \eta_2 \\ &= \tfrac{1}{2} + \tfrac{1}{2}\sqrt{\tfrac{1}{2}(H+\sqrt{H^2-4L})} + \tfrac{1}{2}\sqrt{\tfrac{1}{2}(H-\sqrt{H^2-4L})} \\ &= \tfrac{1}{2} + \tfrac{1}{2}\sqrt{H+2\sqrt{L}}. \end{aligned} \tag{35}$$

The explicit result is

$$\begin{aligned} P_c = \tfrac{1}{2} + \tfrac{1}{2}\Big(p_a^2(1-X^2) + p_c^2(1-Z^2) - 2p_a p_c |Y|^2 + \\ + 2p_a p_c \sqrt{[|Y|^2-(1+XZ)]^2-(X+Z)^2}\Big)^{1/2}. \end{aligned} \tag{36}$$

Check with pure states. The case of pure states is obtained by letting $p_a = 1$ and then $p_c = 0$. In agreement with (12), from (36) we get

$$P_c = \tfrac{1}{2} + \tfrac{1}{2}\sqrt{1-X^2}\,. \tag{37}$$

Case of orthogonality. If $Y = \langle a|d\rangle = 0$, also $W = \langle c|b\rangle = 0$. This leads to a simplification in (36), namely

$$P_c = \tfrac{1}{2} + \tfrac{1}{2}\left(p_a\sqrt{1-X^2} + p_c\sqrt{1-Z^2}\right). \tag{38}$$

B. Numerical example

A choice for a symmetric case may be $n = 10$, $q_0 = 0.5$, $\alpha_0 = -1.26491 \rightarrow N_{\alpha_0} = 1.6$, $\alpha_1 = 1.26491 \rightarrow N_{\alpha_1} = 1.6$, $\mathcal{N} = 0.05$ then, the leading blocks of the Gram matrices are

$$\begin{aligned} G_{00} = \gamma_0^* \gamma_0 &= \begin{bmatrix} 0.476206 & 0 \\ 0 & 0.0226692 \end{bmatrix}, \\ G_{01} = \gamma_0^* \gamma_1 &= \begin{bmatrix} 0.019409 & -0.0107206 \\ -0.0107206 & 0.00498695 \end{bmatrix}. \end{aligned} \tag{39}$$

From (39) we obtain (see (25)) the probabilities: $p_a = 0.95241$, $p_c = 0.04534$, $p_b = 0.95241$, $p_d = 0.04534$ and the inner products $X = 0.04076$, $Y = -0.10318$, $W = -0.10318$, $Z = 0.21999$. The eigenvalues of D are $\{-0.47558,\ -0.0218743,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0.0218743,\ 0.47558\}$ and give $P_c = 0.997268$ and $P_e = 0.00273197$. The eigenvalues of G_s are exactly the non zero eigenvalues of D, in agreement with the theory. The coefficients of the eigenvalue equation (33) are $-\frac{1}{16}L = 0.000108222$, $\frac{1}{4}H = -0.226655$ and the solutions $\{-0.47558, -0.0218743, 0.0218743, 0.47558\}$.

We now check that the same results are obtained from the analytical formulas. The evaluation of (34) gives $H = 0.907538$ and $L = 0.0017714$. Hence, from (36) we obtain $P_c = 0.997924$, in agreement with the above evaluation.

The case $Y = 0$ gives $H = 0.907538$, $L = 0.0017714$, and $P_c = 0.997924$, $P_e = 0.00207581$.

The error probability with pure states (with the same inner product $X = 0.04076$) is $P_e = 0.000415467$, that is one order of magnitude better than with with with rank 2. The degradation is due to the presence of thermal noise.

C. Limit to closed–form results

As established by Evarist Galois two hundred years ago, the solution of algebraic equations can be written explicitly up to the fourth order. We have seen that in the general case of the binary quantum detection the case of rank 2+2 leads just to a quartic equation and so it is the limit. In the presence of GUS the quartic equation is essentially reduced to the second order. In principle, with the GUS it would be possible to solve the case of rank 3+3, where the six order equation can be reduced to a cubic, and also the case of rank 4+4, where the eight order equation can be reduced to a quartic. No more closed form is possible, unless in particular cases.

VII. Conclusions

The Helstrom theory on binary detection is the only general explicit result available in quantum optimization, but its translation into formulas is really possible only with very small dimensions of the density operators. On the other hand, the final results depends only on the inner products of quantum states, which compress the information contained in the density operators. The skew Gram matrix approach, which gives exactly the same results, starts just from the inner products, with a dramatic simplification of the algebra involved, and permits to establish closed–form results not available elsewhere. Furthermore, the skew Gram matrix approach clearly establishes a fundamental truth, at least for binary detection: the inner products, collected in the Gram matrix, give the necessary and sufficient information to achieve the optimal detection.

References

[1] C. W. Helstrom, J. W. S. Liu, and J. P. Gordon, "Quantum mechanical communication theory", *Proceedings of the IEEE*, vol. 58, no. 10, pp. 1578–1598, Oct. 1970.

[2] J. A. Bergou, U. Herzog, and M. Hillery 2004 *Quantum State Estimation*, Lect. Notes Phys. vol. 649, ed M. Paris and J. Řeháček, Berlin: Springer, chap. 11, pp 417–465.

[3] R. J. Glauber, "Coherent and incoherent states of the radiation field", *Phys. Rev.*, vol. 131, pp. 2766–2788, Sept. 1963.

[4] C. W. Helstrom, *Quantum Detection and Estimation Theory*, New York: Academic Press, 1976.

[5] Y. C. Eldar, A. Megretski, and G. C. Verghese, "Optimal detection of symmetric mixed quantum states," *IEEE Trans. on Inf. Theory*, vol. 50, pp. 1198–1207, Jun. 2004.

[6] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge: University Press, 1985.

[7] U. Herzog, "Minimum–error discrimination between a pure and a mixed two–qubit state," *J. Opt. B*, vol. 6, pp. S24–S28, Mar. 2004.

[8] L. P. Hugston, R. Josa, and W. K. Wootters, "A complete classification of quantum ensembles having a given density matrix," *Phisics Letter A*, vol. 183, pp. 14–18, Nov. 1993.

[9] G. Cariolaro and G. Pierobon, "Performance of quantum data transmission systems in the presence of thermal noise", *IEEE Trans. on Commun.*, vol. 58, pp. 623–630 Feb. 2010.