QUANTUM DETECTION AND ESTIMATION THEORY

SANJOY K. NITTER(2) and S.K. YOUNG(3)

(2) Department of Electrical Engineering and Computer Science
MIT, Cambridge, MA 02139

(3) Department of Mathematics, MIT, Cambridge, MA 02139
Now at IBM Corporation, McLean, VA

1. INTRODUCTION

In classical communication theory, the message to be transmitted is
modulated and the resulting signal propagates through a given channel to
produce a received waveform. The function of the receiver is to recover
the signal from the received waveform, perhaps in an optimum manner,
on the basis of some performance criterion. The input to the
receiver may have some additive noise added to the received waveform. It
is assumed that the receiver can be constructed independently of the model
of the received waveform and the additive noise. Moreover, it is assumed
that the optimum receiver can be physically realized.

In communication at optical frequencies neither of these two assump-
tions are valid. No matter what measurement we make of the received field,
the outcome is random whose statistics depend on the measurement being
made. This is a reflection of the laws of quantum physics. Furthermore,
there is no guarantee that the measurement characterizing the receiver
can be actually implemented.

In this paper, we announce some results on the M-ary quantum detec-
tion problem. Full details will be published elsewhere [1]. For related
work, see [2] and [3]. The proof in [2] appears to be incorrect and in
[3], a complete duality theory is not presented.

It will be assumed that the reader is familiar with the notions of convex analysis in infinite dimensional spaces as for example, presented
in [4].

In the classical formulation of detection theory (Bayesian hypothesis
testing) it is desired to decide which one of n possible hypotheses
H1, ..., Hn is true, based on observation of a random variable whose
probability distribution depends on the several hypotheses. The decision
tables certain costs that depend on which hypothesis is selected and which
hypothesis corresponds to the true state of the system. A decision
procedure or strategy prescribes which hypothesis is to be chosen for
each possible outcome of the observed data; in general, it may be
necessary to use a randomized strategy which specifies the probabilities
with which each hypothesis should be chosen as a function of the observed
data. The detection problem is to determine an optimal decision strategy.

In the quantum formulation of the detection problem, each hypothesis
Hj corresponds to a possible state $\rho_j$ of the quantum system under

consideration. Unlike the classical situation, however, it is not possi-
ble to measure all relevant variables associated with the state of
the system and to specify meaningful probability distributions for the
resulting values. For the quantum detection problem it is necessary to
specify not only the procedure for processing the experimental data, but
also what data to measure in the first place. Hence the quantum detec-
tion problem involves determining the entire measurement process, or, in
mathematical terms, determining the probability operator measure cor-
responding to the measurement process (for a discussion see [5]).

2. FORMULATION OF THE QUANTUM DETECTION PROBLEM

Let $\mathcal{H}$ be a separable complex Hilbert space corresponding to
the physical variables of the system under consideration. There are n hy-
potheses $H_1, ..., H_n$ about the state of the system, each corresponding to
a different density operator $\rho_j$ every $\rho_j$ is a nonnegative definite
selfadjoint trace-class operator on $\mathcal{H}$ with trace 1 and is the analog
of the distribution functions in the classical problem. Let $S$ denote
the set $S = \{1, ..., n\}$. A general decision strategy is determined by a
probability operator measure (POM) $m : 2^S \to L_\infty(\mathcal{H})^+$ [2]; in this case the
POM effecting the decision needs only n components $m_1, ..., m_n$ where
each $m_j$ is a positive selfadjoint bounded linear operator on $\mathcal{H}$ and

$$\sum_{j=1}^{n} m_j = I \quad .$$

The measurement outcome is an integer $i \in S$; the conditional probability
that the hypothesis $H_i$ is chosen when the state of the system is $\rho_j$
is given by

$$Pr(i|j) = \text{tr}(\rho_j m_i) \quad , \quad i, j = 1, ..., n \quad .$$

We remark that it is crucial here to formulate the problem in terms of
general probability operator measures rather than resolutions of the
identity. For example, an instrument which simply chooses an arbitrary
hypothesis with probability $1/n$ without even interacting with the
system corresponds to a measurement process with the POM given by

$$m_j = I/n \quad ;$$

these are certainly not projections.

We denote by $C_{ij}$ the cost associated with choosing hypothesis $H_i$
when $H_j$ is true. For a specified decision procedure effecting the POM
$(m_1, ..., m_n)$, the risk function is the conditional expected cost given that

$(1)$ L$_\infty(\mathcal{H})^+$ denotes the space of bounded positive selfadjoint
operators on $\mathcal{H}$.
the system is in the state \( R_j \), i.e.,
\[
R_j = \sum_{i=1}^{n} \mathbf{C}_{ij, n_i} \mu_j.
\]

If now \( \mu_j \) specifies a prior probability for hypothesis \( H_j \), the Bayes cost is the posterior expected cost
\[
R_j = \sum_{i=1}^{n} \mathbf{R}(j) \mu_j = \sum_{i=1}^{n} \mathbf{f}_{ij} \mu_j
\]
where \( \mathbf{f}_{ij} \) is the selfadjoint trace-class operator
\[
\mathbf{f}_{ij} = \sum_{j=1}^{n} c_{ij} \mathbf{f}_j
\]
The quantum detection problem is to find \( n_1, \ldots, n_n \) so as to minimize (3) subject to the constraint (1) and subject to the condition that the operators \( n_j \) be selfadjoint and nonnegative definite, \( n_j \geq 0 \).

3. THE QUANTUM DETECTION PROBLEM AND ITS DUAL

Let \( H \) be a complex Hilbert space. The real linear space of compact self-adjoint operators \( K(H) \) with the operator norm is a Banach space whose dual is isometrically isomorphic to the real Banach space \( \tau_s(H) \) of self-adjoint trace-class operators with the trace norm, i.e.,
\[
\tau_s(H) = \tau_s(H) \text{ under the duality}
\]
\[
\langle A, B \rangle = \text{tr}(AB) \leq |A|_{\tau_s} |B|_{\tau_s} \text{ for } A \in \tau_s(H), B \in K(H).
\]

Here, \( |A| = \sup \{|\mathbf{B} \psi| : \mathbf{B} \in K(H), |\psi| \leq 1\} = \sup \{|\text{tr}(AB) : A \in \tau_s(H), |A| \leq 1|\}
\]
and \( |A|_{\tau_s} \) is the trace norm \( \langle A, A \rangle < \infty \) where \( \tau_s(H) \) and \( |A| \) are the eigenvalues of \( A \) repeated according to multiplicity. The dual of \( \tau_s(H) \) with the trace norm is isometrically isomorphic to the space of all linear bounded self-adjoint operators, i.e., \( \tau_s(H) = \tau_s(H) \text{ under the duality}
\]
\[
\langle A, B \rangle = \text{tr}(AB) \text{ for } A \in \tau_s(H), B \in K(H).
\]

Moreover, the orderings are compatible in the following sense. If \( K_s(H), \tau_s(H) \), and \( L_s(H) \) denote the convex cones of nonnegative definite operators in \( K_s(H) \), \( \tau_s(H) \), and \( L_s(H) \) respectively, then
\[
\langle K_s(H), \tau_s(H) \rangle = \tau_s(H) + \text{ and } \langle L_s(H), \tau_s(H) \rangle = L_s(H)
\]
where the associated dual spaces are to be understood in the sense defined above.

Let \( f_j \) be given elements of \( \tau_s(H) \) (as defined in (4)), \( j = 1, \ldots, n \).

Define the functionals \( F_j : L_s(H) \rightarrow \mathbb{R} \) by
\[
F_j(A) = \delta_{10}(A) + \text{tr}(f_j A) \text{ for } A \in \tau_s(H), j = 1, \ldots, n.
\]

where \( \delta_{10}(A) \) denotes the indicator function for the set \( \tau_s(H) \) of nonnegative definite operators, i.e., \( \delta_{10}(A) = 0 \) if \( A \geq 0 \) and \( +\infty \) otherwise.

Each \( F_j \) is proper convex and \( w^* \)-lower semicontinuous on \( L_s(H) \), since \( L_s(H) \) is a \( w^* \)-closed convex cone and \( A \mapsto \text{tr}(f_j A) \) is a continuous (in fact \( w^* \)-continuous) linear functional on \( L_s(H) \).

Define the function \( G : L_s(H) \rightarrow \mathbb{R} \) by
\[
G(A) = \delta_{00}(A), \quad A \in \tau_s(H),
\]
that is \( G(A) = 0 \) if \( A = 0 \) and \( G(A) = +\infty \) if \( A \neq 0 \); \( G \) is trivially convex and lower semicontinuous. Let \( m = (m_1, \ldots, m_n) \) denote an element of \( L_s(H)^n \), the Cartesian product of \( n \) copies of \( L_s(H) \). Then the quantum detection problem (3) may be written
\[
P = \inf \left\{ \sum_{j=1}^{n} F_j(m_j) + G[I - L_m] : m = (m_1, \ldots, m_n) \in L_s(H)^n \right\}
\]
where \( L : L_s(H)^n \rightarrow L_s(H) \) is the continuous linear operator
\[
\langle L(m), n \rangle = \sum_{j=1}^{n} n_j, \quad m \in L_s(H)^n.
\]
We consider a family of perturbed problems defined by
\[
P(A) = \inf \left\{ \sum_{j=1}^{n} F_j(m_j) + G(\Lambda - L_m) : m = (m_1, \ldots, m_n) \in L_s(H)^n \right\}
\]
where \( P(\cdot) \) is a convex function \( L_s(H) \rightarrow \mathbb{R} \) and \( P = P(I) \). Note that we are taking perturbations in the equality constraint, i.e., the problem \( P(A) \) requires that every feasible \( m \) satisfy \( L_m = 0 \). We remark that \( G(\cdot) \) is nowhere continuous, so that there is certainly no Kuhn-Tucker point \( \bar{x} \) such that \( G(\cdot) \) is continuous at \( \bar{x} \).

In order to construct the dual problem corresponding to the family of perturbed problems (9) we must calculate the conjugate functions of \( F_j \) and \( G \). We would like to pose the dual problem in the space \( \tau_s(H) \), so we consider \( L_s(H) = \tau_s(H)^* \) and compute the pre-conjugates of \( F_j, G \). It can be shown that the dual problem is
\[
(P^*) (\tau_s(H)) = \sup \{ \text{tr}(\tau_s(H)) : \tau_s(H) \}
\]
It is not too difficult to show that the primal problem has a solution and that there is no duality gap. The difficult part is to show that the dual problem has solutions. It turns out that the level sets of the dual cost functions are bounded in \( \tau_s(H) \) but not weakly compact. This suggests that
we imbed \( \mathcal{T}_s(\mathcal{H}) \) in its bidual \( \mathcal{T}_s(\mathcal{H})^{**} = L_s(\mathcal{H})^* \) and extend the dual problem to the larger space; it will then turn out that there are solutions in \( \mathcal{T}_s(\mathcal{H}) \). This approach works because \( \mathcal{T}_s(\mathcal{H}) \) has a natural topological complement as a subset of \( L_s(\mathcal{H})^* \).

**Proposition 1.** \( L_s(\mathcal{H})^* = \mathcal{T}_s(\mathcal{H}) \otimes_1 (\mathcal{K}_s(\mathcal{H})) \) where \( J \) is the canonical imbedding of \( \mathcal{K}_s(\mathcal{H}) \) in \( L_s(\mathcal{H}) \). In other words, every bounded linear functional \( y \) on \( L_s(\mathcal{H}) \) may be uniquely represented in the form \( y = y_{ac} \otimes y_{sg} \) where \( y_{ac} \in \mathcal{T}_s(\mathcal{H}) \) and \( y_{sg} \in \mathcal{K}_s(\mathcal{H}) \), and

\[
|y| = |y_{ac}| |y_{sg}|.
\]

Before calculating the dual problem, it is necessary to determine what the positive linear functions look like in terms of the decomposition provided by Proposition 1.

**Proposition 2.** Let \( y \in L_s(\mathcal{H})^* \). Then \( y \in [L_s(\mathcal{H})^*]^+ \) iff \( y_{ac} \in \mathcal{T}_s(\mathcal{H})^+ \) and \( y_{sg} \in [L_s(\mathcal{H})^*]^+ \).

It can be shown that in this enlarged space the dual problem \(*\) is given by

\[
\Psi(I) = \sup \{ y(I) - P(y) \} \text{ where } \Psi(I) = \sup \{ tr(y_{ac} \otimes y_{sg}) : y_{ac} \in \mathcal{T}_s(\mathcal{H})^+, \ y_{sg} \leq f_j, j = 1, \ldots, n \}.
\]

**Proposition 3.** \( \Psi(I) \) is continuous at \( I \), and hence \( \Psi(I) \neq \emptyset \). In particular, \( *I(I) * = P(I) \) and the dual problem \( *I(I) * \) has solutions.

It is now an easy matter to show that the dual problem actually has solutions in \( \mathcal{T}_s(\mathcal{H}) \), that is solutions in \( L_s(\mathcal{H})^* \) with \( 0 \) singular part. This leads to the main theorem.

**Main Theorem.** Let \( \mathcal{H} \) be a complex Hilbert space and suppose \( (f_1, \ldots, f_n) \in \mathcal{T}_s(\mathcal{H})^n \). Then the quantum detection problem has solutions. Moreover, the following statements are equivalent for \( m \in \{m_1, \ldots, m_n\} \in [1, n] \):

1) \( m \) solves the quantum detection problem

2) \( \sum_{j=1}^{n} m_j I_j = I; \ m_j \geq 0 \text{ for } i=1, \ldots, n \)

3) \( \sum_{j=1}^{n} m_j f_j I_j \leq f_i \text{ for } i=1, \ldots, n \)

Under any of the above conditions it follows that

\[
y = \sum_{j=1}^{n} f_j m_j \text{ is self-adjoint and is the unique solution of the dual problem.}
\]

**4. Final Remarks**

Quantum estimation theory is the continuous counterpart of the problem treated here. The main difficulties here are in problem formulation and having an adequate theory of integration with respect to operator-valued measures. For details, see [1] and [6].

**References**


