Quantum Information Processing (2025) 24:310
https://doi.org/10.1007/s11128-025-04924-1

Quantum decision theory—minimax approach

Andrzej Łuczak[1]

Received: 21 February 2025 / Accepted: 25 August 2025

Abstract
We show a connection between the minimax and Bayes approaches in quantum decision theory in a general setting of normal states on a von Neumann algebra. In particular, the quantum minimax theorem is proven in a fairly general situation, and it is shown that every minimax strategy is Bayes for some a priori distribution on the set of states—a so-called *least favourable prior*. Minimax strategies with constant risk are investigated in some detail. It turns out that in dimension greater than two such a strategy can be Bayes for a non-uniform a priori distribution which, at the same time, is a least favourable prior.

Keywords Quantum decision theory · Minimax and Bayes strategies · Risk function · Normal states of a von Neumann algebra

Mathematics Subject Classification Primary: 46L53 · Secondary: 62C20

Introduction

In this paper, we make a contribution to the theory of quantum statistical decisions in the minimax approach. In spite of more than 30 years of investigations in quantum statistical decisions, this part of the theory is still barely touched upon. As a matter of fact, there are only a few papers dealing in one way or another with the subject. Namely, in [5] equivalence between solving the minimax problem and solving an optimization problem in the Bayesian setup is shown to hold true in finite dimension, and for the problem of correct detection in the case of two states, the existence of an optimal Bayes strategy which is also minimax is proven. Moreover, the question of minimax unambiguous discrimination for pure states is considered. In [25], the quantum minimax theorem is shown, while in [21] the minimax strategy for the problem of correct detection in dimensions 2 and 3 is applied in several cases involving coherent states. The analysis is based on results connecting the minimax strategy with

✉ Andrzej Łuczak
andrzej.luczak@wmii.uni.lodz.pl

1 Faculty of Mathematics and Computer Science, Łódź University, ul. S. Banacha 22, 90-238 Łódź, Poland

Published online: 23 September 2025

the Bayes one obtained in (hardly available) paper [9] (and to some extent considered also in [5]); however, it seems that some corrections to these results should be made. Some questions concerning minimax problems, however, in a different setting, are considered in [23]. An interesting specific approach to discrimination problems, containing also minimax questions, is presented in [20]. Finally, a relative (although, strictly speaking, different) problem of minimax discrimination of two Pauli channels is discussed in [6].

In this respect, it is interesting to observe that the Bayes approach received much more attention, so much more that even quoting all relevant papers in this area seems to be hardly possible. An interested reader may find more information in survey papers [1, 2, 7].

A particular instance of the general setting of quantum decision theory appears when we are dealing with a finite family of states, in which case we speak of *discrimination* between the states. Such an assumption will be adopted in the second part of the paper. In particular, the situation when a minimax strategy has constant risk is analysed in some detail, and it is shown that for pure states in dimension greater than two such a strategy exists which is Bayes for a non-uniform a priori distribution being a least favourable prior (see below). In the first part, we prove the quantum minimax theorem and show that each minimax strategy is Bayes for some a priori distribution on the set of states which maximises the Bayes risk—a so-called *least favourable prior*.

Although our considerations are devoted mainly to the minimax approach, Theorems 16 and 16' give straightforward criteria to determine when a simple strategy for an arbitrary number of pure states is Bayes.

In our analysis, we aim at full generality considering normal states on a von Neumann algebra.

It is probably worth mentioning that today quantum decision theory, in particular, and quantum statistics, in general, has substantially broaden its scope, and not only entered infinite dimension but also moved from the full algebra $\mathbb{B}(\mathcal{H})$ of all bounded operators in a Hilbert space to a general von Neumann algebra. It seems that such an approach may have substantial value for quantum information. Between numerous examples supporting this stand, one can cite, e.g. the papers [3, 10, 11, 14–16] to mention only a few.

1 Preliminaries and notation

Let $\mathcal{M}$ be an arbitrary von Neumann algebra with identity $\mathbb{1}$ acting on a Hilbert space $\mathcal{H}$. $\mathcal{M}_*$ is denoted the predual of $\mathcal{M}$, i.e. the space of all σ-weakly continuous ($\equiv$ normal) linear functionals on $\mathcal{M}$. A linear functional ρ on $\mathcal{M}$ is said to be *positive* if $\rho(A) \geqslant 0$ for every $0 \leqslant A \in \mathcal{M}$. A positive ρ is said to be a *state* if

$$\|\rho\| = \rho(\mathbb{1}) = 1.$$

To each unit vector $\psi \in \mathcal{H}$, there corresponds a normal state, denoted by the same symbol and called a *vector state* (or *pure state*, especially in the case $\mathcal{M} = \mathbb{B}(\mathcal{H})$), defined as

$$\psi(A) = \langle\psi|A\psi\rangle, \quad A \in \mathcal{M}.$$

By a slight abuse of language, we shall speak of the vector ψ itself as a pure state. $P_{[\psi]}$ shall be denoted the projection onto the space spanned by the (unit) vector ψ (in Dirac notation $P_{[\psi]} = |\psi\rangle\langle\psi|$).

In a general setting, the problem to be dealt with may be described as follows. In a von Neumann algebra $\mathcal{M}$ describing the (bounded) observables of a physical system, we are given a family of normal states $\{\rho_\theta : \theta \in \Theta\}$ on $\mathcal{M}$ in which the system may be. When guessing the correct state of the system, we make a decision u from a set of decisions U and we pay the price $L(\theta, u)$ for making our decision u when the state was ρ_θ. (As usual, we assume that the states ρ_θ are *identifiable*, that is, that the correspondence $\theta \mapsto \rho_\theta$ is one to one.) Thus, we have a function $L\colon \Theta \times U \to \mathbb{R}$, called a *loss function*. The decisions we make are based solely on the outcomes of measurements (called also *strategies*) we perform on the system, which are *normalised positive operator valued* (abbr. POV) *measures* (called also *semispectral measures*) M on some σ-field of subsets of U, taking values in the set of positive operators in $\mathcal{M}$. If the values of M are projections, then the measure is called *spectral*. (Or the strategy is called *simple*.) To every strategy M, there corresponds the *risk function* R_M defined as

$$R_M(\theta) = \int_U L(\theta, u)\, \rho_\theta(M(du))$$

which is clearly seen to be the expectation of our loss under the assumption that the state was ρ_θ. If on the set Θ there is some a priori distribution π, then the Bayes risk r for the strategy M is defined as

$$r(M, \pi) = \int_\Theta R_M(\theta)\, \pi(d\theta) = \mathbb{E}_\pi R_M,$$

which is again the expectation of the risk function with respect to the a priori distribution π.

Two fundamental problems considered in the theory of quantum statistical decisions are the minimax problem and the Bayesian problem. The first one consists in finding a strategy $\overline{M}$ that minimises the maximal risk, i.e. such that the formula

$$\sup_{\theta\in\Theta} R_{\overline{M}}(\theta) = \inf_M \sup_{\theta\in\Theta} R_M(\theta)$$

holds, while in the second we minimise the Bayes risk, i.e. an optimal strategy $\overline{M}$ should satisfy

$$r(\overline{M}, \pi) = \inf_M r(M, \pi)$$

for a given a priori distribution π. Such optimal strategies will be called *minimax* and *Bayes*, respectively.

The space of measurements $\mathfrak{M}$ is topologized by the condition: a net of measurements $\{M_i\}$ converges to a measurement M if for every normal state ρ the probability measures $\rho \circ M_i$ weakly converge to the probability measure $\rho \circ M$. It is known that for U compact $\mathfrak{M}$ is compact (cf. [11]). As for the space Π of probability measures on Θ endowed with the topology of weak convergence, it is compact for Θ compact.

2 Examples of minimax strategies

Here, we present some conditions under which a given strategy is minimax. In particular, we give an example of the situation when a Bayes strategy is also minimax. Another type of connection between Bayes and minimax strategies will be employed in next sections. Both the theorems below are quantum counterparts of known results from classical mathematical statistics, the first one due to Lehmann and the second to Hodges and Lehmann, with proofs the same as in the classical case (see, for example, [13, p. 24]).

Theorem 1 *Let M^n be Bayes strategies for a priori distributions π^n, $n = 1, 2, \ldots$. If for some strategy $\overline{M}$ we have*

$$R_{\overline{M}}(\theta) \leqslant \limsup r(M^n, \pi^n)$$

for all $\theta \in \Theta$, then $\overline{M}$ is a minimax strategy.

Proof For any strategy M and any n, we have, since M^n is Bayes for the a priori distribution π^n,

$$\sup_{\theta\in\Theta} R_M(\theta) \geqslant \boldsymbol{E}_{\pi^n} R_M = r(M, \pi^n) \geqslant r(M^n, \pi^n),$$

and thus, for each $\theta' \in \Theta$ we obtain

$$\sup_{\theta\in\Theta} R_M(\theta) \geqslant \limsup r(M^n, \pi^n) \geqslant R_{\overline{M}}(\theta'),$$

showing that $\overline{M}$ is minimax.

From this theorem, we obtain as an immediate corollary

Theorem 2 *If $\overline{M}$ is a Bayes strategy for some a priori distribution π on Θ, and $R_{\overline{M}}(\theta) \leqslant r(\overline{M}, \pi)$ for $\theta \in \Theta$, then $\overline{M}$ is a minimax strategy. In particular, if a Bayes strategy has constant risk, then it is a minimax strategy.*

3 Quantum minimax theorem and related results

The quantum minimax theorem consists in showing the following formula

$$\inf_{M\in\mathfrak{M}}\sup_{\theta\in\Theta} R_M(\theta) = \sup_{\pi\in\Pi}\inf_{M\in\mathfrak{M}} r(M,\pi). \tag{1}$$

We are going to show that this theorem holds under natural, fairly mild, assumptions. Moreover, we show that there exists a minimax strategy.

First we shall prove two simple facts concerning weak convergence of measures.

Lemma 3 *Let f be a nonnegative lower semicontinuous function on a topological space X, and let $\{\mu_\alpha\}$ be a net of probability measures on X converging weakly to a probability measure μ. Then,*

$$\liminf_{\alpha}\int_X f\,d\mu_\alpha \geqslant \int_X f\,d\mu.$$

Proof According to [4, Chapter IV, Section 1.1], there is a family $\{f_i : i \in I\}$ of continuous bounded functions on X such that

$$f = \sup_{i\in I} f_i,$$

and taking $f_i \vee 0$ instead of f_i, we may assume that $f_i \geqslant 0$. Let $\mathfrak{F}$ be a family of all finite subsets of I directed by inclusion. For $F \in \mathfrak{F}$, put

$$f_F = \sup_{i\in F} f_i.$$

Then, $\{f_F : F \in \mathfrak{F}\}$ is an increasing net of bounded nonnegative continuous functions such that

$$f = \sup_{F\in\mathfrak{F}} f_F = \lim_{F\in\mathfrak{F}} f_F.$$

For each α and each F, we have

$$\int_X f\,d\mu_\alpha \geqslant \int_X f_F\,d\mu_\alpha,$$

and passing to the limit, we obtain

$$\liminf_{\alpha}\int_X f\,d\mu_\alpha \geqslant \liminf_{\alpha}\int_X f_F\,d\mu_\alpha = \int_X f_F\,d\mu.$$

Consequently, from the Lebesgue Monotone Convergence Theorem we get

$$\liminf_{\alpha} \int_X f\, d\mu_\alpha \geqslant \lim_{F \in \mathfrak{F}} \int_X f_F\, d\mu = \int_X f\, d\mu.$$

□

In a similar way, we prove

Lemma 4 *Let f be a nonnegative bounded upper semicontinuous function on a topological space X, and let $\{\mu_\alpha\}$ be a net of probability measures on X converging weakly to a probability measure μ. Then,*

$$\limsup_{\alpha} \int_X f\, d\mu_\alpha \leqslant \int_X f\, d\mu.$$

The following simple lemma is extremely useful since it gives another form of the expression $\sup_{\theta \in \Theta} R_M(\theta)$ occurring in the minimax formula.

Lemma 5 *For an arbitrary strategy M, the following formula holds*

$$\sup_{\theta \in \Theta} R_M(\theta) = \sup_{\pi \in \Pi} r(M, \pi).$$

Proof For each $\pi \in \Pi$, we have

$$r(M, \pi) = \int_\Theta R_M(\theta)\, \pi(d\theta) \leqslant \int_\Theta \sup_{\theta \in \Theta} R_M(\theta)\, \pi(d\theta) = \sup_{\theta \in \Theta} R_M(\theta),$$

and thus,

$$\sup_{\pi \in \Pi} r(M, \pi) \leqslant \sup_{\theta \in \Theta} R_M(\theta).$$

On the other hand, assuming that $\sup_{\theta \in \Theta} R_M(\theta) < +\infty$, we have for every $\varepsilon > 0$ and some $\theta_0 \in \Theta$,

$$\begin{aligned}\sup_{\theta \in \Theta} R_M(\theta) < R_M(\theta_0) + \varepsilon &= \int_\Theta R_M(\theta)\, \delta_{\theta_0}(d\theta) + \varepsilon \\ &= r(M, \delta_0) + \varepsilon \leqslant \sup_{\pi \in \Pi} r(M, \pi) + \varepsilon,\end{aligned}$$

giving

$$\sup_{\pi \in \Pi} r(M, \pi) \geqslant \sup_{\theta \in \Theta} R_M(\theta).$$

If $\sup_{\theta \in \Theta} R_M(\theta) = +\infty$, then for arbitrary $m > 0$ there is $\theta_0 \in \Theta$ such that

$$R_M(\theta_0) > m,$$

and consequently,

$$\sup_{\pi\in\Pi} r(M,\pi) \geqslant r(M,\delta_{\theta_0}) = \int_\Theta R_M(\theta)\,\delta_0(d\theta) = R_M(\theta_0) > m,$$

showing that $\sup_{\pi\in\Pi} r(M,\pi) = +\infty$.

Now, we are in a position to prove the quantum minimax theorem which considerably generalises the one obtained in [25, Theorem 3].

Theorem 6 *Assume that*

(i) *the decision space U is compact,*
(ii) *the loss function L is bounded from below and such that for each fixed $\theta\in\Theta$ the function $U \ni u \mapsto L(\theta,u)$ is lower semicontinuous.*

Then, the minimax formula (1) holds. Moreover, there exists a minimax strategy $\overline{M}$, i.e.

$$\inf_{M\in\mathfrak{M}} \sup_{\theta\in\Theta} R_M(\theta) = \sup_{\theta\in\Theta} R_{\overline{M}}(\theta).$$

Proof Since the loss function is bounded from below, $L(\theta,u)\geqslant c$, we may assume, taking the function $L-c$ instead of L, that $L\geqslant 0$. Then, for the risk function we have $R_M \geqslant 0$.

First we shall show the existence of a minimax strategy.

Let $\{M_i\}$ be an arbitrary net of measurements converging to a measurement M. Then for each $\theta\in\Theta$, the probability measures $\rho_\theta\circ M_i$ converge weakly to the probability measure $\rho_\theta\circ M$. On account of Lemma 3, we have

$$\begin{aligned}\liminf_i R_{M_i}(\theta) &= \liminf_i \int_U L(\theta,u)\,\rho_\theta(M_i(du))\\ &\geqslant \int_U L(\theta,u)\,\rho_\theta(M(du)) = R_M(\theta),\end{aligned}$$

showing that for each $\theta\in\Theta$ the function $\mathfrak{M}\ni M\mapsto R_M(\theta)$ is lower semicontinuous. Consequently, for each fixed $\pi\in\Pi$, we have, using Fatou's lemma,

$$\begin{aligned}\liminf_i r(M_i,\pi) &= \liminf_i \int_\Theta R_{M_i}(\theta)\,\pi(d\theta) \geqslant \int_\Theta \liminf_i R_{M_i}(\theta)\,\pi(d\theta)\\ &\geqslant \int_\Theta R_M(\theta)\,\pi(d\theta) = r(M,\pi),\end{aligned}$$

which shows that for each $\pi\in\Pi$ the function $\mathfrak{M}\ni M\mapsto r(M,\pi)$ is lower semicontinuous. Denote

$$g(M) = \sup_{\pi\in\Pi} r(M,\pi), \quad M\in\mathfrak{M}.$$

It follows that g is lower semicontinuous, and since $\mathfrak{M}$ is compact, there is $\overline{M} \in \mathfrak{M}$ such that

$$\inf_{M\in\mathfrak{M}} g(M) = g(\overline{M}),$$

i.e.

$$\inf_{M\in\mathfrak{M}} \sup_{\pi\in\Pi} r(M,\pi) = \inf_{M\in\mathfrak{M}} g(M) = g(\overline{M}) = \sup_{\pi\in\Pi} r(\overline{M},\pi).$$

By virtue of Lemma 5, we get

$$\inf_{M\in\mathfrak{M}} \sup_{\theta\in\Theta} R_M(\theta) = \inf_{M\in\mathfrak{M}} \sup_{\pi\in\Pi} r(M,\pi) = \sup_{\pi\in\Pi} r(\overline{M},\pi) = \sup_{\theta\in\Theta} R_{\overline{M}}(\theta),$$

which shows that $\overline{M}$ is a minimax strategy.

Now, again by virtue of Lemma 5, the formula (1) is equivalent to the formula

$$\inf_{M\in\mathfrak{M}} \sup_{\pi\in\Pi} r(M,\pi) = \sup_{\pi\in\Pi} \inf_{M\in\mathfrak{M}} r(M,\pi), \tag{2}$$

so we are dealing with the minimax theorem for the function r. It is easily seen that this function is affine with respect to each variable; thus, in particular, it is *concave–convexlike* in the terminology of [24]. Since $\mathfrak{M}$ is compact, and the function $\mathfrak{M} \ni M \mapsto r(M,\pi)$ is lower semicontinuous for each $\pi \in \Pi$, [24, Theorem 4.2'] yields the formula (2), and thus the formula (1).

Remark In the quantum minimax theorem obtained in [25] for the full algebra $\mathbb{B}(\mathcal{H})$, it was assumed that:

(i) the spaces Θ and U are compact,
(ii) the loss function is bounded from below and lower semicontinuous,
(iii) the map $\Theta \ni \theta \mapsto \rho_\theta$ is *regular* in the terminology of [25], which in this setting amounts to its norm-continuity.

While our result is more general, it should be mentioned that for the problem of discrimination which deals with a finite number of states and measurements defined on a finite space, the setup of [25, Theorem 3] is sufficient.

Now, we are going to deal with two fundamental questions:

(a) When a minimax strategy is Bayes for some *a priori* distribution?
(b) Does there exist a *least favourable prior*, i.e. a distribution $\overline{\pi}$ such that the formula

$$\sup_{\pi\in\Pi} \inf_{M\in\mathfrak{M}} r(M,\pi) = \inf_{M\in\mathfrak{M}} r(M,\overline{\pi}) \tag{3}$$

holds?

In the lemma that follows, the assumption of compactness alone for the underlying spaces is sufficient; however, this requires considering uniform structures on compact spaces in order to obtain the uniform continuity of a continuous function. For simplicity, we assume that these spaces are metric.

Lemma 7 *Let Θ and U be compact metric spaces with metrics d_Θ and d_U, respectively, let the loss function L be continuous, and let the map $\Theta \ni \theta \mapsto \rho_\theta$ be continuous with respect to the weak topology in $\mathcal{M}_*$. Then, the risk function R_M is continuous for each strategy M.*

Proof Take arbitrary $\theta_0 \in \Theta$, and let $\{\theta_i\}$ be an arbitrary net converging to θ_0. We have

$$\begin{aligned} |R_M(\theta_i) - R_M(\theta_0)| &\leqslant \int_\Theta |L(\theta_i, u) - L(\theta_0, u)|\, \rho_{\theta_i}(M(du)) \\ &+ \Big| \int_\Theta L(\theta_0, u)\, \rho_{\theta_i}(M(du)) - \int_\Theta L(\theta_0, u) \rho_{\theta_0}(M(du)) \Big|. \end{aligned} \tag{4}$$

For arbitrary $\varepsilon > 0$, let $\delta > 0$ be such that for arbitrary (θ', u') and (θ'', u'') the relations

$$d_\Theta(\theta', \theta'') < \delta \quad \text{and} \quad d_U(u', u'') < \delta$$

imply the inequality

$$|L(\theta', u') - L(\theta'', u'')| < \varepsilon,$$

which is possible since L is uniformly continuous on $\Theta \times U$. Choose i_0 such that for $i \geqslant i_0$ we have $d_\Theta(\theta_i, \theta_0) < \delta$. Then for each $u \in U$, we have

$$|L(\theta_i, u) - L(\theta_0, u)| < \varepsilon,$$

which yields the inequality

$$\int_\Theta |L(\theta_i, u) - L(\theta_0, u)|\, \rho_{\theta_i}(M(du)) \leqslant \varepsilon.$$

For arbitrary measurable $E \subset U$, we have $\rho_{\theta_i}(M(E)) \to \rho_{\theta_0}(M(E))$, showing that the probability measures $\rho_{\theta_i} \circ M$ converge weakly to the probability measure $\rho_{\theta_0} \circ M$. Consequently, we have

$$\Big| \int_\Theta L(\theta_0, u)\, \rho_{\theta_i}(M(du)) - \int_\Theta L(\theta_0, u) \rho_{\theta_0}(M(du)) \Big| \to 0,$$

and passing to the limit in the inequality (4), we obtain

$$\lim_i |R_M(\theta_i) - R_M(\theta_0)| \leqslant \varepsilon,$$

which gives the claim.

Lemma 8 *Let the loss function be bounded, $0 \leqslant L \leqslant c$, and such that for each fixed $u \in U$ the function $\Theta \ni \theta \mapsto L(\theta, u)$ is upper semicontinuous, and let the map $\Theta \ni \theta \mapsto \rho_\theta$ be continuous with respect to the norm topology in $\mathcal{M}_*$. Then, the risk function R_M is upper semicontinuous for each strategy M.*

Proof Observe first that for an arbitrary strategy M we have

$$|\rho_{\theta'}(M(E)) - \rho_{\theta''}(M(E))| = |(\rho_{\theta'} - \rho_{\theta''})(M(E))| \leqslant \|\rho_{\theta'} - \rho_{\theta''}\|$$

for every measurable $E \subset U$, so for the signed measure
$(\rho_{\theta'} - \rho_{\theta''}) \circ M$ and its total variation $|(\rho_{\theta'} - \rho_{\theta''}) \circ M|$ we have, in particular,

$$|(\rho_{\theta'} - \rho_{\theta''}) \circ M|(U) \leqslant 2\|\rho_{\theta'} - \rho_{\theta''}\|.$$

For arbitrary $\theta_0 \in \Theta$, let $\{\theta_i\}$ be an arbitrary net converging to θ_0. We have

$$\begin{aligned} R_M(\theta_i) = \int_U L(\theta_i, u)\, \rho_{\theta_i}(M(du)) &= \int_U L(\theta_i, u)\, \rho_{\theta_i}(M(du)) + \\ &\quad - \int_U L(\theta_i, u)\, \rho_{\theta_0}(M(du)) + \int_U L(\rho_{\theta_i}, u)\, \rho_{\theta_0}(M(du)) \\ &= \int_U L(\theta_i, u)\, (\rho_{\theta_i} - \rho_{\theta_0})(M(du)) + \int_U L(\rho_{\theta_i}, u)\, \rho_{\theta_0}(M(du). \end{aligned}$$

The first integral on the right-hand side above is estimated as follows

$$\begin{aligned} \left| \int_U L(\theta_i, u)\, (\rho_{\theta_i} - \rho_{\theta_0})(M(du)) \right| &\leqslant \int_U |L(\theta_i, u)|\, |(\rho_{\theta_i} - \rho_{\theta_0}) \circ M|(du) \\ &\leqslant 2c\|\rho_{\theta_i} - \rho_{\theta_0}\|, \end{aligned}$$

which means that

$$\int_U L(\theta_i, u)\, (\rho_{\theta_i} - \rho_{\theta_0})(M(du)) \underset{i}{\rightarrow} 0.$$

Consequently, we obtain, using Fatou's lemma and the upper semicontinuity of L,

$$\begin{aligned} &\limsup_i R_M(\theta_i) = \limsup_i \int_U L(\theta_i, u)\, \rho_{\theta_0}(M(du)) \\ &\leqslant \int_U \limsup_i L(\theta_i, u)\, \rho_{\theta_0}(M(du)) \leqslant \int_U L(\theta_0, u) \rho_{\theta_0}(M(du)) = R_M(\theta_0), \end{aligned}$$

which proves the upper semicontinuity of R_M.

Theorem 9 *Let Θ and U be compact. Assume that either*

(1) *the loss function L is continuous,*
(2) *the map $\Theta \ni \theta \mapsto \rho_\theta$ is continuous with respect to the weak topology in $\mathcal{M}_*$,*

or

(1') *the loss function is bounded,* $0 \leqslant L \leqslant c$, *and such that for each fixed* $\theta \in \Theta$ *the function* $U \ni u \mapsto L(\theta, u)$ *is lower semicontinuous, and for each fixed* $u \in U$ *the function* $\Theta \ni \theta \mapsto L(\theta, u)$ *is upper semicontinuous,*

(2') *the map* $\Theta \ni \theta \mapsto \rho_\theta$ *is continuous with respect to the norm topology in* $\mathcal{M}_*$.

Then, there exists a least favourable prior, and every minimax strategy is Bayes for this prior.

Proof On account of Lemmas 7 and 8, the assumptions (1) and (2) or (1') and (2') imply the upper semicontinuity and boundedness of the function R_M for every strategy M; moreover, Theorem 6 applies. Let $\pi^0 \in \Pi$ be arbitrary, and let $\{\pi_i\} \subset \Pi$ be an arbitrary net converging weakly to π^0. For each strategy M, we have by virtue of Lemma 4

$$\begin{aligned}\limsup_i r(M, \pi_i) &= \limsup_i \int_\Theta R_M(\theta)\, \pi_i(d\theta) \\ &\leqslant \int_\Theta R_M(\theta)\, \pi^0(d\theta) = r(M, \pi^0),\end{aligned}$$

showing that the function $\Pi \ni \pi \mapsto r(M, \pi)$ is upper semicontinuous for every strategy M.

Denote

$$h(\pi) = \inf_{M \in \mathfrak{M}} r(M, \pi), \quad \pi \in \Pi.$$

From the upper semicontinuity of the function $r(M, \cdot)$, it follows that h is upper semicontinuous; consequently, there is $\overline{\pi} \in \Pi$ such that

$$h(\overline{\pi}) = \sup_{\pi \in \Pi} h(\pi).$$

For an arbitrary minimax strategy $\widetilde{M}$, we have by virtue of Theorem 6

$$\begin{aligned}\sup_{\theta \in \Theta} R_{\widetilde{M}}(\theta) &= \inf_{M \in \mathfrak{M}} \sup_{\theta \in \Theta} R_M(\theta) \\ &= \sup_{\pi \in \Pi} \inf_{M \in \mathfrak{M}} r(M, \pi) = \sup_{\pi \in \Pi} h(\pi) = h(\overline{\pi}).\end{aligned}$$

Moreover,

$$r(\widetilde{M}, \overline{\pi}) \geqslant \inf_{M \in \mathfrak{M}} r(M, \overline{\pi}) = h(\overline{\pi}),$$

and, on account of Lemma 5, also

$$r(\widetilde{M}, \overline{\pi}) \leqslant \sup_{\pi \in \Pi} r(\widetilde{M}, \pi) = \sup_{\theta \in \Theta} R_{\widetilde{M}}(\theta) = h(\overline{\pi}).$$

Hence,

$$r(\widetilde{M},\overline{\pi})=h(\overline{\pi})=\inf_{M\in\mathfrak{M}}r(M,\overline{\pi}),$$

and thus, $\widetilde{M}$ is Bayes for the a priori distribution $\overline{\pi}$. Moreover,

$$\inf_{M\in\mathfrak{M}}r(M,\overline{\pi})=h(\overline{\pi})=\sup_{\pi\in\Pi}\inf_{M\in\mathfrak{M}}r(M,\pi),$$

which means that $\overline{\pi}$ is a least favourable prior.

Remark The existence of a least favourable prior together with a proof that every minimax strategy is Bayes for this prior was shown in [25, Theorem 6] under the assumptions:

(i) the spaces Θ and U are compact,
(ii) the loss function is continuous,
(iii) the map $\Theta \ni \theta \mapsto \rho_\theta$ is regular ($\equiv$ norm-continuous).

Again it is seen that for the problem of discrimination, the setup of [25, Theorem 6] is sufficient.

Corollary 10 *Let the assumptions of Theorem 9 hold true, and let $\overline{\pi}$ be a least favourable prior. Then, for every minimax strategy $\widetilde{M}$ we have*

$$R_{\widetilde{M}}(\theta)=c\quad \overline{\pi}-a.e..$$

Moreover,

$$R_{\widetilde{M}}(\theta)\leqslant c,\quad \textit{for all}\ \theta\in\Theta.$$

Indeed, in the notation of Theorem 9 we have

$$\sup_{\theta\in\Theta}R_{\widetilde{M}}(\theta)=r(\widetilde{M},\overline{\pi})=\int_\Theta R_{\widetilde{M}}(\theta)\,\overline{\pi}(d\theta)=\int_\Theta\sup_{\theta\in\Theta}R_{\widetilde{M}}(\theta)\,\overline{\pi}(d\theta),$$

which yields

$$R_{\widetilde{M}}(\theta)=\sup_{\theta\in\Theta}R_{\widetilde{M}}(\theta)\quad \overline{\pi}-\text{a.e.},$$

showing the claim.

4 Discrimination of states

From now on, we restrict attention to discrimination of states. Assume thus that $\Theta=\{1,\dots,n\}$ and let $\{\rho_\theta:\theta\in\Theta\}=\{\rho_1,\dots,\rho_n\}$ be a family of normal states on

$\mathcal{M}$. As the space of decisions, we take $U = \{1, \dots, n\}$, and thus, the loss function $L\colon \Theta \times U \to \mathbb{R}$ can be written in matrix form $\left[L(i, j)\right]_{i,j=1}^{n}$. The risk function has the form

$$R_M(i) = \sum_{j=1}^{n} L(i, j)\rho_i(M_j), \qquad i = 1, \dots, n,$$

where $M = (M_1, \dots, M_n)$ is a strategy.

The most important instance of this scheme appears for the loss function of the form

$$L(i, j) = 1 - \delta_{ij}, \tag{5}$$

corresponding to no loss when we guess correctly and loss 1 when not. Then,

$$\begin{aligned} R_M(i) &= \sum_{j=1}^{n} L(i, j)\rho_i(M_j) = \sum_{j\neq i} \rho_i(M_j) = \sum_{j=1}^{n} \rho_i(M_j) - \rho_i(M_i) \\ &= 1 - \rho_i(M_i), \end{aligned} \tag{6}$$

which, according to our previous considerations, is seen to be the probability of failing to identify ρ_i. Furthermore, for an *a priori* distribution $\pi = (\pi_1, \dots, \pi_n)$, we have

$$r(M, \pi) = 1 - \sum_{i=1}^{n} \pi_i \rho_i(M_i). \tag{7}$$

Since $\rho_i(M_i)$ is the probability of correctly guessing the state ρ_i, the subtrahend above is the overall probability of correct guess, called the *probability of detection*. We shall denote this probability by $\mathbb{P}_D(M, \pi)$, and thus,

$$\mathbb{P}_D(M, \pi) = \sum_{i=1}^{n} \pi_i \rho_i(M_i).$$

Let $\overline{\pi} = (\overline{\pi}_1, \dots, \overline{\pi}_n)$ be a least favourable prior. Guided by Corollary 10, we are interested in the situation when some of the $\overline{\pi}_i$ are 0. Put

$$I = \{i : \overline{\pi}_i \neq 0\}, \qquad J = \{i : \overline{\pi}_i = 0\}.$$

Lemma 11 *For an arbitrary minimax strategy $\widetilde{M}$, the risk function $R_{\widetilde{M}}$ is constant on I. Moreover,*

$$R_{\widetilde{M}}(i) \geqslant R_{\widetilde{M}}(k), \tag{8}$$

for any $i \in I$, $k \in J$

Proof The proof follows from Corollary 10.

Remark *a priori* it is not clear whether the set J may be nonempty or whether we can have strict inequality in (8). In fact, this second possibility is excluded in the considerations in [21, p. 1693], quoting results of [9] where it is stated that the risk function of a minimax strategy which is also Bayes is constant. As we shall see in Example 1, this is not the case and both possibilities may occur. However, the previous statement is true for the loss function given by the formula (5) and *faithful* states as we shall see in Theorem 14.

For the most important situation when the loss function is given by the formula (5), Lemma 11 gives the following

Proposition 12 *Let the loss function be given by the formula* (5). *For any $i \in I$, $k \in J$ and every minimax strategy $\widetilde{M}$, we have*

$$\rho_i(\widetilde{M}_k) = 0, \qquad \rho_i(\widetilde{M}_i) \leqslant \rho_k(\widetilde{M}_k),$$

and

$$\rho_j(\widetilde{M}_j) = \text{const}, \quad j \in I.$$

Proof We need only to show the equality $\rho_i(\widetilde{M}_k) = 0$ for any $i \in I, k \in J$. Put

$$\widetilde{M}_0 = \sum_{k \in J} \widetilde{M}_k.$$

Fix an arbitrary $i \in I$, and define a strategy N by the formula

$$N_j = \begin{cases} \widetilde{M}_j, & \text{for} \quad j \in I,\ j \neq i \\ \widetilde{M}_i + \widetilde{M}_0, & \text{for} \quad j = i \\ 0, & \text{for} \quad j \in J. \end{cases}$$

Let $\overline{\pi}$ be a least favourable prior. Then,

$$\begin{aligned} \sum_{j=1}^n \overline{\pi}_j \rho_j(N_j) &= \sum_{j \in I} \overline{\pi}_j \rho_j(N_j) = \sum_{\substack{j \in I \\ j \neq i}} \overline{\pi}_j \rho_j(\widetilde{M}_j) \\ &\quad + \overline{\pi}_i \rho_i(\widetilde{M}_i + \widetilde{M}_0) = \sum_{j \in I} \overline{\pi}_j \rho_j(M_j) + \overline{\pi}_i \rho_i(\widetilde{M}_0) \\ &= \sum_{j=1}^n \overline{\pi}_j \rho_j(M_j) + \overline{\pi}_i \rho_i(\widetilde{M}_0) \geqslant \sum_{j=1}^n \overline{\pi}_j \rho_j(M_j). \end{aligned} \tag{9}$$

Consequently,

$$r(\widetilde{M}, \overline{\pi}) = 1 - \sum_{j=1}^n \overline{\pi}_j \rho_j(M_j) \geqslant 1 - \sum_{j=1}^n \overline{\pi}_j \rho_j(N_j) = r(N, \overline{\pi}),$$

and since $\widetilde{M}$ is Bayes, it follows that there must be equality in (9). Thus,

$$\overline{\pi}_i \rho_i(\widetilde{M}_0) = 0,$$

i.e. $\rho_i(\widetilde{M}_0) = 0$ giving $\rho_i(\widetilde{M}_k) = 0$ for each $k \in J$.

Let us illustrate the use of the above result with the following example.

Example 1 Consider pure states $\{\psi_1, \psi_2, \psi_3\}$ on $\mathcal{M} = \mathbb{B}(\mathbb{C}^3)$ given in the standard basis of $\mathbb{C}^3$ by

$$\psi_1 = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix} \qquad \psi_2 = \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \\ 0 \end{pmatrix} \qquad \psi_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Let the loss function be given by the formula (5). We want to find a minimax strategy for these states. Let $M = (M_1, M_2, M_3)$ be such. It is clear that since ψ_3 is orthogonal to the subspace spanned by ψ_1 and ψ_2, we should take $M_3 = P_{[\psi_3]}$ because then we make no mistake while discriminating ψ_3. Let $\overline{\pi} = (\overline{\pi}_1, \overline{\pi}_2, \overline{\pi}_3)$ be a least favourable prior. M is a Bayes strategy for $\overline{\pi}$. We have $\langle\psi_3|M_3\psi_3\rangle = 1$, so Proposition 12 yields $\overline{\pi}_3 = 0$ because otherwise we would have $\langle\psi_1|M_1\psi_1\rangle = \langle\psi_2|M_2\psi_2\rangle = 1$, which is impossible. Consequently, our problem is reduced to the two-dimensional space spanned by ψ_1 and ψ_2 in which the strategy (M_1, M_2) is Bayes for the *a priori* distribution $(\overline{\pi}_1, \overline{\pi}_2)$. First observe that $\overline{\pi}_1 \neq 0$, $\overline{\pi}_2 \neq 0$. Indeed, if, for example, $\overline{\pi}_2 = 0$, then $I = \{1\}$, $J = \{2, 3\}$, and Proposition 12 yields

$$\psi_1(M_2) = 0, \quad \psi_1(M_1) \leqslant \psi_2(M_2),$$

thus

$$1 = \psi_1(\mathbb{1}) = \psi_1(M_1) + \psi_1(M_2) = \psi_1(M_1),$$

consequently,

$$\langle\psi_1|M_1\psi_1\rangle = \psi_1(M_1) = \psi_2(M_2) = \langle\psi_2|M_2\psi_2\rangle = 1,$$

which is impossible.

Taking again into account Lemma 11, we obtain $\langle\psi_1|M_1\psi_1\rangle = \langle\psi_2|M_2\psi_2\rangle$. Now, the celebrated Helstrom result (cf. [7, 8]) yields that M_1 and M_2 are the projections onto the vectors of the standard basis; moreover, $\overline{\pi}_1 = \overline{\pi}_2 = 1/2$. Thus, a minimax strategy has the form

$$M_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad M_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad M_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and $\overline{\pi} = (1/2, 1/2, 0)$ is a (unique) least favourable prior. It is not difficult to show that also this minimax strategy is unique.

As the example shows, while looking for minimax strategies we are led in a natural way to strategies which are Bayes for some *a priori* distribution $\pi = (\pi_1, \ldots, \pi_n)$ such that some of the π_i are equal to zero. This is different from the situation in customary Bayesian inference theory where it is assumed that all *a priori* probabilities are strictly positive.

Lemma 13 *Let the states ρ_i be faithful, let the loss function be given by the formula (5), and let $\widetilde{M}$ be a minimax strategy. Then, $\pi_i^0 \neq 0$, $i = 1, \ldots, n$, for every* a priori *distribution $\pi^0 = (\pi_i^0)$ such that $\widetilde{M}$ is a Bayes strategy for this distribution.*

Proof Suppose that $\pi_k^0 = 0$ for some k. If $\widetilde{M}_k = 0$, then we would have

$$\min_{M \in \mathfrak{M}} \max_{1 \leqslant i \leqslant n} R_M(i) = \max_{1 \leqslant i \leqslant n} [1 - \rho_i(\widetilde{M}_i)] = 1,$$

while for the strategy $N = (N_i)$ defined as $N_i = (1/n)\mathbb{1}$ we have

$$\max_{1 \leqslant i \leqslant n} R_N(i) = \max_{1 \leqslant i \leqslant n} [1 - \rho_i((1/n)\mathbb{1})] = 1 - 1/n,$$

which contradicts the assumption that $\widetilde{M}$ is minimax. Consequently, $\widetilde{M}_k \neq 0$, and thus, $\rho_i(\widetilde{M}_k) > 0$ for each $i = 1, \ldots, n$. Take j such that $\pi_j^0 > 0$, and define the strategy $\widehat{M}$ as

$$\widehat{M}_i = \begin{cases} \widetilde{M}_i, & \text{for } \ i \neq j \text{ and } i \neq k \\ \widetilde{M}_j + \widetilde{M}_k, & \text{for } i = j \\ 0, & \text{for } i = k \end{cases}.$$

We have

$$\begin{aligned} r(\widehat{M}, \pi^0) &= 1 - \Bigg(\sum_{\substack{i \neq j \\ i \neq k}} \pi_i^0 \rho_i(\widetilde{M}_i) + \pi_j^0 \rho_j(\widetilde{M}_j + \widetilde{M}_k) \Bigg) \\ &= 1 - \sum_{i=1}^{n} \pi_i^0 \rho_i(\widetilde{M}_i) - \pi_j^0 \rho_j(\widetilde{M}_k) \\ &= r(\widetilde{M}, \pi^0) - \pi_j^0 \rho_j(\widetilde{M}_k) < r(\widetilde{M}, \pi^0), \end{aligned}$$

which contradicts the assumption that $\widetilde{M}$ is Bayes.

The lemma above implies the following important result.

Theorem 14 *Let the states ρ_i, $i = 1, \ldots, n$, be faithful, and let the loss function be given by the formula (5). Then, every minimax strategy M has the same constant risk function.*

Proof Let $\overline{\pi} = (\overline{\pi}_1, \ldots, \overline{\pi}_n)$ be a least favourable prior, and let M be a minimax strategy. Then, M is Bayes for $\overline{\pi}$, and consequently, from Lemma 13 we obtain

$$\overline{\pi}_i \neq 0, \quad i = 1, \ldots, n,$$

and Corollary 10 yields

$$R_M(i) = c_M, \quad i = 1, \ldots, n,$$

for some constant c_M. For arbitrary two minimax strategies M' and M'', we have

$$c_{M'} = \max_{1 \leqslant i \leqslant n} R_{M'}(i) = \max_{1 \leqslant i \leqslant n} R_{M''}(i) = c_{M''},$$

and the conclusion follows.

As it turns out, the conclusion of Lemma 13 holds also for pure states in dimension two.

Proposition 15 *Let* $\{\psi_1, \psi_2, \psi_3\}$ *be pairwise linearly independent pure states in a two-dimensional Hilbert space, and let* M *be a minimax strategy. Then,* $\pi_i^0 \neq 0$*,* $i = 1, 2, 3$*, for any* a priori *distribution* $\pi^0 = (\pi_1^0, \pi_2^0, \pi_3^0)$ *such that* M *is a Bayes strategy for this distribution.*

Proof First observe that $\pi_i^0 \neq 1$ for $i = 1, 2, 3$. Indeed, if for instance, $\pi_1^0 = 1$, then a Bayes strategy for this distribution would be $M_1^0 = P_{[\psi_1]}$, $M_2^0 = P_{[\psi_1]}^{\perp}$, $M_3 = 0$ with the Bayes risk $r(M^0, \pi^0) = 0$, and consequently,

$$0 = r(M, \pi^0) = \max_{i=1,2,3} [1 - \langle \psi_i | M_i \psi_i \rangle].$$

This yields

$$\langle \psi_i | M_i \psi_i \rangle = 1, \quad i = 1, 2, 3,$$

hence by the Schwarz inequality

$$M_i \psi_i = \psi_i.$$

It follows that

$$M_i \geqslant P_{[\psi_i]},$$

and thus,

$$\mathbb{1} = \sum_{i=1}^{3} M_i \geqslant \sum_{i=1}^{3} P_{[\psi_i]},$$

a contradiction.

Now, if, for instance, $\pi_1^0 = 0$, then we would have $\pi_2^0 \neq 0$ and $\pi_3^0 \neq 0$, so

$$\langle \psi_2 | M_1 \psi_2 \rangle = \langle \psi_3 | M_1 \psi_3 \rangle = 0,$$

giving $M_1 = 0$, which as we have seen is impossible for a minimax strategy.

Let us now consider two pure states $\{\psi_1, \psi_2\}$ in a two-dimensional Hilbert space. A minimax strategy $M = (M_1, M_2)$ for these states is Bayes for a least favourable prior $\overline{\pi} = (\overline{\pi}_1, \overline{\pi}_2)$, and as we have seen in Example 1, $\overline{\pi}_1 \neq 0$, $\overline{\pi}_2 \neq 0$. Consequently, Helstrom's result cited in Example 1 yields the equality $M_1 = P_{[\xi_1]}$, $M_2 = P_{[\xi_2]}$ for an orthonormal basis $\{\xi_1, \xi_2\}$. From Lemma 11, we know that the risk function is constant, i.e. $|\langle \xi_1 | \psi_1 \rangle|^2 = |\langle \xi_2 | \psi_2 \rangle|^2$, so using again Helstrom's result we obtain the equality $\overline{\pi}_1 = \overline{\pi}_2 = 1/2$. These considerations lead us to the following question. Let $\{\psi_1, \dots, \psi_n\}$ be linearly independent pure states in an n-dimensional Hilbert space $\mathcal{H}$, and let $M = (M_1, \dots, M_n)$ be a minimax strategy for these states. On account of Theorem 9, this strategy is Bayes for a least favourable prior $\overline{\pi} = (\overline{\pi}_1, \dots, \overline{\pi}_n)$. Assume that $\overline{\pi}_i \neq 0$ for all $i = 1, \dots, n$. Then, Kennedy's results [17, 18] yield that $M_i = P_{[\xi_i]}$, $i = 1, \dots, n$ for an orthonormal basis $\{\xi_i : i = 1, \dots, n\}$ and that this optimal strategy is unique. (The reader can also consult Helstrom's book [8].) From Lemma 11, it follows that the risk function is constant, so we have a simple Bayes strategy with constant risk function. A natural conjecture is that also in this multidimensional case the least favourable prior is uniform, i.e. $\overline{\pi}_i = 1/n$, $i = 1, \dots, n$. It turns out that this is not the case; however, to get a deeper insight into the problem we must digress for a while and investigate more thoroughly some questions in Bayesian inference.

5 Simple Bayes strategies for pure states

In this section, we adopt a more general approach; namely, we shall consider a finite or infinite number of pure states $\{\psi_i\}$ in a Hilbert space $\mathcal{H}$ with an *a priori* distribution $\pi = (\pi_i)$, and a simple strategy $\left(P_{[\xi_i]}\right)$, where $\{\xi_i\}$ is an orthonormal basis in $\mathcal{H}$. Observe that results analogous to those of Kennedy's mentioned above hold true in infinite dimension; namely, for an infinite sequence of strongly linearly independent pure states there exists a unique simple Bayes strategy maximising the probability of detection (cf. [19]). (By *strong linear independence* is meant the property

$$\psi_i \notin \overline{\mathrm{Lin}\{\psi_j : j \neq i\}} \quad \text{for each } i.)$$

Let $k \geqslant 2$ be an arbitrary fixed natural number, and let $i_1, \dots, i_k$ be an arbitrary choice of distinct indices. Set

$$P_{i_1 \dots i_k} = P_{[\xi_{i_1}]} + \dots + P_{[\xi_{i_k}]}, \qquad \mathcal{H}_{i_1 \dots i_k} = P_{i_1 \dots i_k} \mathcal{H}$$

$$\widehat{\psi}_{i_r} = \frac{P_{i_1 \dots i_k} \psi_{i_r}}{\| P_{i_1 \dots i_k} \psi_{i_r} \|}, \qquad p_r = \frac{\pi_{i_r} \| P_{i_1 \dots i_k} \psi_{i_r} \|^2}{\sum_{j=1}^k \pi_{i_j} \| P_{i_1 \dots i_k} \psi_{i_j} \|^2}, \qquad r = 1, \dots, k.$$

Then, $(P_{[\xi_{i_1}]}, \ldots, P_{[\xi_{i_k}]})$ is a simple strategy in the space $\mathcal{H}_{i_1 \ldots i_k}$, $\{\widehat{\psi}_{i_1}, \ldots, \widehat{\psi}_{i_k}\}$ are pure states in this space, and $p = (p_1, \ldots, p_k)$ is some *a priori* distribution. (In this notation, we tacitly assume that the projection $P_{[\xi_r]}$ 'corresponds' to the state $\widehat{\psi}_{i_r}$ which in turn 'corresponds' to the probability p_r.)

Theorem 16 *Let $\{\psi_i\}$ be pure states in a Hilbert space $\mathcal{H}$ with an* a priori *distribution $\pi = (\pi_i)$, $\pi_i \neq 0$, let the loss function be given by the formula (5), let $\left(P_{[\xi_i]}\right)$ be a simple strategy, and let $k \geqslant 2$ be an arbitrary fixed natural number. The following conditions are equivalent*

(i) *$\left(P_{[\xi_i]}\right)$ is a Bayes strategy for the states $\{\psi_i\}$ with the a priori distribution $\pi = (\pi_i)$.*
(ii) *For any distinct $i_1, \ldots, i_k$, $P = \left(P_{[\xi_{i_1}]}, \ldots, P_{[\xi_{i_k}]}\right)$ is a Bayes strategy in $\mathcal{H}_{i_1 \ldots i_k}$ for the states $\{\widehat{\psi}_{i_1}, \ldots, \widehat{\psi}_{i_k}\}$ with the a priori distribution $p = (p_1, \ldots, p_k)$.*

Proof (i)$\Longrightarrow$(ii). We have

$$\mathbb{P}_D(P, p) = \sum_{r=1}^{k} p_r |\langle \xi_{i_r} | \widehat{\psi}_{i_r} \rangle|^2 = \frac{1}{\sum_{r=1}^{k} \pi_{i_r} \| P_{i_1 \ldots i_k} \psi_{i_r} \|^2} \sum_{r=1}^{k} \pi_{i_r} |\langle \xi_{i_r} | \psi_{i_r} \rangle|^2.$$

From the above-mentioned Kennedy's result (cf. [17] or [8]), it follows that there exists a simple Bayes strategy $M = \left(P_{[\eta_1]}, \ldots, P_{[\eta_k]}\right)$ in $\mathcal{H}_{i_1 \ldots i_k}$ for the states $\{\widehat{\psi}_{i_1}, \ldots, \widehat{\psi}_{i_k}\}$ with the a priori distribution $p = (p_1, \ldots, p_k)$. For this strategy, we have

$$\mathbb{P}_D(M, p) = \sum_{r=1}^{k} p_r |\langle \eta_r | \widehat{\psi}_{i_r} \rangle|^2 = \frac{1}{\sum_{r=1}^{k} \pi_{i_r} \| P_{i_1 \ldots i_k} \psi_{i_r} \|^2} \sum_{r=1}^{k} \pi_{i_r} |\langle \eta_r | \psi_{i_r} \rangle|^2.$$

Now, if the strategy P were not Bayes, then we would have $\mathbb{P}_D(P, p) < \mathbb{P}_D(M, p)$, and it is easily seen that taking $\eta_1, \ldots, \eta_k$ instead of $\xi_{i_1}, \ldots, \xi_{i_k}$, respectively, and keeping the remaining ξ_i unchanged, we would obtain a better simple strategy which contradicts the fact that $\left(P_{[\xi_i]}\right)$ is Bayes.

(ii)$\Longrightarrow$(i). First observe that, regardless the dimension of $\mathcal{H}$, there exists a Bayes strategy $M = (M_i)$ for the states $\{\psi_i\}$ with the a priori distribution $\pi = (\pi_i)$. This follows from results in [12] or [22]. (The reader may also consult [19] for a short proof and a discussion of the problem.) Assume that $\left(P_{[\xi_i]}\right)$ is not Bayes. Then,

$$\mathbb{P}_D(M, \pi) = \sum_i \pi_i \langle \psi_i | M_i \psi_i \rangle > \mathbb{P}_D(P, \pi) = \sum_i \pi_i |\langle \xi_i | \psi_i \rangle|^2.$$

We shall employ the following simple lemma whose proof is left to the reader.

Lemma *Let (a_i) and (b_i) be sequences of real numbers such that the series $\sum_{i=1}^{\infty} a_i$, $\sum_{i=1}^{\infty} b_i$ are convergent and $\sum_{i=1}^{\infty} a_i < \sum_{i=1}^{\infty} b_i$. Let k be an arbitrary natural number. There exist natural numbers $i_1 < \cdots < i_k$ such that $\sum_{r=1}^{k} a_{i_r} < \sum_{r=1}^{k} b_{i_r}$.*

From the lemma, it follows that there exist natural numbers $i_1 < \cdots < i_k$ such that

$$\sum_{r=1}^{k} \pi_{i_r} \langle \psi_{i_r} | M_{i_r} \psi_{i_r} \rangle > \sum_{r=1}^{k} \pi_{i_r} |\langle \xi_{i_r} | \psi_{i_r} \rangle|^2.$$

After rewriting the left- and right-hand sides of the above inequality, we obtain

$$\begin{aligned} &\sum_{r=1}^{k} \pi_{i_r} \| P_{i_1 \dots i_k} \psi_{i_r} \|^2 \sum_{r=1}^{k} p_r \langle \widehat{\psi}_{i_r} | P_{i_1 \dots i_k} M_{i_r} P_{i_1 \dots i_k} \widehat{\psi}_{i_r} \rangle \\ &\quad > \sum_{r=1}^{k} \pi_{i_r} \| P_{i_1 \dots i_k} \psi_{i_r} \|^2 \sum_{r=1}^{k} p_r |\langle \xi_{i_r} | \widehat{\psi}_{i_r} \rangle|^2. \end{aligned} \tag{10}$$

Denote

$$N_r = P_{i_1 \dots i_k} M_{i_r} P_{i_1 \dots i_k}, \qquad r = 1, \dots, k.$$

Then from the inequality (10), we get

$$\sum_{r=1}^{k} p_r \langle \widehat{\psi}_{i_r} | N_r \widehat{\psi}_{i_r} \rangle > \sum_{r=1}^{k} p_r |\langle \xi_{i_r} | \widehat{\psi}_{i_r} \rangle|^2.$$

$N = (N_1, \dots, N_k)$ is 'almost' a strategy in $\mathcal{H}_{i_1 \dots i_k}$ in the sense that

$$\sum_{r=1}^{k} N_r \leqslant P_{i_1 \dots i_k}.$$

Completing N to a true strategy (for instance, by adding $P_{i_1 \dots i_k} - \sum_{r=1}^{k} N_r$ to N_1), and keeping the same notation, we obtain

$$\mathbb{P}_D(N, p) = \sum_{r=1}^{k} p_r \langle \widehat{\psi}_{i_r} | N_r \widehat{\psi}_{i_r} \rangle > \sum_{r=1}^{k} p_r |\langle \xi_{i_r} | \widehat{\psi}_{i_r} \rangle|^2 = \mathbb{P}_D(P, p),$$

which contradicts the optimality of P.

The above theorem is especially effective in the case $k = 2$ because then we have a simple method which allows us to verify its condition (ii).

Theorem 16′ *Let $\{\psi_i\}$ be pure states in a Hilbert space $\mathcal{H}$ with an a priori distribution $\pi = (\pi_i)$, $\pi_i \neq 0$, let the loss function be given by the formula (5), and let $\left(P_{[\xi_i]}\right)$ be a simple strategy. The following conditions are equivalent*

(i) $\left(P_{[\xi_i]}\right)$ *is a Bayes strategy for the states $\{\psi_i\}$ with the a priori distribution $\pi = (\pi_i)$.*

(ii) *For any $j, l,\ j \neq l$,*

$$\begin{aligned} &\pi_j \langle \psi_j | \xi_j \rangle \langle \xi_l | \psi_j \rangle = \pi_l \langle \xi_l | \psi_l \rangle \langle \psi_l | \xi_j \rangle \\ &\pi_j |\langle \xi_j | \psi_j \rangle|^2 \geqslant \pi_l |\langle \xi_j | \psi_l \rangle|^2 \\ &\pi_l |\langle \xi_l | \psi_l \rangle|^2 \geqslant \pi_j |\langle \xi_l | \psi_j \rangle|^2. \end{aligned} \tag{11}$$

Remark In principle, the problem of finding a Bayes strategy for two pure states in a two-dimensional Hilbert space was solved long ago (cf. for instance [8] or [7]), but in a somewhat different setting. Namely, for given two pure states there are formulas describing an explicit form of two orthogonal vectors such that the projections onto these vectors are a Bayes strategy. However, our aim is different since we want to verify whether *given* two orthogonal vectors give rise to a Bayes strategy. Of course, it would be possible to solve this task by writing the corresponding formulas for the vectors of the optimal strategy and comparing them with the given vectors but it turns out much more complicated than simply checking the formulas (11).

Proof of Theorem In view of Theorem 16, we need only to show that the formulas (11) are equivalent to the fact that the strategy $(P_{[\xi_j]}, P_{[\xi_l]})$ is Bayes for the states $\{\widehat{\psi}_j, \widehat{\psi}_l\}$ with the a priori probabilities (p_j, p_l) in the two-dimensional Hilbert space $\mathcal{H}_{jl}$. The Lagrange operator for our problem (see [7] or [8]) has the form

$$\Gamma = p_j P_{[\xi_j]} P_{[\widehat{\psi}_j]} + p_l P_{[\xi_l]} P_{[\widehat{\psi}_l]} = p_j \langle \xi_j | \widehat{\psi}_j \rangle E_{\xi_j, \widehat{\psi}_j} + p_l \langle \xi_l | \widehat{\psi}_l \rangle E_{\xi_l, \widehat{\psi}_l},$$

where by $E_{\xi,\eta}$ we denote the operator defined as

$$E_{\xi,\eta} \zeta = \langle \eta | \zeta \rangle \xi, \qquad \zeta \in \mathcal{H}_{jl}.$$

Now, the necessary and sufficient conditions for the strategy $(P_{[\xi_j]}, P_{[\xi_l]})$ to be a Bayes one are

(i) $P_{[\xi_j]}(p_j P_{[\widehat{\psi}_j]} - p_l P_{[\widehat{\psi}_l]}) P_{[\xi_l]} = 0$,
(ii) $\Gamma \geqslant p_r P_{[\widehat{\psi}_r]}$ for $r = j, l$

(cf. [1, 7]; see also [8, 10] for a longer discussion). The operator on the left-hand side of the condition (i) is easily seen to be $(p_j \langle \xi_j | \widehat{\psi}_j \rangle \langle \widehat{\psi}_j | \xi_l \rangle - p_l \langle \xi_j | \widehat{\psi}_l \rangle \langle \widehat{\psi}_l | \xi_l \rangle) E_{\xi_j, \xi_l}$ and thus the condition (i) is equivalent to the equality

$$p_j \langle \xi_j | \widehat{\psi}_j \rangle \langle \widehat{\psi}_j | \xi_l \rangle = p_l \langle \xi_j | \widehat{\psi}_l \rangle \langle \widehat{\psi}_l | \xi_l \rangle, \tag{12}$$

which after inserting the expressions for p_j, p_l, $\widehat{\psi}_j$ and $\widehat{\psi}_l$ gives equivalently the first equality in (11).

Operator Γ has the following matrix form in the basis $\{\xi_j, \xi_l\}$

$$\Gamma = \begin{bmatrix} p_j |\langle \xi_j | \widehat{\psi}_j \rangle|^2 & p_j \langle \xi_j | \widehat{\psi}_j \rangle \langle \widehat{\psi}_j | \xi_l \rangle \\ p_l \langle \xi_l | \widehat{\psi}_l \rangle \langle \widehat{\psi}_l | \xi_j \rangle & p_l |\langle \xi_l | \widehat{\psi}_l \rangle|^2 \end{bmatrix},$$

while for the operators $p_j P_{[\widehat{\psi}_j]}$ and $p_l P_{[\widehat{\psi}_l]}$ we have

$$p_j P_{[\widehat{\psi}_j]} = \begin{bmatrix} p_j |\langle \xi_j | \widehat{\psi}_j \rangle|^2 & p_j \langle \xi_j | \widehat{\psi}_j \rangle \langle \widehat{\psi}_j | \xi_l \rangle \\ p_j \langle \xi_l | \widehat{\psi}_j \rangle \langle \widehat{\psi}_j | \xi_j \rangle & p_j |\langle \xi_l | \widehat{\psi}_j \rangle|^2 \end{bmatrix},$$

$$p_l P_{[\widehat{\psi}_l]} = \begin{bmatrix} p_l |\langle \xi_j | \widehat{\psi}_l \rangle|^2 & p_l \langle \xi_j | \widehat{\psi}_l \rangle \langle \widehat{\psi}_l | \xi_l \rangle \\ p_l \langle \xi_l | \widehat{\psi}_l \rangle \langle \widehat{\psi}_l | \xi_j \rangle & p_l |\langle \xi_l | \widehat{\psi}_l \rangle|^2 \end{bmatrix}.$$

The equalities (12) show that the off-diagonal entries of $p_j P_{[\widehat{\psi}_j]}$ and $p_l P_{[\widehat{\psi}_l]}$ are the same as those of Γ, so the condition (ii) takes the form

$$p_l |\langle \xi_l | \widehat{\psi}_l \rangle|^2 \geqslant p_j |\langle \xi_l | \widehat{\psi}_j \rangle|^2$$
$$p_j |\langle \xi_j | \widehat{\psi}_j \rangle|^2 \geqslant p_l |\langle \xi_j | \widehat{\psi}_l \rangle|^2,$$

which again after inserting the expressions for p_j, p_l, $\widehat{\psi}_j$ and $\widehat{\psi}_l$ gives equivalently the last two inequalities in (11).

The next example shows that in dimension greater than two it is possible to find pure states and a simple minimax strategy with constant risk function such that this strategy is Bayes with respect to a non-uniform a priori distribution being a least favourable prior.

Example 2 Let

$$\psi_1 = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{4}\sqrt{\frac{5}{2}} \\ \frac{1}{4}\sqrt{\frac{3}{2}} \end{pmatrix}, \quad \psi_2 = \begin{pmatrix} \frac{3}{10}\sqrt{\frac{5}{2}} \\ \frac{\sqrt{3}}{2} \\ \frac{1}{10}\sqrt{\frac{5}{2}} \end{pmatrix}, \quad \psi_3 = \begin{pmatrix} \frac{3}{8}\sqrt{\frac{3}{2}} \\ \frac{1}{8}\sqrt{\frac{5}{2}} \\ \frac{\sqrt{3}}{2} \end{pmatrix}$$

be pure states in the three-dimensional Hilbert space $\mathbb{C}^3$, and let

$$\xi_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \xi_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \xi_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

be orthonormal vectors determining a simple strategy M, i.e.

$$M_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad M_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Let $\pi = (\pi_1, \pi_2, \pi_3)$, $\pi_i \neq 0$, be an a priori distribution such that the simple strategy $M = \left(P_{[\xi_1]}, P_{[\xi_2]}, P_{[\xi_3]}\right)$ is Bayes for π. From the conditions (11), we obtain

$$\pi_1 = \frac{2}{5}, \quad \pi_2 = \frac{1}{3}, \quad \pi_3 = \frac{4}{15},$$

so such a distribution does exist; moreover, it is the unique a priori distribution satisfying the conditions $\pi_i \neq 0$, $i = 1, 2, 3$, for which the strategy M is Bayes. We have

$$\psi_i(M_i) = |\langle \xi_i | \psi_i \rangle|^2 = \frac{3}{4}, \quad i = 1, 2, 3,$$

which shows that the risk function for this strategy is constant,

$$R_M(i) = 1 - \psi_i(M_i) = \frac{1}{4}, \quad i = 1, 2, 3.$$

Theorem 2 implies that M is a minimax strategy. Since

$$\psi_i(M_k) = |\langle \xi_k | \psi_i \rangle|^2 \neq 0 \quad \text{for } i, k = 1, 2, 3,$$

Proposition 12 yields that for a least favourable prior $\overline{\pi} = \left(\overline{\pi}_1, \overline{\pi}_2, \overline{\pi}_3\right)$ we have $\overline{\pi}_i \neq 0$, $i = 1, 2, 3$. On account of Theorem 9, for a least favourable prior the strategy M is Bayes, and thus, $\pi = (2/5, 1/3, 4/15)$ is the only least favourable prior.

6 Discussion

The paper is devoted mainly to the minimax approach in quantum decision theory. In particular, a quantum minimax theorem is proven, the existence of a minimax strategy and a least favourable prior is shown, and for the case of discrimination of faithful states with respect to the probability of detection, it is demonstrated that every minimax strategy has the same constant risk function. However, some connection between two basic modes of quantum statistical inference: minimax and Bayesian, is also presented. To this end, the equivalent condition for the existence of a simple (i.e. consisting of projections) Bayes strategy for pure states is obtained. For the connection, it is shown that every minimax strategy is Bayes for any least favourable prior. (Note that a minimax strategy is defined in the way independent of an *a priori* distribution on the set of the states considered.) Moreover, in dimension greater than two, it is possible to find pure states and a simple minimax strategy with constant risk function such that this strategy is Bayes with respect to a *non-uniform a priori* distribution being a least favourable prior, which is impossible in dimension two.

Funding No funding was received to assist with the preparation of this manuscript.

Declarations

Conflict of interest The author has no relevant financial or non-financial interests to disclose.

References

1. Barnett, S.M., Croke, S.: Quantum state discrimination. Adv. Opt. Photonics 1, 238–278 (2009); preprint, arXiv: 0810.1970
2. Bergou, J., Herzog, U., Hillery, M.: Discrimination of quantum states. In: Paris, M., Řeháček, J. (eds.) Quantum State Estimation. Lecture Notes Phys. 649, pp. 417–465. Springer, Berlin (2004)
3. Berta, M., Scholz, V.B., Tomamichel, M.: Rényi divergences as weighted non-commutative vector-valued L_p-spaces. Ann. Henri Poincaré **19**, 1843–1867 (2018)
4. Bourbaki, N.: Elements of Mathematics: Integration I. Springer, Berlin (2000)
5. D'Ariano, G.M., Sacchi, M.S., Kahn, J.: Minimax quantum-state discrimination. Phys. Rev. A **72**, 032310 (2005)
6. D'Ariano, G.M., Sacchi, M.S., Kahn, J.: Minimax discrimination of two Pauli channels. Phys. Rev. A **72**, 052302 (2005)
7. Chefles, A.: Quantum state discrimination. Contemporary Phys. **41**, 401–430 (2000)
8. Helstrom, C.W.: Quantum Detection and Estimation Theory. Academic Press, New York-San Francisco-London (1976)
9. Hirota, O., Ikehara, S.: Minimax strategy in quantum detection theory and its application to optical communications. IEEE Trans. IECE Jpn. E **65**(11), 627–633 (1982)
10. Holevo, A.S.: Statistical decision theory for quantum systems. J. Multivariate Anal. **3**, 337–394 (1973)
11. Holevo, A.S.: Isl'edovanija po obščej teorii statističeskich rešenij. Trudy Matematičeskogo Instituta imeni V.A. Steklova 124, Nauka, Moskva (1976); English transl. Investigations in the general theory of statistical decisions. Proceedings of the Steklov Institute of Mathematics 124 (issue 3) (1978)
12. Hollands, S.: Trace- and improved data processing inequalities for von Neumann algebras. Publ. RIMS Kyoto Univ. **59**, 687–729 (2023)
13. Ibragimov, I.A., Has'minskii, R.Z.: Statistical Estimation. Asymptotic Theory. Springer, New York (1981)
14. Jakšić, V., Ogata, Y., Pillet, C.-A., Seiringer, R.: Quantum hypothesis testing and non-equlibrium statistical mechanics. Rev. Math. Phys. **24**(6), 1230002 (2012)
15. Jenčová, A.: Rényi relative entropies and noncommutative L_p-spaces. Ann. Henri Poincaré **19**, 2513–2542 (2018)
16. Jenčová, A., Petz, D.: Sufficiency in quantum statistical inference. Comm. Math. Phys. **263**, 259–276 (2006)
17. Kennedy, R.S.: On the optimum receiver for the M-ary linearly independent pure state problem. M.I.T. Res. Lab. Electron. Quart. Prog. Rep. 110, 142–146 (1973)
18. Kennedy, R.S.: Uniqueness of the optimum receiver for the M-ary pure state problem, M.I.T. Res. Lab. Electron. Quart. Prog. Rep.113, 129–130 (1974)
19. Łuczak, A.: Maximizing the probability of detection for pure states. J. Math. Phys. **50**, 053502 (2009)
20. Nakahira, K., Kato, K., Usuda, T.S.: Generalized quantum state discrimination problems. Phys. Rev. A **91**, 052304 (2015)
21. Osaki, M., Ban, M., Hirota, O.: Derivation and physical interpretation of the optimum detection operators for coherent-state signals. Phys. Rev. A **54**, 1691–1701 (1996)
22. Ozawa, M.: Optimal measurements for general quantum systems. Rep. Math. Phys. **18**, 11–28 (1980)
23. Quadeer, M., Tomamichel, M., Ferrie, C.: Minimax quantum state estimation under Bregman divergence. Quantum **3**, 126 (2019)
24. Sion, M.: On general minimax theorems. Pacific J. Math. **8**(1), 171–176 (1958)
25. Tanaka, F.: Quantum minimax theorem. preprint; arXiv: 1410.3639