

Wootters' distance revisited: a new distinguishability criterium

A. Majtey¹, P.W. Lamberti^{1,a}, M.T. Martin², and A. Plastino²

¹ Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba, Ciudad Universitaria, 5000 Córdoba, Argentina

² Instituto de Física (IFLP), Universidad Nacional de La Plata and CONICET, C.C. 727, La Plata 1900, Argentina

Received 3 August 2004 / Received in final form 11 November 2004

Published online 25 January 2005 – © EDP Sciences, Società Italiana di Fisica, Springer-Verlag 2005

Abstract. The notion of distinguishability between quantum states has shown to be fundamental in the frame of quantum information theory. In this paper we present a new distinguishability criterium by using an information theoretic quantity: the Jensen-Shannon divergence (JSD). This quantity has several interesting properties, both from a conceptual and a formal point of view. Previous to define this distinguishability criterium, we review some of the most frequently used distances defined over quantum mechanics' Hilbert space. In this point our main claim is that the JSD can be taken as a unifying distance between quantum states.

PACS. 02.50.-r Probability theory, stochastic processes, and statistics – 03.65.-w Quantum mechanics – 89.70.+c Information theory and communication theory

1 Introduction

The problem of measurement is an issue of central importance in quantum theory, that, since the pioneering days of the twenties has given rise to controversies [1]. Many of the most astonishing results of quantum mechanics are related to the particular properties of the measurement processes. In recent years, the unique character of quantum measurement has led to a new field of research: quantum information technology [2]. From a formal point of view, a measurement in quantum theory is described by means of an Hermitian operator. If the eigenstates of this operator are $|\phi_k\rangle$ and the state of the system to be measured is $|\Psi\rangle = \sum c_k |\phi_k\rangle$, then, according to the axioms of the quantum theory, the result of the measurement will, with probability $|c_k|^2$, be the corresponding eigenvalue a_k , represented physically by an appropriate state of the measuring device \mathcal{A} .

A close related theme is that of the distinguishability between states, that is, just how can we discern between two states $|\Psi^{(1)}\rangle$ and $|\Psi^{(2)}\rangle$ of a given physical system by using the measuring device \mathcal{A} . In a seminal paper, Wootters investigated this problem and introduced a “distinguishability-distance” between pure states in the associated Hilbert space [4]. Braunstein and Caves extended this distance to density operators for mixed states [5]. Wootters distinguishability-criterium can be established, within the framework of probability theory (in-

dependently of any quantum interpretation), in the following way [4]: two probability distributions, say, $p^{(1)} = (p_1, p_2, \dots, p_N)$ and $p^{(2)} = (q_1, q_2, \dots, q_N)$ are distinguishable after L trials ($L \rightarrow \infty$) if and only if the condition

$$\frac{\sqrt{L}}{2} \left\{ \sum_{i=1}^N \frac{(\delta p_i)^2}{p_i} \right\}^{1/2} > 1 \quad (1)$$

with $\delta p_i = p_i - q_i$, is satisfied. This distinguishability-criterium involves a distance defined over the space of probability distributions

$$ds(p^{(1)}, p^{(2)}) = \frac{1}{2} \sqrt{\sum_i \frac{(\delta p_i)^2}{p_i}}. \quad (2)$$

Statisticians call to the square of this form the χ^2 distance. Wootters maps this distance into the associated Hilbert space and establishes a correspondence with the usual notion of distance between states in Hilbert's space.

In addition to its relevance with regards to the distinguishability issue, the concept of distance between different states in a Hilbert space plays an important role in a diversity of circumstances

- the study of the geometric properties of the quantum evolution sub-manifold [6, 7],

^a e-mail: lamberti@fis.uncor.edu

- in discussing squeezed coherent states or generalized coherent spin states [8],
- in ascertaining the quality of approximate treatments [9].

It has recently been recognized that the concept of distinguishability is basic to manipulate information in the sense that being able to discern between different physical states of a given system allows one to determine just how much information can be encoded into that system, so that the notion of distinguishability builds a bridge between quantum theory and information theory [3].

In this work, we will try to strengthen this connection by investigating the relation between Wootters' distance and a suitable metric for the probability-distributions' space that is used in information theory: the Jensen-Shannon divergence (JSD). Recently, the JSD has been exhaustively studied in different contexts [10]. It has many interesting interpretations, both in the framework of information theory as in the context of mathematical statistics. One of its basic properties is that its square root is a true metric in the probability-distributions' space, i.e., its square root is a distance that verifies the triangle inequality [11]. This fact is quite relevant, since metric properties are crucial for the application of many important convergence theorems that one needs when iterative algorithms are studied.

The purpose of this paper is twofold:

1. first, we pursue a pedagogical objective by reviewing some distances and metrics commonly used in quantum theory. Even though many of the results presented here are known, they are not always presented from an unified perspective, at least in physics literature,
2. second, we formulate a distinguishability criterium for quantum mechanics based on the JSD.

Finally, some conclusions are drawn.

2 A primer on Hilbert space distances

Let $|\phi_1\rangle, \dots, |\phi_N\rangle$ be the eigenstates of a given Hermitian operator associated with the measuring instrument \mathcal{A} . For simplicity's sake we assume that no degeneration exists. Thus, in a given measurement N possible results may ensue. If we have prepared the system in the (normalized) state $|\Psi^{(1)}\rangle$, each of these results can be found with probability $|\langle\phi_i|\Psi^{(1)}\rangle|^2$. If we prepare it, instead, in the state $|\Psi^{(2)}\rangle$, this probability is $|\langle\phi_i|\Psi^{(2)}\rangle|^2$. Since the basis $|\phi_i\rangle$ is complete

$$\sum_i |\phi_i\rangle\langle\phi_i| = I, \tag{3}$$

one has

$$\sum_i |\langle\phi_i|\Psi^{(1)}\rangle|^2 = \sum_i |\langle\phi_i|\Psi^{(2)}\rangle|^2 = 1. \tag{4}$$

Let us write

$$p_i^{(1)} = |\langle\phi_i|\Psi^{(1)}\rangle|^2, \quad p_i^{(2)} = |\langle\phi_i|\Psi^{(2)}\rangle|^2. \tag{5}$$

An alternative way of looking at things is as follows. Let

$$X_N^+ = \left\{ (p_1, \dots, p_N); 0 \leq p_i \leq 1; \sum_i p_i = 1 \right\} \tag{6}$$

be the set of discrete probability distributions (generalization to continuous ones being straightforward) and let \mathcal{S} be the set of normalized states in the Hilbert space \mathcal{H}^{n+1} , $n + 1 = N$. To each state $|\Psi\rangle$ in \mathcal{S} (indeed to a ray $\lambda|\Psi\rangle$, $\lambda = e^{i\varphi}$) we assign an element $\{p_i\}$ of X_N^+ through the application $\mathcal{F}_\mathcal{A}$ given by:

$$\mathcal{F}_\mathcal{A} : \mathcal{S} \subset \mathcal{H}^{n+1} \rightarrow X_N^+ \\ |\Psi\rangle \rightarrow \{p_i\} \quad \text{such that } p_i = |\langle\phi_i|\Psi\rangle|^2. \tag{7}$$

Obviously, the application $\mathcal{F}_\mathcal{A}$ is consistent with expressions (4) and (5).

Let $s_X(p^{(1)}, p^{(2)})$ be a distance defined on the space of probability distributions X_N^+ , that is, an application from $X_N^+ \times X_N^+$ into \mathfrak{R} such that is symmetric and $s_X(p^{(1)}, p^{(2)}) = 0$ if and only if $p^{(1)} = p^{(2)}$. One can associate to $s_X(p^{(1)}, p^{(2)})$ a distance in the space \mathcal{H}^{n+1} , $s_{\mathcal{H}}^{\mathcal{A}}(|\Psi^{(1)}\rangle, |\Psi^{(2)}\rangle)$ through the application $\mathcal{F}_\mathcal{A}$. Let us note that this distance depends upon the measuring instrument \mathcal{A} . Our objective is to find a representative distance of $s_X(p^{(1)}, p^{(2)})$ in Hilbert's space independently of the basis $|\phi_k\rangle$. This will be attained by looking for the maximum of the associated distance $s_{\mathcal{H}}^{\mathcal{A}}$. We discuss some examples below. The pertinent distances are given proper names (e.g., Wootters), according to common usage.

Notation remark. We will use the following notation: s_X denotes a distance defined over X_N^+ ; $s_{\mathcal{H}}^{\mathcal{A}}$ denotes the corresponding distance over \mathcal{H}^{n+1} obtained from the correspondence induced by application $\mathcal{F}_\mathcal{A}$; $S_{\mathcal{H}}$ denotes the maximum of $s_{\mathcal{H}}^{\mathcal{A}}$.

2.1 Wootters distance

The Wootters distance between two probability distributions, $p^{(1)}$ and $p^{(2)}$ is defined as

$$s_X^W(p^{(1)}, p^{(2)}) = \arccos \left(\sum_i \sqrt{p_i^{(1)} p_i^{(2)}} \right). \tag{8}$$

When $p^{(1)} \rightarrow p^{(2)}$, the form (2) is re-obtained.

By using the correspondence (7), we can write

$$s_{\mathcal{H}}^{W, \mathcal{A}}(|\Psi^{(1)}\rangle, |\Psi^{(2)}\rangle) = \arccos \left(\sum_i |\langle\phi_i|\Psi^{(1)}\rangle| |\langle\phi_i|\Psi^{(2)}\rangle| \right). \tag{9}$$

Note that $\arccos(x)$ decreases in $[0, 1]$. Also, the following inequality

$$\sum_i \left| \langle \phi_i | \Psi^{(1)} \rangle \right| \left| \langle \phi_i | \Psi^{(2)} \rangle \right| \geq \left| \langle \Psi^{(1)} | \Psi^{(2)} \rangle \right|, \quad (10)$$

is true for all $\{|\phi_i\rangle\}$. Indeed, assume $|\Psi^{(1)}\rangle = \sum_k a_k |\phi_k\rangle$, and $|\Psi^{(2)}\rangle = \sum_k b_k |\phi_k\rangle$. Then,

$$\begin{aligned} \left| \langle \Psi^{(1)} | \Psi^{(2)} \rangle \right| &= \left| \sum_k a_k b_k^* \right| \leq \sum_k |a_k b_k^*| \\ &\leq \sum_k \left| \langle \phi_k | \Psi^{(1)} \rangle \right| \left| \langle \phi_k | \Psi^{(2)} \rangle \right|. \end{aligned} \quad (11)$$

Inequality (11), together with the \arccos -function decreasing nature, imply that the distance

$$S_{\mathcal{H}}^W \left(|\Psi^{(1)}\rangle, |\Psi^{(2)}\rangle \right) = \arccos \left(\left| \langle \Psi^{(1)} | \Psi^{(2)} \rangle \right| \right), \quad (12)$$

maximizes $s_{\mathcal{H}}^{W,A}$. In this way we arrive at the distance associated to the Wootters' one in Hilbert's space. Geometrically, it gives the angle between the two states (rays) $|\Psi^{(1)}\rangle$ and $|\Psi^{(2)}\rangle$.

2.2 Hellinger' distance

Let s_X^H be a distance in X_N^+ such that its square reads

$$(s_X^H)^2 \left(p^{(1)}, p^{(2)} \right) = \frac{1}{2} \sum_i \left| \sqrt{p_i^{(1)}} - \sqrt{p_i^{(2)}} \right|^2. \quad (13)$$

Its \mathcal{H}^{n+1} -counterpart $s_{\mathcal{H}}^{H,A}$ satisfies

$$(s_{\mathcal{H}}^{H,A})^2 = \frac{1}{2} \sum_i \left\{ \left| \langle \phi_i | \Psi^{(1)} \rangle \right| - \left| \langle \phi_i | \Psi^{(2)} \rangle \right| \right\}^2, \quad (14)$$

that can be cast as

$$1 - \sum_i \left| \langle \phi_i | \Psi^{(1)} \rangle \right| \left| \langle \phi_i | \Psi^{(2)} \rangle \right|. \quad (15)$$

We see that, according to the inequality (10), the distance

$$(S_{\mathcal{H}}^H)^2 \left(|\Psi^{(1)}\rangle, |\Psi^{(2)}\rangle \right) = 1 - \left| \langle \Psi^{(1)} | \Psi^{(2)} \rangle \right|, \quad (16)$$

is the maximum of the associated distance s_X^H . It is known as Hellinger-distance and it represents the sine of the half angle between the two Hilbert space vectors $|\Psi^{(1)}\rangle$ and $|\Psi^{(2)}\rangle$ [12].

2.3 Bhattacharyya' distance

Another distinguishability measure arises from Bhattacharyya coefficients. For two probability distributions $p^{(1)}$ and $p^{(2)}$, the Bhattacharyya coefficients are defined by [13]

$$B \left(p^{(1)}, p^{(2)} \right) = \sum_i \sqrt{p_i^{(1)}} \sqrt{p_i^{(2)}}. \quad (17)$$

Out of these coefficients we can define a distance between probability distributions:

$$s_X^B \left(p^{(1)}, p^{(2)} \right) = -\ln \left(B \left(p^{(1)}, p^{(2)} \right) \right). \quad (18)$$

Note that the Wootters' distance can be also expressed in terms of the coefficients $B(p^{(1)}, p^{(2)})$ as $s_X^W(p^{(1)}, p^{(2)}) = \arccos(B(p^{(1)}, p^{(2)}))$. It is worth mentioning that neither Wootters' nor the distance (18) are metrics because they do not verify the triangle inequality.

The associated distance to (18) in Hilbert's space is

$$s_{\mathcal{H}}^{B,A} = -\ln \sum_i \left| \langle \phi_i | \Psi^{(1)} \rangle \right| \left| \langle \phi_i | \Psi^{(2)} \rangle \right|. \quad (19)$$

Now, since the function $-\ln(x)$ decreases with x , on the basis of (10) we gather that

$$S_{\mathcal{H}}^B \left(|\Psi^{(1)}\rangle, |\Psi^{(2)}\rangle \right) = -\ln \left| \langle \Psi^{(1)} | \Psi^{(2)} \rangle \right|, \quad (20)$$

is the maximum of Bhattacharyya's distance.

In these examples we focused attention upon the maximums. Also, we have been able to cast all these distances as a function of a Riemannian Hilbert-space metric: an "angle" between rays, the only one that remains invariant under the action of the time-evolution unitary operator.

2.4 Fubini-Study's metric

Let us recall that the Hilbert space \mathcal{H}^{n+1} is isomorphic to the n -dimensional complex projective space \mathcal{P}^n , that is, the quotient space

$$\mathcal{P}^n = (C^{n+1} - \{0\}) / \sim. \quad (21)$$

with \sim the equivalence relation given by

$$|\psi\rangle \sim |\phi\rangle \text{ iff } \exists \lambda \in C - 0 \text{ such that } |\psi\rangle = \lambda |\phi\rangle. \quad (22)$$

In this example we start with a \mathcal{H}^{n+1} -distance and construct one in X_N^+ (previously we proceeded in reverse fashion). In \mathcal{P}^n one defines the Fubini-Study metric θ_{FS} according to

$$\cos^2 \left(\frac{\theta_{FS}}{2} \right) \equiv \frac{\langle \psi | \eta \rangle \langle \eta | \psi \rangle}{\langle \psi | \psi \rangle \langle \eta | \eta \rangle}. \quad (23)$$

For $|\psi\rangle \sim |\phi\rangle$, one has $\theta_{FS} = 0$. Maximum separation between two states is attained for $\theta_{FS} = \pi$. Let (i) $\mathcal{S} \subset \mathcal{P}^n$ be the set of normalized states in \mathcal{P}^n while (ii) $|\psi\rangle$ and $|\psi\rangle + |d\psi\rangle$ are two very close states in \mathcal{S} . Normalization implies

$$2\text{Re}(\langle \psi | d\psi \rangle) = -\langle d\psi | d\psi \rangle. \quad (24)$$

From (23), by putting $|\eta\rangle = |\psi\rangle + |d\psi\rangle$, we can evaluate the Fubini-Study distance between two infinitely close states:

$$\cos^2 \left(\frac{d\theta_{FS}}{2} \right) \simeq \left(1 - \frac{1}{2!} \left(\frac{d\theta_{FS}^2}{2} \right) + \dots \right)^2 \simeq 1 - \frac{(d\theta_{FS}^2)}{4}, \quad (25)$$

so that

$$d\theta_{FS}^2 = 4(\langle d\psi|d\psi\rangle - |\langle\psi|d\psi\rangle|^2). \quad (26)$$

If $|d\psi_\perp\rangle \equiv |d\psi\rangle - |\psi\rangle\langle\psi|d\psi\rangle$ is the orthogonal projection onto $|\psi\rangle$ of $|d\psi\rangle$, the Fubini-Study metrics acquires the aspect [5]

$$d\theta_{FS}^2 = 4\langle d\psi_\perp|d\psi_\perp\rangle. \quad (27)$$

An alternative approach to the Fubini-Study metric can be found in reference [6].

Assume now the following expansions for $|\psi\rangle$ and $|\eta\rangle = |\psi\rangle + |d\psi\rangle$:

$$|\psi\rangle = \sum_i \sqrt{p_i} |\phi_i\rangle, \quad |\eta\rangle = \sum_i \sqrt{p_i + dp_i} |\phi_i\rangle, \quad (28)$$

noticing that one might add appropriate phases in both equations. These phases, however, can be eliminated by a proper basis-transformation (see Ref. [5]). The Fubini-Study distance between these states, up to second order in dp_i becomes

$$d\theta_{FS}^2(|\psi\rangle, |\eta\rangle) = \frac{1}{4} \sum_i \frac{dp_i^2}{p_i} \quad (29)$$

which can be thought as the corresponding Fubini-Study metric between the distributions $\{p_i\}$ and $\{p_i + dp_i\}$ over the space X_N^+ .

3 Jensen-Shannon divergence

Information theoretic measures allow one to build up quantitative entropic divergences between two probability distributions. A common entropic measure is the Kullback-Leibler divergence:

$$s_X^K(p^{(1)}, p^{(2)}) = \sum_i p_i^{(1)} \ln \frac{p_i^{(1)}}{p_i^{(2)}}. \quad (30)$$

This distance, however, is (i) not symmetric, (ii) unbounded, and (iii) not always well defined. To overcome these limitations Rao and Lin introduced a symmetrized version of the Kullback-Leibler divergence the Jensen-Shannon divergence (JSD), which is defined as

$$s_X^{JS}(p^{(1)}, p^{(2)}) = H\left(\frac{p^{(1)} + p^{(2)}}{2}\right) - \frac{1}{2}H(p^{(1)}) - \frac{1}{2}H(p^{(2)}), \quad (31)$$

where $H(p) = -\sum_i p_i \ln p_i$ stands for Shannon's entropy [14,15].

The minimum of the JSD occurs at $p^{(1)} = p^{(2)}$ and its maximum is reached when $p^{(1)}$ and $p^{(2)}$ are two distinct deterministic distributions. In this case $s_X^{JS} = \ln 2$. As it was mentioned previously, one of the JSD main properties is that of being the square of a metric. A proof of this fact can be found in reference [11]. Alternatively, this can be proved starting from some classical results of harmonic analysis due to Schoenberg [16,17]. The basic property of the JSD that makes Schoenberg theorem

applicable is that s_X^{JS} is a definite negative kernel, that is, for all finite collection of real number $(\zeta_i)_{i \leq N}$ and for all corresponding finite sets $(x_i)_{i \leq N}$ of points in X_N^+ , the implication

$$\sum_i \zeta_i = 0 \Rightarrow \sum_{i,j} \zeta_i \zeta_j s_X^{JS}(x_i, x_j) \leq 0 \quad (32)$$

is valid [18].

Another consequence of Schoenberg's theorems is that the metric space $(X_N^+, \sqrt{s_X^{JS}})$ can be isometrically mapped into a subset of a Hilbert space. This result establishes a connection between information theory and differential geometry [19], which could have interesting consequences in the realm of quantum information theory.

Consider once again the states $|\psi\rangle$ and $|\eta\rangle$ given by (28) in order to evaluate the JSD between the concomitant probability distributions $p^{(1)}(|\psi\rangle)$, $p^{(2)}(|\eta\rangle)$. By doing so we are evaluating the associated distance in Hilbert's space $s_{\mathcal{H}}^{JS,A}$ between the states $|\psi\rangle$ and $|\eta\rangle$. Expanding the pertinent JSD in dp_i -terms, one easily ascertains that the first non-vanishing contributions are the quadratic ones

$$ds_{\mathcal{H}}^{JS,A}(|\psi\rangle, |\eta\rangle) = \frac{1}{8} \sum_i \frac{dp_i^2}{p_i}, \quad (33)$$

which coincides with (a half of) the Fubini-Study (29) instance up to this order in dp_i . Up to same order a similar relation exists between the JSD and both the Wootters' and the Bhattacharyya' distances, that is

$$ds_{\mathcal{H}}^{JS,A} = \frac{1}{2} \left(ds_{\mathcal{H}}^{W,A} \right)^2 = \frac{1}{2} \left(ds_{\mathcal{H}}^{B,A} \right)^2, \quad (34)$$

which can be easily checked by inspection. Incidentally, it is worth mentioning that, when we have a continuous probability distribution $p(x)$, the JSD between $p(x)$ and its shifted version $p(x + \delta)$ is related to the Fisher information measure I through the expression

$$s_X^{JS}(p(x), p(x + \delta)) \simeq \frac{\delta}{2} \sqrt{\frac{I}{2}} \quad (35)$$

with

$$I[p(x)] = \int \frac{[dp(x)/dx]^2}{p(x)} dx. \quad (36)$$

Equations (34) have been established up to second order in dp_i . Let us proceed to higher orders. To do this let us consider a binary system (a generalization to a system with a greater number of states is straightforward). Let $p^{(1)} = (p, q)$ and $p^{(2)} = (p + dp, q - dp)$ with $p + q = 1$ two neighboring probability distributions and evaluate the pertinent JSD up to order dp^4 . We get

$$ds_X^{JS} = -\frac{1}{8} \frac{1}{(p-1)p} dp^2 + \frac{1}{16} \frac{2p-1}{p^2(p-1)^2} dp^3 - \frac{7}{192} \frac{3p^2-3p+1}{p^3(p-1)^3} dp^4 + o(dp^5). \quad (37)$$

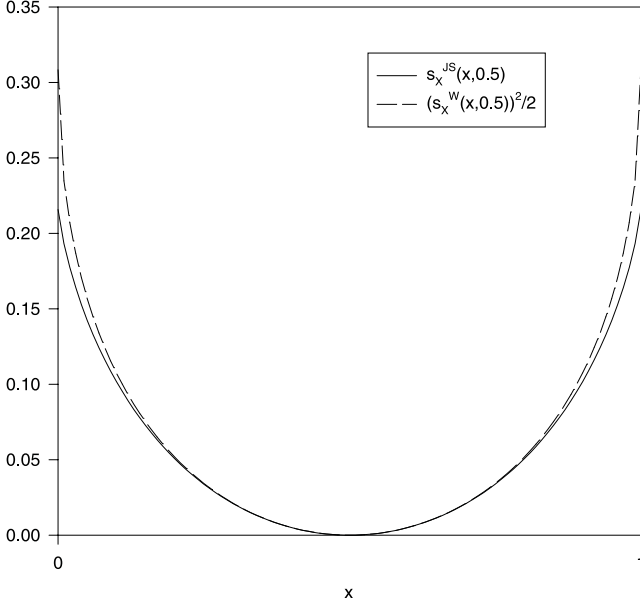


Fig. 1. Plots of s_X^{JS} and $(s_X^W)^2/2$. See text for details.

In turn, the corresponding Wootters' distance squared, up to the same order is

$$\begin{aligned} \frac{1}{2}(ds_X^W)^2 = & -\frac{1}{8} \frac{1}{(p-1)p} dp^2 + \frac{1}{16} \frac{2p-1}{p^2(p-1)^2} dp^3 \\ & - \frac{1}{384} \frac{44p^2 - 44p + 15}{p^3(p-1)^3} dp^4 + o(dp^5). \end{aligned} \quad (38)$$

We detect coincidence between (37) and (38) up to order dp^3 . The fourth order difference equals $1/192$. In other words, the relation

$$ds_X^{JS} = \frac{1}{2}(ds_X^W)^2 \quad (39)$$

can be established up to third order in dp . Figure 1 shows how s_X^{JS} and $(s_X^W)^2/2$ approach one to each other for $p^{(1)} \approx p^{(2)}$. We took $p^{(1)} = (a, 1-a)$ and $p^{(2)} = (b, 1-b)$ and evaluated the corresponding distances as a function of b by fixing $a = 0.5$.

Going back to Wootters' distinguishability criterium (1), with equation (39) in mind, we are in a position to enunciate an alternative criterium: two probability distributions $P^{(1)}$ and $P^{(2)}$ are distinguishable after L trials ($L \rightarrow \infty$) if and only if

$$\left(s_X^{JS}(P^{(1)}, P^{(2)}) \right)^{1/2} > \frac{1}{\sqrt{2L}}. \quad (40)$$

There exist formal arguments in favor of this last statement, namely (i) $(s_X^{JS})^{1/2}$ is a true metric for the space X_N^+ and (ii) this criterium is established in terms of an information theoretic quantity, the JSD. Obviously inequality (40) is equivalent to inequality (1) for two distributions "close" enough.

In the context of Section 2 the following question emerges: what metric is the representative of s_X^{JS} in

Hilbert's space \mathcal{H}^{n+1} ? Equivalently: what is the maximum of the metric $s_{\mathcal{H}}^{JS, \mathcal{A}}$? In this case it is difficult (or impossible) to obtain an analytical expression for both metrics, $s_{\mathcal{H}}^{JS, \mathcal{A}}$ and its upper bound $S_{\mathcal{H}}^{JS}$. Anyway, it is possible to deduce an upper bound for $s_{\mathcal{H}}^{JS}$. Let us consider a Hilbert space of dimension $2D$ and let $|\Psi^{(1)}\rangle$ and $|\Psi^{(2)}\rangle$ be two arbitrary, normalized states (the extension to a greater number of dimensions is straightforward). We set $|\langle \Psi^{(1)} | \Psi^{(2)} \rangle| = \cos \varphi$ for $\varphi \in [0, \pi/2]$, that is, φ is the Wootters distance between $|\Psi^{(1)}\rangle$ and $|\Psi^{(2)}\rangle$.

Let $\{|\phi_i\rangle\}_{i=1}^2$ be an orthonormal basis for \mathcal{H}^2 . Any other orthonormal basis $\{|\tilde{\phi}_i\rangle\}_{i=1}^2$ can be related to $\{|\phi_i\rangle\}$ via the rotation

$$\begin{aligned} |\tilde{\phi}_1(\theta)\rangle &= \frac{e^{i\theta}}{\sqrt{2}}|\phi_1\rangle + \frac{e^{-i\theta}}{\sqrt{2}}|\phi_2\rangle \\ |\tilde{\phi}_2(\theta)\rangle &= -\frac{e^{i\theta}}{\sqrt{2}}|\phi_1\rangle + \frac{e^{-i\theta}}{\sqrt{2}}|\phi_2\rangle \end{aligned} \quad (41)$$

with $\theta \in [0, 2\pi]$. We set $p_i^{(j)} \equiv |\langle \phi_i | \Psi^{(j)} \rangle|^2$ and $\tilde{p}_i^{(j)}(\theta) \equiv |\langle \tilde{\phi}_i(\theta) | \Psi^{(j)} \rangle|^2$. Also, $\langle \phi_i | \Psi^{(j)} \rangle = \sqrt{p_i^{(j)}} e^{i\alpha_i^{(j)}}$ (via application of (7)). A little algebra then leads to

$$\tilde{p}_1^{(1)}(\theta) = \frac{p_1^{(1)} + p_2^{(1)}}{2} + \sqrt{p_1^{(1)} p_2^{(1)}} \cos(2\theta + \alpha_2^{(1)} - \alpha_1^{(1)}), \quad (42)$$

and

$$\tilde{p}_1^{(2)}(\theta) = \frac{p_1^{(2)} + p_2^{(2)}}{2} + \sqrt{p_1^{(2)} p_2^{(2)}} \cos(2\theta + \alpha_1^{(2)} - \alpha_2^{(2)}) \quad (43)$$

with $\alpha_i^{(j)}$ are real numbers. Moreover, $\tilde{p}_2^{(1)} = 1 - \tilde{p}_1^{(1)}$ and $\tilde{p}_2^{(2)} = 1 - \tilde{p}_1^{(2)}$. Without loss of generality we can take $|\phi_1\rangle = |\Psi^{(1)}\rangle$, so that $p_1^{(1)} = 1$, $\alpha_1^{(1)} = 0$, $p_2^{(1)} = 0$, $\sqrt{p_1^{(2)}} = \cos \varphi$ and $\sqrt{p_2^{(2)}} = \sin \varphi$. Thus, we can compute $\sqrt{2s_{\mathcal{H}}^{JS, \tilde{\phi}}(\tilde{p}^{(2)}, \tilde{p}^{(1)})}$ as a function of θ . Figure 2 plots such a function for different φ -values. Figure 3 depicts a 3D-plot of $\sqrt{2s_{\mathcal{H}}^{JS, \tilde{\phi}}}$ as a function of θ and φ . In both cases we put $\alpha_2^{(1)} = \alpha_1^{(2)} = \alpha_2^{(2)} = 0$.

Out of these figures we conclude that Wootters' distance (φ) is an upper bound to $\sqrt{2s_{\mathcal{H}}^{JS, \tilde{\phi}}(\tilde{p}^{(2)}, \tilde{p}^{(1)})}$. For $\varphi \rightarrow 0$, both quantities tend to coincide. In other words, we can state the inequalities

$$\begin{aligned} s_{\mathcal{H}}^W(|\Psi^{(1)}\rangle, |\Psi^{(2)}\rangle) &\geq s_{\mathcal{H}}^{W, \mathcal{A}}(|\Psi^{(1)}\rangle, |\Psi^{(2)}\rangle) \\ &\geq \sqrt{2s_{\mathcal{H}}^{JS, \mathcal{A}}(|\Psi^{(1)}\rangle, |\Psi^{(2)}\rangle)} \end{aligned} \quad (44)$$

for any measure device \mathcal{A} .

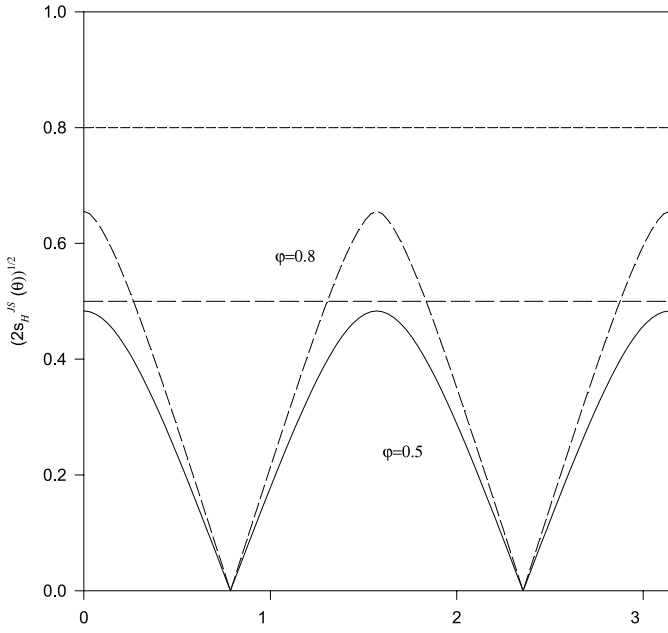


Fig. 2. $\sqrt{2s_{\mathcal{H}}^{JS, \tilde{\phi}}(\tilde{p}^{(2)}, \tilde{p}^{(1)})}$ as a function of θ for $\varphi = 0.5$ and $\varphi = 0.8$.

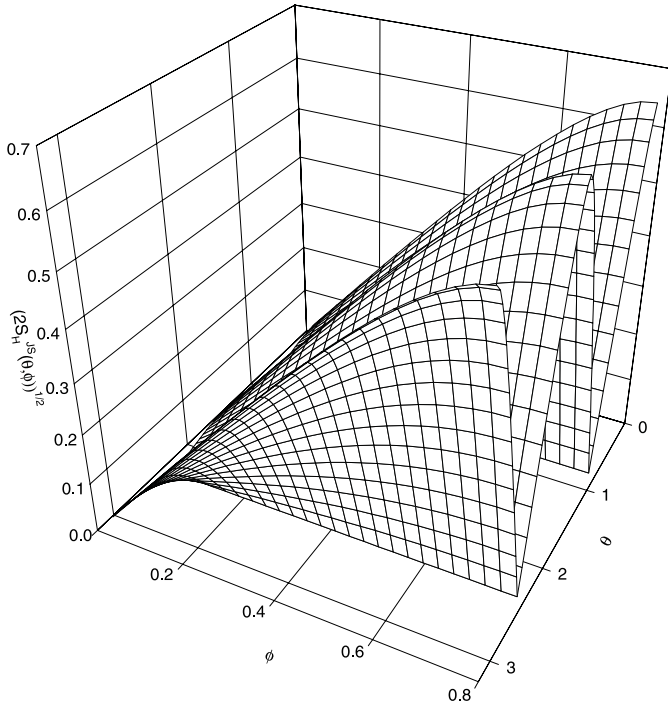


Fig. 3. 3D Plot of $\sqrt{2s_{\mathcal{H}}^{JS, \tilde{\phi}}(\tilde{p}^{(2)}, \tilde{p}^{(1)})}$ as a function of θ and φ . One clearly appreciates the bound in the plane $z = \varphi$.

Inequalities (44) allow us to conclude that $S_{\mathcal{H}}^W(|\Psi^{(1)}\rangle, |\Psi^{(2)}\rangle)$ “represents” (as the maximum, that is as the lowest upper bound) to $\sqrt{2s_{\mathcal{H}}^{JS, A}}$ in the

Hilbert space. Furthermore two states distinguishable under the “Jensen-Shannon criterium” are obviously distinguishable under the Wootters’ ones.

4 Conclusions

We have proposed an alternative distinguishability criterium for quantum states. This distinguishability criterium is established in terms of an information theoretical quantity: the JSD, that exhibits many interesting properties, such as a metric character and its boundedness. This provides for a better formal context. In some sense we feel that the JSD could be taken as a unified measure of distinguishability in the framework of quantum information theory.

In the present work we focused on the case of pure states. An extension to mixed states can be easily attained. In fact, by replacing in equation (31) the Shannon entropy by the von Neumann entropy, $H_N(\rho) = -\text{Tr}(\rho \ln \rho)$, we can evaluate the JSD between two states described by the density operators ρ_1 and ρ_2 :

$$S_{\mathcal{H}}^{JS}(\rho_1, \rho_2) = H_N\left(\frac{\rho_1 + \rho_2}{2}\right) - \frac{1}{2}H_N(\rho_1) - \frac{1}{2}H_N(\rho_2). \quad (45)$$

Remarkably, this quantity is always well defined unlike the corresponding Kullback-Leibler divergence that requires that the support of ρ_1 is equal to or larger than that of ρ_2 [20]. A more detailed study of the properties of JSD for mixed states will be presented elsewhere.

Finally it is worth to mention that the JSD can be also interpreted in a Bayesian probabilistic sense. In fact, the JSD gives both lower and upper bounds to Bayes’ probability error. Therefore, it deserves careful scrutiny in the light of some alternative quantum descriptions [21].

We are grateful to Secretaria de Ciencia y Tecnica de la Universidad Nacional de Córdoba for financial assistance. AM is a fellowship holder of SECYT-UNC and PWL and AP are members of CONICET. This work was partially supported by Grant BIO2002-04014-C03-03 from Spanish Government.

References

1. A. Wheeler, W. Zurek, *Quantum Theory and Measurement* (Princeton University Press, Princeton, 1983)
2. N. Gisin, G. Ribordy, W. Tittel, H. Zbinden, *Rev. Mod. Phys.* **74**, 145 (2002)
3. V. Vedral, *Rev. Mod. Phys.* **74**, 197 (2002)
4. W. Wootters, *Phys. Rev. D* **23**, 357 (1981)
5. S. Braunstein, C. Caves, *Phys. Rev. Lett.* **72**, 3439 (1994)
6. J. Anandan, Y. Aharonov, *Phys. Rev. Lett.* **65**, 1697 (1990)
7. S. Abe, *Phys. Rev. A* **48**, 4102 (1993)

8. L.C. Kwek, Ch. Oh, Xiang-Bin Wang, *J. Phys. A* **32**, 6613 (1999)
9. M. Ravicule, M. Casas, A. Plastino, *Phys. Rev. A* **55**, 1695 (1997)
10. I. Grosse, P. Bernaola-Galvan, P. Carpena, R. Roman Roldan, J. Oliver, H.E. Stanley, *Phys. Rev. E* **65**, 041905 (2002)
11. D. Endres, J. Schindelin, *IEEE Trans. Inf. Theory* **49**, 1858 (2003)
12. D. Brady, L. Hughston, *The Geometric Universe, Science, Geometry and the Work of R. Penrose*, edited by S.A. Huggett, L. Mason, K. Tod, S.-T. Tsou, N. Woodhouse (Oxford University Press, 1998)
13. A. Bhattacharyya, *Bull. Calcutta Math. Soc.* **35**, 99 (1943)
14. C. Rao, *IMS-Lect. Notes* **10**, 217 (1987)
15. J. Lin, *IEEE Trans. Inf. Theory* **37**, 145 (1991)
16. I.J. Schoenberg, *Trans. Am. Math. Soc.* **44**, 3 (1938)
17. C. Berg, J. Christensen, P. Ressel, *Harmonic Analysis on Semigroups* (Springer-Verlag, New York, 1984)
18. F. Topsøe, "Inequalities for the Jensen-Shannon Divergence", draft available at <http://www.math.ku.dk/topsoe/>
19. B. Fuglede, F. Topsøe, "Jensen-Shannon Divergence and Hilbert space embedding", Draft available at <http://www.math.ku.dk/topsoe/>
20. G. Lindblad, *Comm. Math. Phys.* **33**, 305 (1973)
21. C. Caves, C. Fuchs, R. Schack, *Phys. Rev. A* **65**, 022305 (2002)