Bhattacharyya statistical divergence of quantum observables

V. Majerník

Institute of Mathematics, Slovak Academy of Sciences, Bratislava, Štefánikova 47, Slovak Republic

S. Shpyrko

Department of Optics, Faculty of Science, Palacký University, Tr. 17. listopadu 50, CZ-77207 Olomouc, Czech Republic and
Institute of Nuclear Research, Ukrainian Academy of Sciences, pr. Nauki 47 Kiev, Ukraine

In this article we exploit the Bhattacharyya statistical divergence to determine the similarity of probability distributions of quantum observables. After brief review of useful characteristics of the Bhattacharyya divergence we apply it to determine the similarity of probability distributions of two non-commuting observables. An explicit expression for the Bhattacharyya statistical divergence is found for the case of two observables which are the x- and z-components of the angular momentum of a spin-1/2 system. Finally, a note is given of application of the considered statistical divergence to the specific physical measurement.

PACS numbers: 03.65.Ta; 02.50.-r

Keywords: statistical divergence, probability systems, quantum observables

I. INTRODUCTION

One of the important problems in the probability theory is to find an appropriate measure of the difference or the statistical divergence of two probability distributions \( P \) and \( P' \). This measure quantifies the degree of the similarity between them. In the mathematical statistics the divergence of two probability distributions is introduced as follows: If \( [X, P] \) and \( [X, P'] \) are two probability spaces then the so-called (Csiszár’s) f-divergence of probability distributions \( P \) and \( P' \) is given as

\[
D_f(P; P') = \sum_{x \in X} P'(x) f \left( \frac{P(x)}{P'(x)} \right),
\]

where \( f(u) \) represents a convex function in the interval \((0, \infty)\) and strictly convex for \( u = 1 \) [1].

Among the existing divergence measures of two discrete probabilities, \( P \equiv [p_1, p_2, \ldots, p_n] \) and \( Q \equiv [q_1, q_2, \ldots, q_n] \), the Kullback-Leibler statistical divergence [3]

\[
D_K(P; Q) = \sum_{i=1}^{n} p_i \log \left( \frac{p_i}{q_i} \right),
\]

is perhaps best known and most widely used. This is why this measure has several desirable properties, such as nonnegativity and additivity, which are crucial in its applications. \( D(P; Q) \) is not symmetrical regarding the exchange of \( P \) and \( Q \). For \( D_K(P; Q) \), the inequality holds

\[
\sum_{i=1}^{n} p_i \log \left( \frac{p_i}{q_i} \right) \geq 0.
\]

The minimum of \( D_K(P; Q) \) is obtained iff \( p_i = q_i \) (see, e.g. [14]).

Apart of the Kullback-Leibler statistical divergence, a number of other divergence measures, depending on certain parameters, have been proposed and intensively studied by Rényi [3], Kapur [2], Kullback and Leiber [4], Havrda and Charvat [6], Tsallis [7], [8]. Some of them satisfy the convexity condition only for restricted values of the corresponding parameters.

*Electronic address: fyziemar(at)savba.sk
†Electronic address: serge_shp(at)yahoo.com
II. THE BHATTACHARYYA STATISTICAL DIVERGENCE

The Bhattacharyya statistical divergence of the discrete probability distributions

\[ P = p_1, p_2, \ldots, p_n; \quad Q = q_1, q_2, \ldots, q_2 \]

is defined as \[ S(P, Q) = \sum_{i=1}^{n} (p_i q_i)^{1/2}. \] (1)

This divergence has the following properties:
(i) It becomes its maximal value equal to 1 when the probability distributions \( P \) and \( Q \) are identical.
(ii) Its minimal value is zero when the components of \( P \) and \( Q \) do not overlap.
(iii) Its value lies in the interval \([0, 1]\) and expresses the degree how much the probability distributions of \( P \) and \( Q \) are similar.
(iv) \( S(P, Q) \) is symmetrical regarding the exchange of \( P \) and \( Q \).
(v) \( S(P, Q) \) satisfies the properties of nonnegativity, finiteness and boundedness.
(vi) It can be straightforwardly extended for more than two probability distributions [15].

The Bhattacharyya statistical divergence of \( P \) and \( Q \) has a simple geometrical interpretation. Consider the following vectors \( \hat{P} = \{ \sqrt{p_1}, \sqrt{p_2}, \ldots, \sqrt{p_n} \} \) and \( \hat{Q} = \{ \sqrt{q_1}, \sqrt{q_2}, \ldots, \sqrt{q_n} \} \). According to Eq. (1), the similarity measure of \( \hat{P} \) and \( \hat{Q} \) is simply the scalar product of \( \hat{P} \) and \( \hat{Q} \) in \( \mathbb{R}^{(m)}_{+} \). Since \( \hat{P} \) and \( \hat{Q} \) represent the unit vectors in \( \mathbb{R}^{(m)}_{+} \) its scalar product is equal to the cosine of angle between \( \hat{P} \) and \( \hat{Q} \) which, of course, has the properties (i)-(iv).

III. APPLICATION OF THE BHATTACHARYYA STATISTICAL DIVERGENCE TO QUANTUM MECHANICS

Consider two observables \( A, B \) with Hermitian operators \( \hat{A}, \hat{B} \) in an \( N \)-dimensional Hilbert space, whose corresponding complete orthonormal sets of eigenvectors \( \{ |x_i^{(A)}\rangle \}, \{ |x_i^{(B)}\rangle \}, \ldots \ (i = 1, 2, \ldots, N) \) are disjointed and have nondegenerate spectra. Let \( |\phi\rangle \) be a normalized state vector of \( N \)-dimensional Hilbert space then it holds

\[
|\phi\rangle = \sum_{i}^{N} a_i |x_i^{(A)}\rangle, \quad |\phi\rangle = \sum_{j}^{N} b_j |x_j^{(B)}\rangle, \ldots
\]

Accordingly, the components of the probability distributions \( P(A) \) and \( P(B) \) associated with the observables \( A \) and \( B \) are

\[
P_i(A) = |a_i|^2 = |\langle x_i^{(A)} | \phi \rangle|^2; \quad \sum_{i=1}^{n} P_i(A) = \sum_{i=1}^{n} |a_i|^2 = 1 \quad (2a)
\]

\[
P_i(B) = |b_i|^2 = |\langle x_j^{(B)} | \phi \rangle|^2; \quad \sum_{i=1}^{n} P_i(B) = \sum_{i=1}^{n} |b_i|^2 = 1 \quad (2b)
\]

Inserting Eqs.(2a) and (2b) into Eq.(1) we get

\[
S(P(A), P(B)) = \sum_{i=1}^{n} (P_i(A) P_i(B))^{1/2} = \sum_{i=1}^{n} |a_i||b_i|.
\]
If $A = B$ then $|a_i| = |b_i|$ for $i = 1, 2, ..., n$ and Eqs. (2a) and (2b) becomes

$$S(P(A), P(B)) = \sum_{i=1}^{n} P_i(a) = \sum_{i=1}^{n} |a_i|^2 = \sum_{i=1}^{n} P_i(B) = \sum_{i=1}^{n} |b_i|^2 = 1.$$ 

Hence, for $A \equiv B$ it follows $S(P(A), P(B)) = 1$. Given the state vector $|\phi\rangle$ and operators $\hat{A}, \hat{B}$ the considered statistical divergence of their probability distributions can be generally determined. To each operator, a ray in the Hilbert space can be assigned. The quantity $S(P(A), P(B))$ gives the closeness of different rays in Hilbert space. If these rays are identical then their Bhattacharyya divergence $S(P(A), P(B))$ is equal to 1. If they are perpendicular to each other then $S(P(A), P(B))$ becomes zero. Generally, $S(P(A), P(B)) A \equiv \hat{B}$ and the corresponding rays of these operators in Hilbert space are identical, i.e. the cosine of angle between them is equal to 1. Therefore, $S(P(A), P(B)) = 1$.

Next, we consider the case of two non-commuting observables. Consider two observables $A$ and $B$ with non-commuting Hermitian operators $\hat{A}$ and $\hat{B}$ in an $N$-dimensional Hilbert space, whose corresponding complete orthonormal sets of eigenvectors $\{|x_i\rangle, \{y_i\rangle \ (i = 1, 2, ..., N)$ are disjoint and have nondegenerate spectra. Let $|\phi\rangle$ be a normalized state vector of $N$-dimensional Hilbert space then it holds

$$|\phi\rangle = \sum_{i}^{N} a_i |x_i\rangle, \quad |\phi\rangle = \sum_{j}^{N} b_j |y_j\rangle.$$ 

According the quantum transformation theory we have

$$|\phi\rangle = \left( \sum_{i}^{N} a_i \langle x_i|y_1\rangle |y_1\rangle \right) |y_1\rangle + \left( \sum_{i}^{N} a_i \langle x_i|y_2\rangle |y_2\rangle \right) |y_2\rangle + \cdots = \sum_{j}^{N} \sum_{i}^{N} a_i \langle x_i|y_j\rangle |y_j\rangle$$

$$P_i(A) = |\langle x_i|\phi\rangle|^2 = |a_i|^2, \quad Q_j(B) = |\langle y_j|\phi\rangle|^2 = \left| \sum_{i}^{N} a_i \langle x_i|y_j\rangle \right|^2,$$

(3)

where $\langle x_i|y_j\rangle \ i, j = 1, 2, 3, ..., N$ are the elements of the transformation matrix $T$ between the observables $A$ and $B$

$$T = \begin{pmatrix} \langle x_1|y_1\rangle & \langle x_1|y_2\rangle & \cdots & \langle x_1|y_n\rangle \\ \langle x_2|y_1\rangle & \cdots & \langle x_2|y_n\rangle \\ \vdots & \ddots & \vdots \\ \langle x_n|y_1\rangle & \cdots & \langle x_n|y_n\rangle \end{pmatrix}$$

Inserting Eq.(3) into Eq.(1) we get for the Bhattacharyya statistical divergence of the probability distributions $P(A)$ and $Q(B)$ the expression

$$S(P(A), Q(B)) = \sum_{i=1}^{n} \sqrt{P_i(A), Q_i(B)} = \sum_{i=1}^{n} |a_i||b_i| = \sum_{i=1}^{n} |\langle x_i|\phi\rangle||\langle y_i|\phi\rangle| = \sum_{j=1}^{n} |a_j| \left( \sum_{i=1}^{n} a_i \langle x_i|y_j\rangle \right). \quad (1)$$

Given the state vector $|\phi\rangle$ and the components of $T$, the divergence of the probability distributions of $A$ and $B$ can be generally determined.

Next, we present an example for determining of Bhattacharyya divergence of two concrete complementary observables describing a simple quantum system.

**IV. AN EXAMPLE**

For the sake of simplicity, we will consider the complementary observables in a two-dimensional Hilbert space. Such system represents a particle with spin $\hbar/2$ [13]. Determining the probability distributions of the components $J_x$ and $J_y$ we can calculate the similarity measure of their probability distributions.

The state vector of this quantum system is spinor

$$|\Psi\rangle = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix},$$
where 
\[ a_1a_1^* + a_2a_2^* = 1. \]

Its wave functions in z-representation takes the form \(|\Psi\rangle_z = a_1|z_1\rangle + a_2|z_2\rangle\). According Eq.(3), the probability \(P_{z_1}\) and \(P_{z_2}\) that \(J_z\) is projected on \(|z_1\rangle\) and \(|z_2\rangle\) is \(|a_1|^2 = a_1a_1^*\) and \(|a_2|^2 = a_2a_2^*\), respectively, so the corresponding probabilistic schema becomes
\[
\frac{J_z}{P_{a_1a_1^*}} \quad a_1a_1^* \quad a_2a_2^*
\]

Using the transfer transformation matrix
\[
T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1, & 1 \\ 1, & -1 \end{pmatrix}
\]
we obtain \(\Psi\) in its x-representation \(|\Psi\rangle_x = \frac{a_1 + a_2}{\sqrt{2}} |x_1\rangle + \frac{a_1 - a_2}{\sqrt{2}} |x_2\rangle\). Similarly, the probabilistic scheme for \(J_x\) turns out to be
\[
\frac{J_x}{P_{2^{-1}(a_1 + a_2)(a_1^* + a_2^*)} P_{2^{-1}(a_1 - a_2)(a_1^* - a_2^*)}} \quad 2^{-1}(a_1 + a_2)(a_1^* + a_2^*) \quad 2^{-1}(a_1 - a_2)(a_1^* - a_2^*)
\]
Now, we express \(a_1\) and \(a_2\) by means of new variables \(r\) and \(\varphi\) in the following way
\[ a_1 = \sqrt{r} \exp(i\varphi_1), \quad a_2 = (\sqrt{1-r}) \exp(i\varphi_2), \quad \varphi = \varphi_1 - \varphi_2. \]
In these variables, we obtain for \(J_x\) and \(J_z\) the following probability distributions
\[ Q(J_x) = \{r^2, (1-r)^2\} \]
and
\[ P(J_z) = \left\{ \frac{1}{2} (1 + 2\sqrt{r - r^2} \cos \varphi), \frac{1}{2} (1 - 2\sqrt{r - r^2} \cos \varphi) \right\}. \]
The similarity measure of these probability distributions consists of two terms
\[ S(Q(J_x), P(J_z)) = T_1 + T_2, \tag{4} \]
where
\[ T_1 = \sqrt{\frac{1}{2}(1 + 2\sqrt{r - r^2} \cos \varphi)} \quad T_2 = \sqrt{\frac{1-r}{2}(1 - 2\sqrt{r - r^2} \cos \varphi)}. \]
If \(J_x\) occurs in one of its eigenstates, i.e. \(r = 1\) or \(0\), then the first or second term in Eq.(4) becomes zero and we obtain the minimal value of \(S(P(J_x), Q(J_z))\) equal to \(T_1 = T_2 = \sqrt{1/2}\). \(S(P(J_x), Q(J_z))\) attains its maximal value equal to 1 for \(r = \sqrt{1/2}\) and \(\varphi = \pi/2\). The 3D-plot \(S(P(J_x), Q(J_z))\) is given in Fig. 1.

This graph shows that in the vicinity of \(r = 0.85\) and for \(\varphi \in [0, \pi/2]\) a hump occurs, where the probability distributions are almost similar. \(S(P(J_x), Q(J_z))\) never drops under the value \(\sqrt{1/2}\), therefore it holds
\[ S(P(J_x), Q(J_z)) \geq \sqrt{1/2}. \]
While \(S(P:Q)\) for two commuting observables in two-dimensional Hilbert space can attain arbitrary value, the similarity measure of the probability distributions of the investigating non-commuting observables, \(J_x\) and \(J_z\), is bounded by the value \(\sqrt{1/2}\). We note that two observables \(A, B\) in an \(N\)-dimensional Hilbert space are said to be complementary (to each other) if their transformation matrix has the form \([11]\]
\[ |⟨a_i|b_j⟩| = N^{-1/2} \quad (i, j = 1, ..., N). \]
Complementary observables can be considered as a generalization to higher dimensions of spin-1/2 orthogonal system \([12]\). Hence, we can proceed similarly also for two complementary observables in a general \(N\)-dimensional Hilbert space.

The concept of Bhattacharyya statistical divergence is quite independent of quantum mechanics and can be defined in any probability space. Hence, apart from the application of the similarity measure in quantum physics it can also be applied to theory of the physical measurement. It may serve as a certain degree of the reliance of a physical measurement. Suppose that in two different laboratories the probabilities of the decay modes of an elementary particle are measured. Generally, the different probability distributions of the individual decay modes are obtained in each laboratory. To determine the reliance of the measurement we can insert the measured probability distributions in formula (1). Here, the simple rule holds: the large is the value of \(S(P_1, P_2)\) the more reliable is the corresponding measurement.
FIG. 1: 3D-plot of the similarity measure of the probabilities $P_{J_1}$ and $P_{J_2}$ as function of $r$ and $\varphi$.

Acknowledgments

Partial support by the Grant Agency VEGA No. 2/6087/26 is highly acknowledged.

[1] I. Vajda, *Theory of statistical Inference and Information*. Dordrecht: Kluwer, 185.
[2] J. W. Karpur, J. Math. Phys. Sci. 17, (1983), 103.
[3] S. Kullback, *Information Theory and Statistics*. (John Willey: New York) 1959.
[4] S. Kullback, R. A. Leibler, Ann. Math. Stat. 22, (1951), 79. (1951).
[5] A. Rényi, "On the Measures of Entropy and Information". In: *4th Berkeley Symp.Math.Stat.Probability*, 1 (1961), 541.
[6] J. H. Havrda and F. Charvat, Kybernetika 3 (1967), 30.
[7] C. Tsallis, J. Stat. Phys. 52 (1988), 479.
[8] C. Tsallis, Brazilian J. Phys. 29 (1999), 1-87.
[9] V. Majerník, *Elementary theory of organization*. (Palacký University Press, Olomouc: 2001).
[10] A. Bhattacharyya, *On a Measure of Divergence Between Two Statistical Populations Defined by their Probability Distributions*, Bull. Calcutta Math.Soc. 35. (1943), 99-109.
[11] K. Kraus, Phys. Rev. D 35 (1987), 3070.
[12] J. Sánchez, Phys. Lett. A 173 (1993), 233.
[13] E. Merzbacher, *Quantum Mechanics*, 2nd Ed. (Wiley, New York, 1972).
[14] G. Aczél and Dároczy, *On Measures of Information and Their Characterization*, Academic Press, New York, 1975.
[15] V. Majerník, Intern. J. Gen. Systems 33 (2004), 673.