OPEN ACCESS

University of BRISTOL

Tamir, B., & Cohen, E. (2015). A Holevo-Type Bound for a Hilbert Schmidt Distance Measure. Journal of Quantum Information Science, 5(4), 127-133. DOI: 10.4236/jqis.2015.54015

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
10.4236/jqis.2015.54015

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research
General rights

Scientific
Research
Publishing

A Holevo-Type Bound for a Hilbert Schmidt Distance Measure

Boaz Tamir[1], Eliahu Cohen[2,3]

[1]Faculty of Interdisciplinary Studies, Bar-Ilan University, Ramat-Gan, Israel
[2]H.H. Wills Physics Laboratory, University of Bristol, Bristol, UK
[3]School of Physics and Astronomy, Tel-Aviv University, Tel-Aviv, Israel
Email: canjlm@actcom.co.il, eliahuco@post.tau.ac.il

Abstract

We prove a new version of the Holevo bound employing the Hilbert-Schmidt norm instead of the Kullback-Leibler divergence. Suppose Alice is sending classical information to Bob by using a quantum channel while Bob is performing some projective measurements. We bound the classical mutual information in terms of the Hilbert-Schmidt norm by its quantum Hilbert-Schmidt counterpart. This constitutes a Holevo-type upper bound on the classical information transmission rate via a quantum channel. The resulting inequality is rather natural and intuitive relating classical and quantum expressions using the same measure.

1. Introduction

Holevo's theorem [1] is one of the pillars of quantum information theory. It can be informally summarized as follows: "It is not possible to communicate more than n classical bits of information by the transmission of n qubits alone". It therefore sets a useful upper bound on the classical information rate using quantum channel.

Suppose Alice prepares a state ρ_x in some systems Q, where $x \in X = \{0, \cdots, n\}$ with probabilities $p(x)$. Bob performs a measurement described by the POVM elements $E_Y = E_0, \cdots, E_m$ on that state, with measurement outcome Y. Let

$$\rho = \rho^Q = \sum_x p(x) \rho_x. \tag{1}$$

The Holevo bound states that [2]

$$H(X:Y) \leq S(\rho) - \sum_x p(x) S(\rho_x), \tag{2}$$

where S is the von Neumann entropy and $H(X:Y)$ is the Shannon mutual information of X and Y. Recent proofs of the Holevo bound can be found in [3] [4].

Consider the following trace distance between two probability distributions $p(x)$ and $q(x)$ on X (note that the trace distance here is different than the one used in [2], chapter 9, by a square)

$$d(p\|q) = \frac{1}{2} \sum_x (p(x) - q(x))^2. \tag{3}$$

We can extend the definition to density matrices ρ and σ

$$d(\rho\|\sigma) = \frac{1}{2} tr(\rho - \sigma)^2, \tag{4}$$

where we use A^2 for $A^\dagger A$. This is known as the Hilbert-Schmidt (HS) norm [5]-[7] (in fact this is one half of the HS norm). Recently, the above distance measure was coined the "logical divergence" of two densities [8].

We prove a Holevo-type upper bound on the mutual information of X and Y, where the mutual information is written this time in terms of the HS norm instead of the Kullback-Leibler divergence. It is recently suggested by Ellerman [8] that employing the HS norm in the formulation of classical mutual information is natural. This is consistent with the identification of information as a measure of distinction [8]. Note that employing the Kullback-Leibler divergence in the standard form of the Holevo bound gives an expression which can be identified with quantum mutual information, however, the "coherent information" is considered as a more appropriate expression (see also [2] chapter 12). In view of the above we hereby take Ellerman's idea a step further and write a Holevo-type bound based on the HS norm.

The question whether $d(\rho\|\sigma)$ was the right measure of quantum mutual information was discussed in [9], within the context of area laws. It was used there to provide an upper bound on the correlations between two distant operators M_A and M_B, where A is a region inside a spin grid and B is its complement:

$$d(\rho^{A,B}\|\rho^A \otimes \rho^B) \geq \frac{C(M_A, M_B)^2}{2\|M_A\|^2 \|M_B\|^2}, \tag{5}$$

where $C(M_A, M_B) = \langle M_A \otimes M_B \rangle - \langle M_A \rangle \langle M_B \rangle$ is the correlation function of M_A and M_B.

In addition, the HS norm was suggested as an entanglement measure [6] [10], however, this was criticized in [11], claiming it did not fulfill the so called CP non-expansive property (*i.e.* non-increasing under every completely-positive trace-preserving map).

In the following, we will prove a Holevo-type bound on the above HS distance between the probability $p(x,y)$ on the product space (X,Y) and the product of its marginal probabilities $p(x) \cdot p(y)$:

$$\frac{1}{q} d(X,Y\|X \otimes Y) \leq d(\rho^{P,Q}\|\rho^P \otimes \rho^Q), \tag{6}$$

where

$$\rho^{P,Q} = \sum_x p(x)|x\rangle\langle x| \otimes \rho_x, \tag{7}$$

and where

$$\rho^Q = \sum_x p(x)\rho_x, \tag{8}$$

$$\rho^P = \sum_x p(x)|x\rangle\langle x| \tag{9}$$

are the partial traces of $\rho^{P,Q}$, and q is the dimension of the space Q. Note that both sides of Inequality (6) are measures of mutual information. Therefore, our claim is that the classical HS mutual information is bounded by the corresponding quantum one multiplied by the dimension of the quantum density matrices used in the channel. We will also show that

$$\frac{1}{q} d(X,Y\|X \otimes Y) \leq \frac{1}{2} L(\rho^P) L(\rho^Q), \tag{10}$$

where $L(\rho^Q)$ and $L(\rho^P)$ are the Tsallis entropies [12]

$$L(\rho) = tr(\rho(1-\rho)),$$ (11)

also known as the linear entropy, purity [13] or logical entropy of ρ [8].

All the above is proved for the case of projective measurements. However, we expect similar results in the general case of POVM, in light of Naimark's dilation theorem (see [14] or [15] for instance).

In the next section we review some basic properties of quantum logical divergence and then use these properties to demonstrate the new Holevo-type bound.

2. The HS Norm and the Holevo-Type Bound

Let

$$d(\rho\|\sigma) = \frac{1}{2}tr(\rho-\sigma)^2.$$ (12)

In what follows we recall some basic properties of the HS distance measure, then we state and prove the main result of this paper.

Theorem 2.1. Contractivity of the HS norm with respect to projective measurements

Let ρ and σ be two density matrices of a system S. Let $\mathcal{E}(\rho)$ be the trace preserving operator

$$\mathcal{E}(\rho) = \sum_i P_i \rho P_i,$$ (13)

where the projections P_i satisfy $\sum_i P_i = I$, $P_i^\dagger = P_i$ and $P_i^2 = P_i$ for every i, then

$$d(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \le d(\rho\|\sigma).$$ (14)

Proof: We now write $X = \rho - \sigma$. Then X is Hermitian with bounded spectrum, and using Lemma 2 in [16] we conclude that

$$tr(\mathcal{E}\rho - \mathcal{E}\sigma)^2 < tr(\rho-\sigma)^2.$$ (15)

Theorem 2.2. The joint convexity of the HS norm

The logical divergence $d(\rho\|\sigma)$ is jointly convex.

Proof: First observe that $tr(\rho^2)$ is convex from the convexity of x^2 and the linearity of the trace. Next we can write

$$
\begin{aligned}
&d(\lambda\rho_1 + (1-\lambda)\rho_2\|\lambda\sigma_1 + (1-\lambda)\sigma_2)\\
&= \frac{1}{2}tr((\lambda\rho_1 + (1-\lambda)\rho_2) - (\lambda\sigma_1 + (1-\lambda)\sigma_2))^2\\
&= \frac{1}{2}tr(\lambda(\rho_1-\sigma_1) + (1-\lambda)(\rho_2-\sigma_2))^2\\
&\le \frac{\lambda}{2}tr(\rho_1-\sigma_1)^2 + \frac{1-\lambda}{2}tr(\rho_2-\sigma_2)^2\\
&= \lambda d(\rho_1\|\sigma_1) + (1-\lambda)d(\rho_2\|\sigma_2),
\end{aligned}
$$ (16)

where the inequality is due to the convexity of $tr(\rho^2)$. This constitutes the joint convexity.

Theorem 2.3. The monotonicity of the HS norm with respect to partial trace

Let $\rho^{A,B}$ and $\sigma^{A,B}$ be two density matrices, then

$$d(\rho^A \otimes I/b\|\sigma^A \otimes I/b) \le d(\rho^{A,B}\|\sigma^{A,B}),$$ (17)

where b is the dimension of B.

Proof: One can find a set of unitary matrices U_j over B and a probability distribution p_j such that

$$\rho^A \otimes I/b = \sum_j p_j U_j \rho^{A,B} U_j^\dagger$$ (18)

$$\sigma^A \otimes I/b = \sum_j p_j U_{j} \sigma^{A,B} U_{j}^{\dagger}, \tag{19}$$

(see [2] chapter 11). Now since $d(\rho\|\sigma)$ is jointly convex on both densities, we can write

$$d\left(\rho^A \otimes I/b \| \sigma^A \otimes I/b\right) \leq \sum_j p_j \cdot d\left(U_j \rho^{A,B} U_j^{\dagger} \| U_j \sigma^{A,B} U_j^{\dagger}\right). \tag{20}$$

Observe now that the divergence is invariant under unitary conjugation, and therefore the sum in the right hand side of the above inequality is $d\left(\rho^{A,B}\|\sigma^{A,B}\right)$.

We can now state the main result:

Theorem 2.4. A Holevo-type bound for the HS trace distance between $p(x,y)$ **and** $p(x)\cdot p(y)$

Suppose Alice is using a distribution $p(x)$, where x is in $X = \{1,\cdots,n\}$, to pick one of n densities ρ_x in Q. She then sends the signal in a quantum physical channel to Bob. We can add an artificial quantum system P and write (P,Q) for Alice as:

$$\rho^{P,Q} = \sum_x p(x)|x\rangle\langle x| \otimes \rho_x, \tag{21}$$

where the vectors $|x\rangle$ are orthogonal. Let ρ^P and ρ^Q be the partial traces of $\rho^{P,Q}$. Suppose Bob is measuring the system using a projective measurement as in Theorem 2.1, then

$$\frac{1}{q}d\left(X,Y\|X \otimes Y\right) \leq d\left(\rho^{P,Q}\|\rho^P \otimes \rho^Q\right), \tag{22}$$

where q is the dimension of the space Q.

Proof: First we consider one more auxiliary quantum system, namely M for the measurement outcome for Bob. Initially the system M is in the state $M_0 = |0\rangle\langle 0|$. Let $\mathcal{E}$ be the operator defined by Bob's measurement as in Theorem 2.1 above: let $\{P_y\}_y$ on Q be defined such that $\sum_y P_y = I$ and

$$\mathcal{E}(\rho_x) = \sum_y P_y \rho_x P_y. \tag{23}$$

One can easily extend $\mathcal{E}$ to the space (Q,M) by

$$\mathcal{E}(\rho \otimes |0\rangle\langle 0|) = \sum_y P_y \rho P_y \otimes |y\rangle\langle y|. \tag{24}$$

This can be done by choosing a set of operators, conjugating $|0\rangle\langle 0|$ to $|y\rangle\langle y|$. It amounts to writing the measurement result in the space M (see also Ch. 12.1.1 in [2]). If we now trace out Q we find

$$tr_Q \mathcal{E}(\rho \otimes |0\rangle\langle 0|) = \sum_y p(y)|y\rangle\langle y|. \tag{25}$$

Moreover, $\mathcal{E}$ can be extended to $\rho^{(P,Q,M)}$ by

$$\mathcal{E}\left(\sum_x p(x)|x\rangle\langle x| \otimes \rho_x \otimes |0\rangle\langle 0|\right) = \sum_x \sum_y p(x)|x\rangle\langle x| \otimes P_y \rho_x P_y \otimes |y\rangle\langle y|. \tag{26}$$

If we trace out Q we arrive at

$$tr_Q \mathcal{E}\left(\rho^{(P,Q,M_0)}\right) = \sum_{x,y} p(y/x) p(x)|x\rangle\langle x| \otimes |y\rangle\langle y| = \sum_{x,y} p(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y|. \tag{27}$$

Finally, we can extend $\mathcal{E}$ to $\rho^P \otimes \rho^{(Q,M)}$ by

$$\mathcal{E}\left(\sum_x p(x)|x\rangle\langle x| \otimes \rho \otimes |0\rangle\langle 0|\right) = \sum_x \sum_y p(x)|x\rangle\langle x| \otimes P_y \rho P_y \otimes |y\rangle\langle y|. \tag{28}$$

If we trace out Q we get

$$tr_Q \mathcal{E}\left(\rho^P \otimes \rho^{(Q,M_0)}\right) = \sum_x \sum_y p_x|x\rangle\langle x| \otimes p(y)|y\rangle\langle y| = \left(\sum_x p_x|x\rangle\langle x|\right) \otimes \left(\sum_y p_y|y\rangle\langle y|\right). \tag{29}$$

We can now use the properties stated in the above theorems, Equation (27) and Equation (29) to deduce

$$d\left(\rho^{P,Q}\|\rho^P\otimes\rho^Q\right)=d\left(\rho^{P,Q,M_0}\|\rho^P\otimes\rho^{Q,M_0}\right)\geq d\left(\mathcal{E}\left(\rho^{P,Q,M_0}\right)\|\mathcal{E}\left(\rho^P\otimes\rho^{Q,M_0}\right)\right)$$
$$\geq d\left(tr_Q\mathcal{E}\left(\rho^{P,Q,M_0}\right)\otimes I_q\|tr_Q\mathcal{E}\left(\rho^P\otimes\rho^{(Q,M_0)}\right)\otimes I_q\right)=\frac{1}{q}d\left(X,Y\|X\otimes Y\right),\tag{30}$$

where in the first inequality we have used Theorem 2.1 and in the second inequality Theorem 2.3. The final equality is an easy consequence of the definition of the HS norm.

Corollary: Suppose Alice is sending classical information to Bob using a quantum channel Q, Bob measures the quantum state using a projective measurement defined above (having results in space Y). Under all the above assumptions

$$\frac{1}{q}d\left(X,Y\|X\otimes Y\right)\leq\frac{1}{2}L\left(\rho^P\right)L\left(\rho^Q\right),\tag{31}$$

where $L\left(\rho^Q\right)$ and $L\left(\rho^P\right)$ are Tsallis entropies of the second type (the quantum logical entropies) of ρ^Q and ρ^P.

Proof: Clearly (see also [17])

$$d\left(\rho^{P,Q}\|\rho^P\otimes\rho^Q\right)=tr\left(\rho^{P,Q}\left(1-\rho^P\otimes\rho^Q\right)\right)-\frac{1}{2}L\left(\rho^P\otimes\rho^Q\right)-\frac{1}{2}L\left(\rho^{P,Q}\right).\tag{32}$$

It is easy to see (by a matrix representation) that for $\rho^{P,Q}$ as in Equation (21)

$$tr\left(\rho^{P,Q}\left(1-\rho^P\otimes\rho^Q\right)\right)=L\left(\rho^{P,Q}\right),\tag{33}$$

therefore

$$d\left(\rho^{P,Q}\|\rho^P\otimes\rho^Q\right)=\frac{1}{2}L\left(\rho^{P,Q}\right)-\frac{1}{2}L\left(\rho^P\otimes\rho^Q\right).\tag{34}$$

However, $L\left(\rho^{P,Q}\right)\leq L\left(\rho^P\right)+L\left(\rho^Q\right)$, and $L\left(\rho^P\otimes\rho^Q\right)=L\left(\rho^P\right)+L\left(\rho^Q\right)-L\left(\rho^P\right)\cdot L\left(\rho^Q\right)$ (see [17] Theorem II.2.4 and Theorem II.3), therefore

$$d\left(\rho^{P,Q}\|\rho^P\otimes\rho^Q\right)\leq\frac{1}{2}L\left(\rho^P\right)\cdot L\left(\rho^Q\right).\tag{35}$$

Combining this with Theorem 2.4 we find

$$\frac{1}{q}d\left(X,Y\|X\otimes Y\right)\leq\frac{1}{2}L\left(\rho^P\right)L\left(\rho^Q\right).\tag{36}$$

Example: Suppose Alice sends the state $|0\rangle$ with probability 1/2 and the state $|\psi\rangle=\cos\theta|0\rangle+\sin\theta|1\rangle$ with probability 1/2, then

$$\rho^{P,Q}=\frac{1}{2}|0\rangle\langle0|\otimes\rho_0+\frac{1}{2}|1\rangle\langle1|\otimes\rho_1,\tag{37}$$

where $\rho_0=|0\rangle\langle0|$ and $\rho_1=|\psi\rangle\langle\psi|$. By partial tracing we get

$$\rho^Q=\frac{1}{2}\begin{pmatrix}1&0\\0&0\end{pmatrix}+\frac{1}{2}\begin{pmatrix}\cos^2\theta&\cos\theta\sin\theta\\\cos\theta\sin\theta&\sin^2\theta\end{pmatrix},\tag{38}$$

and ρ^P is a balanced coin. The eigenvalues of ρ^Q are $(1\pm\cos\theta)/2$ and therefore

$$L\left(\rho^Q\right)=1-\left[(1+\cos\theta)^2/4+(1-\cos\theta)^2/4\right]=\frac{1}{2}\sin^2\theta.\tag{39}$$

Also $L\left(\rho^P\right)=1/2$ and $q=2$, hence

$$d\left(X,Y\|X\otimes Y\right)\leq\frac{1}{4}\sin^2\theta\leq1/4.\tag{40}$$

The left hand side of the above inequality is a measure of the classical mutual information according to the HS norm between *X* and *Y*. The very fact that it is smaller than the Tsallis information measure of *X* (which is 1/2) means that the quantum channel restricts the rate of classical information transfer, where the mutual information is measured by the HS norm and the source of information *X* is measured by Tsallis entropy. This is analogous to Holevo's upper bound in the framework of Tsalis/linear entropy. We find this result similar in spirit to the well-known limitation on the rate of classical information transmission via a quantum channel (without utilizing entanglement): one cannot send more than one bit for each use of the channel using a one qubit channel.

In the above example, if ρ_0 and ρ_1 are mixed states, then by the same argument we can show that:

$$d\left(X,Y\|X\otimes Y\right)\leq\frac{1}{2}L\left(\rho^Q\right) \tag{41}$$

This gives a bound on the classical mutual information using the quantum "logical entropy" (the Tsallis entropy).

3. Discussion

We proved a Holevo-type bound employing the Hilbert-Schmidt distance between the density matrices on the product space $\left(X,Y\right)$ and the tensor of the two marginal density matrices on $X\otimes Y$. Using a different measure of mutual information, we showed that this Holevo-type upper bound on classical information transmission could be written as an inequality between the classical mutual information expression and its quantum counterpart.

It seems that by utilizing Naimark's dilation [14] [15], the above result can be generalized to any POVM if one is willing to employ the suitable channel in a higher dimensional Hilbert space.

As was claimed in [17], the divergence distance used above is the natural one in the context of quantum logical entropy [8]. Being the "right" measure of mutual information in quantum channels passing classical information, we expect that this formalism would be helpful in further studies of various problems such as channel capacity theory, entanglement detection and area laws.

Acknowledgements

References

[1] Holevo. A.S. (1973) Bounds for the Quantity of Information Transmitted by a Quantum Communication Channel. *Problemy Peredachi Informatsii*, **9**, 3-11.

[2] Nielsen, M.A. and Chuang, I.L. (2010) Quantum Computation and Quantum Information. Cambridge University Press, Cambridge, UK. http://dx.doi.org/10.1017/CBO9780511976667

[3] Cerf, N.J. and Adami, C. (1996) Accessible Information in Quantum Measurement. http://arxiv.org/abs/quant-ph/9611032

[4] Winter, A. (1999) Coding Theorem and Strong Converse for Quantum Channels. *IEEE Transactions on Information Theory*, **45**, 2481-2485. http://dx.doi.org/10.1109/18.796385

[5] Buzek, V. and Hillery, M. (1996) Quantum Copying: Beyond the No-Cloning Theorem. *Physical Review A*, **54**, 1844-1852. http://dx.doi.org/10.1103/PhysRevA.54.1844

[6] Vedral, V. and Plenio, M.B. (1998) Entanglement Measures and Purification Procedures. *Physical Review A*, **57**, 1619-1633. http://dx.doi.org/10.1103/PhysRevA.57.1619

[7] Perez-Garcia, D., Wolf, M.M., Petz, D. and Ruskai, M.B. (2006) Contractivity of Positive and Trace-Preserving Maps under Lp Norms. *Journal of Mathematical Physics*, **47**, Article ID: 083506. http://dx.doi.org/10.1063/1.2218675

[8] Dllerman, E. (2013) Information as Distinction: New Foundation for Information Theory. http://arxiv.org/abs/1301.5607

[9] Wolf, M.M., Verstraete, F., Hasting, M.B. and Cirac, J.I. (2008) Area Law in Quantum Systems: Mutual Information and Correlations. *Physical Review Letters*, **100**, Article ID: 070502. http://dx.doi.org/10.1103/PhysRevLett.100.070502

[10] Witte, C. and Trucks, M. (1999) A New Entanglement Measure Induced by the Hilbert-Schmidt Norm. *Physics Letters A*, **257**, 14-20. http://dx.doi.org/10.1016/S0375-9601(99)00279-0

[11] Ozawa, M. (2000) Entanglement Measures and the Hilbert-Schmidt Distance. *Physics Letters A*, **268**, 158-160. http://dx.doi.org/10.1016/S0375-9601(00)00171-7

[12] Tsallis, C. (1988) Possible Generalization of Boltzmann-Gibbs Statistics. *Journal of Statistical Physics*, **52**, 479-487. http://dx.doi.org/10.1007/BF01016429

[13] Peters, N.A., Wei, T.C. and Kwiat, P.G. (2004) Mixed State Sensitivity of Several Quantum Information Benchmarks. *Physical Review A*, **70**, Article ID: 052309. http://dx.doi.org/10.1103/PhysRevA.70.052309

[14] Paulsen, V. (2002) Completely Bounded Maps and Operator Algebras. Cambridge University Press, Cambridge, UK.

[15] Matsumoto, K., Shimono, T. and Winter, A. (2004) Remarks on Additivity of the Holevo Channel Capacity and of the Entanglement of Formation. *Communications in Mathematical Physics*, **246**, 427-442. http://dx.doi.org/10.1007/s00220-003-0919-0

[16] Rastegin, A.E. (2010) Some General Properties of Unified Entropies. *Journal of Statistical Physics*, **143**, 1120-1135. http://dx.doi.org/10.1007/s10955-011-0231-x

[17] Tamir, B. and Cohen, E. (2014) Logical Entropy for Quantum States. http://arxiv.org/abs/1412.0616