

Optimal entropic uncertainty relation for successive measurements in quantum information theory

M D SRINIVAS

Centre for Policy Studies, 27, Rajasekaran Street, Mylapore, Chennai 600 004, India
Email: policy@vsnl.com

MS received 8 July 2002; revised 9 December 2002; accepted 20 December 2002

Abstract. We derive an optimal bound on the sum of entropic uncertainties of two or more observables when they are sequentially measured on the same ensemble of systems. This optimal bound is shown to be greater than or equal to the bounds derived in the literature on the sum of entropic uncertainties of two observables which are measured on distinct but identically prepared ensembles of systems. In the case of a two-dimensional Hilbert space, the optimum bound for successive measurements of two-spin components, is seen to be strictly greater than the optimal bound for the case when they are measured on distinct ensembles, except when the spin components are mutually parallel or perpendicular.

Keywords. Uncertainty relations; information-theoretic entropy; optimum bounds; successive measurements; influence of measurements on uncertainties.

PACS Nos 03.65.Bz; 03.67.-a

1. Uncertainty relations for distinct and successive measurements

Heisenberg, in his celebrated paper on uncertainty relations [1], written seventy-five years ago, seems to have interpreted these relations as expressing the influence of measurement of one observable on the uncertainty in the outcome of a succeeding measurement of another observable performed on the same system. Heisenberg states:

“At the instant when position is determined – therefore, at the moment when the photon is scattered by the electron – the electron undergoes a discontinuous change in momentum. This change is the greater the smaller the wavelength of the light employed – that is, the more exact the determination of the position. At the instant at which the position of the electron is known, its momentum therefore can be known up to magnitudes, which correspond to that discontinuous change. Thus, the more precisely the position is determined, the less precisely the momentum is known, and conversely.”

However, the standard mathematical formulation of the uncertainty relations, as first derived by Robertson in 1929 [2] from the first principles of quantum theory, does not in any way support the above interpretation. If A , B are two observables and ρ is the

density operator characterizing the state of the system, then the standard formulation of the uncertainty relation is given by the inequality

$$(\Delta^{\rho}A)^2(\Delta^{\rho}B)^2 \geq \frac{1}{4}|\text{Tr}(\rho[A, B])|^2, \quad (1)$$

where the variances $(\Delta^{\rho}A)^2$ and $(\Delta^{\rho}B)^2$ are given by

$$(\Delta^{\rho}A)^2 = \text{Tr}(\rho A^2) - [\text{Tr}(\rho A)]^2, \quad (2a)$$

$$(\Delta^{\rho}B)^2 = \text{Tr}(\rho B^2) - [\text{Tr}(\rho B)]^2. \quad (2b)$$

Since the variance $(\Delta^{\rho}A)^2$ of an observable A is non-negative and vanishes only when the state ρ is a mixture of eigenstates of A that are associated with the same eigenvalue, it is a reasonable measure of the uncertainty in the outcome of an A -measurement carried out on an ensemble of systems in state ρ . However, $(\Delta^{\rho}A)^2$ and $(\Delta^{\rho}B)^2$ as given by eqs (2a) and (2b) refer to uncertainties in the outcomes of A and B when they are measured on *distinct* though *identically prepared* ensembles of systems in state ρ . Thus the inequality (1) refers to a situation where the observables A, B are measured on distinct ensembles of systems and may hence be referred to as ‘uncertainty relations for distinct measurements’.

The current understanding of uncertainty relations (1) is succinctly summarized in the recent treatise on quantum computation and quantum information by Nielsen and Chuang as follows [3] (we have altered the symbols appearing in the quotation so as to correspond to inequality (1) above; emphasis as in the original):

“You should be vary of a common misconception about the uncertainty principle, that measuring an observable A to some ‘accuracy’ $\Delta^{\rho}A$ causes the value of B to be ‘disturbed’ by an amount $\Delta^{\rho}B$ in such a way that some sort of inequality similar to (1) is satisfied. While it is true that measurements in quantum mechanics cause disturbance to the system being measured, this is most emphatically *not* the content of the uncertainty principle.

The correct interpretation of the uncertainty principle is that if we prepare a large number of quantum systems in identical states, ρ , and then perform measurements of A on some of those systems and of B on others, then the standard deviation $\Delta^{\rho}A$ of the A results times the standard deviation $\Delta^{\rho}B$ of the results for B will satisfy the inequality (1).”

In fact, the uncertainty relation (1) essentially expresses the limitations imposed by quantum theory on the preparation of ensembles of systems in identical states, as has been clearly explained in the recent monograph on the conceptual foundations of quantum theory by Home [4]:

“To test uncertainty relation (1), we require a repeatable state preparation procedure leading to an ensemble of identically prepared particles, all corresponding to the state being studied. Then on each such system we must measure one or the other of the two dynamical variables (either A or B). The statistical distributions of the measured results of A and B are characterized by variances satisfying (1). The significance of the uncertainty principle is stated as follows: It is impossible to prepare an ensemble of identical particles, all in the same state, such that the product of ...[variances] of any two noncommuting dynamical variables has a value less than the lower bound given by relation (1).”

In order to explore the influence of the measurement of one observable on the uncertainties in the outcomes of another, we have to formulate an ‘uncertainty relation for successive measurements’, which is different from (1). For this purpose, we have to consider

Influence of measurements on uncertainties

the uncertainties in the outcomes of the A and B measurements, when they are performed sequentially on the same ensemble. If A and B are observables with purely discrete spectra with spectral resolution

$$A = \sum_i a_i P^A(a_i) \quad (3a)$$

$$B = \sum_j b_j P^B(b_j), \quad (3b)$$

then the joint probability $\text{Pr}_{A,B}^\rho(a_i, b_j)$ that the outcomes a_i, b_j are realised in a sequential measurement of the observables A, B , is given by the well-known Wigner formula [5,6]:

$$\text{Pr}_{A,B}^\rho(a_i, b_j) = \text{Tr}[P^B(b_j)P^A(a_i)\rho P^A(a_i)P^B(b_j)]. \quad (4)$$

In the above, and in what follows, we adopt the Heisenberg picture of time evolution, where the observables carry the entire burden of time evolution in the absence of measurements. An uncertainty relation for successive measurements can easily be derived [7] by calculating the variances $\Delta_{A,B}^\rho(A)$ and $\Delta_{A,B}^\rho(B)$, which correspond to the uncertainties in the outcomes of these observables when they are measured sequentially, by employing the joint probability (4):

$$\Delta_{A,B}^\rho(A)^2 \Delta_{A,B}^\rho(B)^2 \geq |\text{Tr}[\rho A \varepsilon(B)] - \text{Tr}[\rho A] \text{Tr}[\rho \varepsilon(B)]|^2, \quad (5)$$

where

$$\varepsilon(B) = \sum_i P^A(a_i) B P^A(a_i). \quad (6)$$

In a seminal paper, Deutsch [8] pointed out that, except in the case of canonically conjugate variables, inequality (1) does not adequately express the quantum uncertainty principle. The RHS of (1) crucially depends on the state ρ of the system except in the case of canonically conjugate observables. In order to get a nontrivial lower bound on the uncertainty in the outcome of B given the uncertainty in the outcome of A , we need to take the infimum of RHS of (1) over all states ρ , and this invariably vanishes whenever A or B has a single discrete eigenvalue. Therefore inequality (1) does not give any nontrivial lower bound on the uncertainties in most physical situations involving angular momenta, spin or finite-level systems. The variance form of the uncertainty relation for successive measurements (5) is even more ineffective, for the infimum of RHS of (5) identically vanishes as the operators A and $\varepsilon(B)$ always commute.

Deutsch argued that, in order to have a meaningful formulation of the uncertainty principle, we should have a nontrivial lower bound on the product/sum of the uncertainties of two observables, which, unlike as in (1), is independent of the state and vanishes essentially only when the two observables have a common eigenvector. He also showed that it is possible to formulate such an uncertainty relation by making use of the information-theoretic entropy, instead of variance, as the measure of uncertainty.

2. Entropic uncertainty relations for distinct measurements

For a classical discrete random variable X which takes on n different values with associated probabilities $\{p_i\}$, the information-theoretic entropy given by

$$S(X) = - \sum_i p_i \log(p_i), \quad (7)$$

is a good measure of the uncertainty, or spread in the probability distribution. In (7) we assume that $0 \log(0) = 0$ always. Since, $0 \leq p_i \leq 1$ and $\sum p_i = 1$, we have

$$0 \leq S(X) \leq \log(n). \quad (8)$$

The lower bound in (8) is attained when all the p_i vanish except for some p_k , $1 \leq k \leq n$; the upper bound is attained when all the p_i are equal to $1/n$.

The above definition can be extended to cases where the number of outcomes is not finite and also to the case when X is a continuous random quantity with probability density $p(x)$. In the latter case the information-theoretic entropy is usually defined as follows [9]:

$$S(X) = - \int p(x) \log p(x) dx. \quad (9)$$

Apart from the fact that $S(X)$ as defined above is not non-negative, it is also not physically meaningful as it has inadmissible physical dimensions. While the discrete probabilities $\{p_i\}$ and the corresponding entropy (7) are mere dimensionless numbers, the probability density $p(x)$ has dimension $(1/D)$, where D is the physical dimension of the random quantity X , and hence $S(X)$ as defined by eq. (9) has inadmissible physical dimension $\log(D)$. Therefore it has been suggested [10] that for continuous variables we should employ, instead of $S(X)$, its exponential $E(X)$ given by

$$E(X) = \exp \left[- \int p(x) \log p(x) dx \right]. \quad (10)$$

The exponential entropy (10) has a physically meaningful dimension D , the same as the random quantity X . It is a monotonic function of $S(X)$ and is also non-negative.

It may be of interest to note that the first entropic uncertainty relation in quantum theory was derived much before the work of Deutsch, and it was formulated for continuous observables, position and momentum. An entropic uncertainty relation for position and momentum was conjectured by Everett in his famous thesis [11] in 1957 and by Hirschman [12] the same year. It was proved in 1975 by Beckner [13] and Bialynicki-Birula and Mycielski [14]. If $\psi(x)$ is the wave function of a particle in one-dimension, and $\varphi(p)$ the corresponding momentum space wave function, then the entropic uncertainty relation obtained by these authors may be expressed in the form

$$- \int |\psi(x)|^2 \log |\psi(x)|^2 dx - \int |\varphi(p)|^2 \log |\varphi(p)|^2 dp \geq \log(\pi \hbar e), \quad (11)$$

where e is the base of natural logarithms. Clearly, both the LHS and RHS in (11) have inadmissible physical dimensions. However, this can be corrected by re-expressing the above relation in terms of the exponential entropies as follows:

$$E^\rho(Q)E^\rho(P) \geq \pi\hbar e. \quad (12)$$

The inequality (12) is the appropriate entropic uncertainty relation for position and momentum, in fact for any pair of canonically conjugate observables, and it can also be shown to be stronger than the well-known variance form for the uncertainty relation for such observables.

We now turn to the recent work on the entropic uncertainty relations for observables with purely discrete spectra following the work of Deutsch. Let us consider two observables A, B , with purely discrete spectra and spectral resolution as given by (3a), (3b). Then the entropic uncertainties of A and B as measured in distinct but identically prepared ensembles in state ρ , are given by

$$S^\rho(A) = - \sum_i \text{Tr}[\rho P^A(a_i)] \log \text{Tr}[\rho P^A(a_i)], \quad (13a)$$

$$S^\rho(B) = - \sum_j \text{Tr}[\rho P^B(b_j)] \log \text{Tr}[\rho P^B(b_j)]. \quad (13b)$$

It can easily be seen that the entropic uncertainty $S^\rho(A)$ (as well as $S^\rho(B)$) is non-zero and vanishes only when the state ρ is a mixture of eigenstates of A (correspondingly B) that are associated with the same eigenvalue. Now the optimum uncertainty relation for distinct measurements is of the form

$$S^\rho(A) + S^\rho(B) \geq \Lambda_D(A, B), \quad (14)$$

where

$$\Lambda_D(A, B) = \inf_{\rho} [S^\rho(A) + S^\rho(B)] \quad (15)$$

is the optimum lower bound. So far, it has not been possible to obtain a constructive expression for $\Lambda_D(A, B)$ in the general case; but sharper and sharper lower bounds have been obtained and of course all of them imply that $\Lambda_D(A, B)$ vanishes essentially only when A, B have a common eigenvector.

The first lower bound on $\Lambda_D(A, B)$ was obtained nearly 20 years ago by Deutsch [8] for the case when A, B have totally non-degenerate spectra. Then (3a), (3b) reduce to

$$A = \sum_i a_i |a_i\rangle\langle a_i|, \quad (16a)$$

$$B = \sum_j b_j |b_j\rangle\langle b_j|. \quad (16b)$$

For such observables, Deutsch showed that

$$\Lambda_D(A, B) \geq 2 \log \left[2 / \left(1 + \sup_{i,j} |\langle a_i | b_j \rangle| \right) \right]. \quad (17)$$

Clearly the RHS in (17) is non-negative and vanishes only when

$$\sup_{i,j} |\langle a_i | b_j \rangle| = 1, \quad (18)$$

which happens only when A, B have a common eigenstate or have eigenstates arbitrarily close to each other. Deutsch inequality (17) was soon generalized to the case when the spectra of A, B have degeneracies also, by Partovi [15], who showed that

$$\Lambda_D(A, B) \geq 2 \log \left[2 / \sup_{i,j} \| P^A(a_i) + P^B(b_j) \| \right], \quad (19)$$

where the symbol $\| \cdot \|$ stands for the operator norm. Partovi's inequality (19) reduces to Deutsch's inequality (17) when A, B have totally non-degenerate spectra.

It was Kraus [16], who pointed out that the inequalities (17) and (19) are not optimal. Based on explicit calculations for 2, 3 and 4-dimensional examples, he conjectured that a much stronger inequality holds in finite-dimensions for the case of the so-called 'complementary observables'. Two observables, A, B with totally non-degenerate spectra are said to be complementary, if their eigenvectors satisfy

$$|\langle a_i | b_j \rangle| = 1/\sqrt{n}, \quad (20)$$

for all $1 \leq i, j \leq n$, where n is the dimension of the Hilbert space. Kraus' conjecture was that such complementary observables obey the optimal uncertainty relation

$$S^p(A) + S^p(B) \geq \Lambda_D(A, B) = \log(n). \quad (21)$$

Kraus' conjecture was proved by Maassen and Uffink [17], who obtained the following lower bound on the sum of entropic uncertainties for any two observables with non-degenerate spectra in a finite-dimensional Hilbert space:

$$\Lambda_D(A, B) \geq \log \left[1 / \max_{i,j} |\langle a_i | b_j \rangle|^2 \right]. \quad (22)$$

For the case of complementary observables, which satisfy (20), Maassen–Uffink bound (22) reduces to the form (21) conjectured by Kraus.

Recently Krishna and Parthasarathy [18] have generalized the result of Maassen and Uffink to obtain the following result, which is valid for any pair of observables in a finite-dimensional Hilbert space:

$$\Lambda_D(A, B) \geq \log \left[1 / \max_{i,j} \| P^A(a_i) P^B(b_j) \|^2 \right]. \quad (23)$$

Clearly, Krishna–Parthasarathy bound (23) reduces to the Maassen–Uffink (and Kraus) bound (22) when the observables A, B have non-degenerate spectra. Following the remarks made by Maassen and Uffink [17], it seems possible to extend the inequalities (22), (23) (by replacing 'max' with 'sup'), so as to be applicable also to the case when the Hilbert space is infinite dimensional, as long as we restrict ourselves to observables with purely discrete spectra. Also, since

$$\| P^A(a_i) P^B(b_j) \|^2 = \| P^A(a_i) P^B(b_j) P^A(a_i) \| \leq \frac{1}{4} \| P^A(a_i) + P^B(b_j) \|^2, \quad (24)$$

the Krishna–Parthasarathy bound (23) is stronger than the Partovi bound (19) and is thus the sharpest available bound on the sum of entropic uncertainties of two observables in the case of distinct measurements.

However, Garrett and Gull [19] showed in 1990 that the Maassen–Uffink bound (22) is not optimal in the case of a two-dimensional Hilbert space when the two observables A, B , are not commuting or complementary. Recently Sanchez-Ruiz [20] has shown that the bound obtained by Garrett and Gull is indeed the optimal entropic uncertainty bound in two-dimensional Hilbert space for distinct measurements. We shall consider this bound in §4.

3. Optimal entropic uncertainty relation for successive measurements

All the uncertainty relations discussed in §2, refer to a situation where the observables A, B are measured on two distinct but identically prepared ensembles of systems in state ρ . We shall now consider a situation where the observables A, B are measured sequentially on the same ensemble of systems in state ρ . Then the joint probability $\Pr_{A,B}^\rho(a_i, b_j)$ that the outcomes are a_i, b_j respectively, is given by the Wigner formula (4); and the probabilities $\Pr_{A,B}^\rho(a_i), \Pr_{A,B}^\rho(b_j)$ for obtaining different outcomes in the A and B measurements, are the marginals of the above joint probability and are given by

$$\Pr_{A,B}^\rho(a_i) = \text{Tr}[\rho P^A(a_i)], \quad (25a)$$

$$\Pr_{A,B}^\rho(b_j) = \text{Tr} \left[\sum_i (P^A(a_i) \rho P^A(a_i)) P^B(b_j) \right] = \text{Tr}[\varepsilon(\rho) P^B(b_j)], \quad (25b)$$

where, as in (6), $\varepsilon(\rho)$ is given by

$$\varepsilon(\rho) = \sum_i P^A(a_i) \rho P^A(a_i). \quad (26)$$

Here, we may note the important feature of quantum theory that, unless $\varepsilon(\rho) = \rho$, we have

$$\Pr_{A,B}^\rho(b_j) \neq \text{Tr}[\rho P^B(b_j)]. \quad (27)$$

Equation (27) shows that the probability (25b) that the outcome b_j is realised in a B -measurement, when an ensemble of systems prepared in state ρ is subjected to the sequence of measurements A, B , is different from the probability $\text{Tr}[\rho P^B(b_j)]$ that the same outcome is realised when there is no intervening measurement of observable A prior to the measurement of B . This important feature of quantum theory has been referred to as the ‘quantum interference of probabilities’ by de Broglie [21] (see also [6,22]) and clearly expresses how a prior measurement influences the probability distributions of the outcomes of later measurements.

The uncertainties in the outcomes of A and B , in a sequential measurement, are given by the information-theoretic entropies associated with the probability distributions (25a) and (25b)

$$S_{A,B}^\rho(A) = - \sum_i \Pr_{A,B}^\rho(a_i) \log \Pr_{A,B}^\rho(a_i), \quad (28a)$$

$$S_{A,B}^\rho(B) = - \sum_j \Pr_{A,B}^\rho(b_j) \log \Pr_{A,B}^\rho(b_j). \quad (28b)$$

We have the standard properties of the eigenprojectors of the observable A ,

$$P^A(a_i)P^A(a_j) = \delta_{ij}P^A(a_i), \quad (29a)$$

$$\sum_i P^A(a_i) = I \quad (29b)$$

we can show the following:

$$S_{A,B}^\rho(A) = S^\rho(A) = S^{\varepsilon(\rho)}(A), \quad (30a)$$

$$S_{A,B}^\rho(B) = S^{\varepsilon(\rho)}(B), \quad (30b)$$

where $\varepsilon(\rho)$ is given by (26). From (30b) it follows that the entropies $S_{A,B}^\rho(B)$ and $S^\rho(B)$ are in general different unless $\varepsilon(\rho) = \rho$, which happens only when ρ commutes with all the eigenprojectors of A .

The optimal uncertainty relation for successive measurements can be expressed in the form

$$S_{A,B}^\rho(A) + S_{A,B}^\rho(B) \geq \Lambda_S(A, B), \quad (31)$$

where the optimum bound is given by

$$\Lambda_S(A, B) = \inf_\rho [S_{A,B}^\rho(A) + S_{A,B}^\rho(B)]. \quad (32)$$

From (30a) and (30b), we obtain

$$\Lambda_S(A, B) = \inf_{\varepsilon(\rho)} [S^{\varepsilon(\rho)}(A) + S^{\varepsilon(\rho)}(B)]. \quad (33)$$

The map, $\rho \rightarrow \varepsilon(\rho)$, given by (26) maps the class of all density operators into a proper subset of itself, namely the set of all those density operators which commute with all the eigenprojectors $\{P^A(a_i)\}$ of A (see for instance ([6], chapter 7)). These are precisely those density operators, which are expressible as mixtures of eigenstates of A . Thus the infimum on the RHS of (33) is over a much smaller class of density operators (namely those of the form $\varepsilon(\rho)$) than the set of all density operators which occurs in the RHS of (15). Thus, we are led to the inequality

$$\Lambda_S(A, B) \geq \Lambda_D(A, B), \quad (34)$$

which shows that the optimal bound on the sum of uncertainties for successive measurements of two observables is always greater than or equal to the optimal bound for sum of uncertainties when the same observables are measured on distinct ensemble of systems. While there is no general constructive expression for the latter, we shall show that the optimal bound $\Lambda_S(A, B)$ can be expressed in terms of the eigenstates of the observables A, B , when they have totally non-degenerate spectra.

A lower bound on $\Lambda_S(A, B)$ was obtained some time ago [10] making use of the joint entropy $S_{A,B}^\rho(A, B)$, defined in terms of the joint probabilities (4):

Influence of measurements on uncertainties

$$S_{A,B}^\rho(A,B) = - \sum_{i,j} \Pr_{A,B}^\rho(a_i, b_j) \log \Pr_{A,B}^\rho(a_i, b_j). \quad (35)$$

Since the probabilities $\Pr_{A,B}^\rho(a_i)$ and $\Pr_{A,B}^\rho(b_j)$, given by (25a), (25b), are marginals of the joint probability (4), we have the sub-additivity inequality:

$$S_{A,B}^\rho(A) + S_{A,B}^\rho(B) \geq S_{A,B}^\rho(A,B). \quad (36)$$

Incidentally, we also have the inequalities,

$$S_{A,B}^\rho(A,B) \geq S_{A,B}^\rho(A) = S^\rho(A), \quad (37a)$$

$$S_{A,B}^\rho(A,B) \geq S_{A,B}^\rho(B), \quad (37b)$$

which, along with (36), can be used to define conditional entropies and mutual information, for sequential measurements of two observables in quantum theory.

Now, from equations (35) and (4) which define the joint entropy and the joint probabilities, we can easily obtain the inequality

$$S_{A,B}^\rho(A,B) \geq \log \left[1 / \sup_{i,j} \| P^A(a_i) P^B(b_j) P^A(a_i) \| \right]. \quad (38)$$

From (36) and (38), we obtain the uncertainty relation

$$S_{A,B}^\rho(A) + S_{A,B}^\rho(B) \geq \log \left[1 / \sup_{i,j} \| P^A(a_i) P^B(b_j) P^A(a_i) \| \right], \quad (39)$$

where the RHS is the same as the Krishna–Parthasarathy bound on the sum of uncertainties for distinct measurements.

We can obtain a much stronger bound than (39) for the sum of uncertainties for successive measurements. Our result on the optimal uncertainty bound for successive measurements is contained in the following Theorem 1.

Theorem 1. *The optimal bound $\Lambda_S(A,B)$ on the sum of entropic uncertainties of two observables A, B , with purely discrete spectra, when they are measured sequentially on the same ensemble of systems, is given by*

$$\Lambda_S(A,B) = \inf_i \inf_{P^A(a_i)|\psi\rangle=|\psi\rangle} - \sum_j \langle \psi | P^B(b_j) | \psi \rangle \log \langle \psi | P^B(b_j) | \psi \rangle, \quad (40a)$$

where the states $|\psi\rangle$ are assumed to be normalized. When the observable A has totally non-degenerate spectrum, eq. (40a) reduces to

$$\Lambda_S(A,B) = \inf_i - \sum_j \langle a_i | P^B(b_j) | a_i \rangle \log \langle a_i | P^B(b_j) | a_i \rangle. \quad (40b)$$

When both A and B have totally non-degenerate spectra, the optimal lower bound further reduces to

M D Srinivas

$$\Lambda_S(A, B) = \inf_i - \sum_j |\langle a_i | b_j \rangle|^2 \log |\langle a_i | b_j \rangle|^2. \quad (40c)$$

Proof. We start with the expression for the optimum bound $\Lambda_S(A, B)$, given by (33), where the infimum is to be taken only over states of the form $\varepsilon(\rho)$, given by (26). Since the entropies $S^{\varepsilon(\rho)}(A)$ and $S^{\varepsilon(\rho)}(B)$, are concave functions of ρ , we need to take the infimum in (33) only over pure states which occur in the decomposition of $\varepsilon(\rho)$. Since, as we have already noted, states of the form $\varepsilon(\rho)$ are mixtures of eigenstates of A , we need to take infimum in (33) only over the eigenstates of A . In each eigenstate $|\psi\rangle$ of A , the entropy $S^{|\psi\rangle}(A)$ vanishes and therefore eq. (33) reduces to

$$\Lambda_S(A, B) = \inf_i \inf_{P^A(a_i)|\psi\rangle=|\psi\rangle} S_{A,B}^{|\psi\rangle}(B). \quad (41)$$

Now, if we employ (28b) for the entropy $S_{A,B}^{|\psi\rangle}(B)$ in eq. (41), we are immediately led to (40a) for the optimal bound. Equations (40b) and (40c) are direct consequences of (40a) when the spectra of A and B turn out to be totally non-degenerate. This completes the proof of Theorem 1.

From (40c) it follows that for observables with non-degenerate spectra

$$\Lambda_S(A, B) \geq \log \left[1 / \sup_{i,j} |\langle a_i | b_j \rangle|^2 \right]. \quad (42)$$

The fact that the optimal uncertainty bound for successive measurements is greater than the Maassen–Uffink bound for distinct measurements could have anyway been inferred from eqs (34) and (22).

The bound appearing on the RHS of (40c), for the sum of uncertainties in successive measurements, has also been obtained in a recent investigation by Cerf and Adami [23]. They have also noted the important fact that this bound is greater than the Maassen–Uffink bound for distinct measurements. Our derivation above shows that the RHS of (40c) actually gives the optimal bound on the sum of uncertainties in the outcomes of successive measurements of observables with totally non-degenerate spectra.

From the above derivation it also follows that $S_{A,B}^\rho(A) + S_{A,B}^\rho(B)$ attains its infimum in eigenstates of observable A , states in which $S_{A,B}^\rho(A)$ vanishes. Therefore, the optimal lower bound of $S_{A,B}^\rho(A) + S_{A,B}^\rho(B)$ is also the optimal lower bound of $S_{A,B}^\rho(B)$, or equivalently,

$$S_{A,B}^\rho(B) \geq \Lambda_S(A, B). \quad (43)$$

In particular, when A, B have totally non-degenerate spectra,

$$S_{A,B}^\rho(B) \geq \inf_i - \sum_j |\langle a_i | b_j \rangle|^2 \log |\langle a_i | b_j \rangle|^2. \quad (43a)$$

When the RHS of (43) is non-zero, which invariably happens when A, B do not have a common eigenvector, the inequality (43) gives a non-trivial lower bound on the uncertainty in the outcome of B -measurement, which arises essentially because the ensemble of systems prepared in state ρ , has been first subjected to an A -measurement. Had there been

no such intervening measurement on the ensemble, then the uncertainty $S^\rho(B)$ in the outcome of B -measurement can be made arbitrarily small, in fact zero by choosing the initial state ρ to be a mixture of eigenstates of B that are associated with the same eigenvalue. Equation (43) shows that whenever A, B do not have a joint eigenstate, the outcome of a B -measurement which follows an A -measurement is always uncertain whatever be the initial state of the system, and the optimal lower bound on this uncertainty is again given by the RHS of (40a)–(40c).

Using the method outlined above, we can obtain optimal bounds on the sum of entropic uncertainties of any arbitrary sequence of measurements, provided we restrict ourselves to observables with purely discrete spectra. For instance, if C is another observable with spectral decomposition

$$C = \sum_k c_k P^C(c_k), \quad (44)$$

then the joint probability $\Pr_{A,B,C}^\rho(a_i, b_j, c_k)$, that the outcomes a_i, b_j, c_k , are realised in a sequential measurement of A, B, C , is given by the Wigner formula:

$$\Pr_{A,B,C}^\rho(a_i, b_j, c_k) = \text{Tr}[P^C(c_k)P^B(b_j)P^A(a_i)\rho P^A(a_i)P^B(b_j)P^C(c_k)]. \quad (45)$$

From (45), we obtain the joint entropy

$$S_{A,B,C}^\rho(A, B, C) = - \sum_{i,j,k} \Pr_{A,B,C}^\rho(a_i, b_j, c_k) \log \Pr_{A,B,C}^\rho(a_i, b_j, c_k). \quad (46)$$

The uncertainty in the outcomes of the C -measurement, in a sequential measurement of A, B, C , is given by

$$S_{A,B,C}^\rho(C) = - \sum_k \Pr_{A,B,C}^\rho(c_k) \log \Pr_{A,B,C}^\rho(c_k). \quad (47)$$

Using the standard properties of spectral projectors of A and B , we can show that

$$S_{A,B,C}^\rho(C) = S^{\Gamma^\varepsilon(\rho)}(C), \quad (48)$$

where, ε is given by eq. (26) and Γ is given by

$$\Gamma(\rho) = \sum_j P^B(b_j)\rho P^B(b_j) \quad (49)$$

From the fact that the probability distributions of the observables A, B, C , are marginals of the joint probability $\Pr_{A,B,C}^\rho(a_i, b_j, c_k)$ given by (45), we can deduce the sub-additivity and the strong sub-additivity properties

$$S_{A,B,C}^\rho(A) + S_{A,B,C}^\rho(B) + S_{A,B,C}^\rho(C) \geq S_{A,B,C}^\rho(A, B, C), \quad (50)$$

$$S_{A,B,C}^\rho(A, B) + S_{A,B,C}^\rho(B, C) \geq S_{A,B,C}^\rho(A, B, C) + S_{A,B,C}^\rho(B), \quad (51)$$

in the same manner as in classical information theory.

The optimal entropic uncertainty relation for the successive measurement of three observables A, B, C , is given by

$$S_{A,B,C}^\rho(A) + S_{A,B,C}^\rho(B) + S_{A,B,C}^\rho(C) \geq \Lambda_S(A, B, C), \quad (52)$$

where

$$\Lambda_S(A, B, C) = \inf_{\rho} [S_{A,B,C}^\rho(A) + S_{A,B,C}^\rho(B) + S_{A,B,C}^\rho(C)]. \quad (53)$$

From (30a), (30b) and (48), we obtain

$$\Lambda_S(A, B, C) = \inf_{\varepsilon(\rho)} [S^{\varepsilon(\rho)}(A) + S^{\varepsilon(\rho)}(B) + S^{\Gamma\varepsilon(\rho)}(C)]. \quad (54)$$

Again, in the RHS of eq. (54), we need to take infimum only over states of the form $\varepsilon(\rho)$, and following the line of argument in the proof of Theorem 1, we can easily establish the following Theorem 2, where for the sake of simplicity we give only the result for observables with totally non-degenerate spectra.

Theorem 2. *The optimal bound $\Lambda_S(A, B, C)$ on the sum of entropic uncertainties of three observables A, B, C , with purely discrete and totally non-degenerate spectra, which are sequentially measured on the same ensemble of systems, is given by*

$$\begin{aligned} \Lambda_S(A, B, C) = & \inf_i - \sum_j |\langle a_i | b_j \rangle|^2 \log |\langle a_i | b_j \rangle|^2 \\ & + \inf_i - \sum_k \left\{ \sum_j [|\langle a_i | b_j \rangle|^2 |\langle b_j | c_k \rangle|^2] \right. \\ & \left. \times \log \left[\sum_j |\langle a_i | b_j \rangle|^2 |\langle b_j | c_k \rangle|^2 \right] \right\}. \end{aligned} \quad (55)$$

There are two terms in the optimal bound equation (55). The first is nothing but the lower bound $\Lambda_S(A, B)$ that was derived earlier in eq. (40c), for the sequential measurement of the two observables, A, B ; it is also the optimal lower bound (43a) on the uncertainty in the outcomes of the B -measurement, $S_{A,B}^\rho(B) = S_{A,B,C}^\rho(B)$ in a sequential measurement of A, B, C . The second term (55) can be easily shown to be the optimal bound on the uncertainty in the outcomes of C -measurement when a sequential measurement of A, B, C is carried out. In other words, we have

$$\begin{aligned} S_{A,B,C}^\rho(C) \geq & \inf_i - \sum_k \left\{ \sum_j [|\langle a_i | b_j \rangle|^2 |\langle b_j | c_k \rangle|^2] \right. \\ & \left. \times \log \left[\sum_j |\langle a_i | b_j \rangle|^2 |\langle b_j | c_k \rangle|^2 \right] \right\}. \end{aligned} \quad (56)$$

Since the matrix, $U_{jk} = |\langle b_j | c_k \rangle|^2$, is a doubly stochastic matrix, it follows [9] that the lower bound on $S_{A,B,C}^\rho(C)$ given by the RHS of (56) is greater than or equal to the lower bound on $S_{A,B}^\rho(B)$ given by the RHS of (43a). The above results (55), (56) can easily be extended to arbitrary sequence of measurements of observables with purely discrete and totally non-degenerate spectra.

4. Optimal uncertainty relations in two-dimensional Hilbert space

We have seen eq. (34) that the optimal bound $\Lambda_S(A, B)$ on the sum of uncertainties of observables A, B , when they are measured sequentially on the same ensemble of systems, is greater than or equal to the optimal bound $\Lambda_D(A, B)$ when they are measured on distinct but identically prepared ensemble of systems. However, the question still remains as to whether $\Lambda_S(A, B)$ is strictly greater than $\Lambda_D(A, B)$ for at least some pair of observables A, B . In order to decide this issue we need to consider a situation for which the optimal bound $\Lambda_D(A, B)$ has been determined for a large enough class of observables.

Recently Sanchez-Ruiz [20] has shown that the uncertainty bounds derived sometime ago by Garrett and Gull [19] for observables in a two-dimensional Hilbert space are in fact optimal. We shall utilise this bound to show that $\Lambda_S(A, B)$ is indeed strictly greater than $\Lambda_D(A, B)$ for a large class of observables.

Let A, B refer to the components of spin, for a spin- $\frac{1}{2}$ system, along the directions $\mathbf{n}_1, \mathbf{n}_2$, i.e. $A = \boldsymbol{\sigma} \cdot \mathbf{n}_1$ and $B = \boldsymbol{\sigma} \cdot \mathbf{n}_2$. If $\mathbf{n}_1 \cdot \mathbf{n}_2 = \cos \theta$, we have

$$|\langle a_1 | b_1 \rangle|^2 = |\langle a_2 | b_2 \rangle|^2 = \cos^2 \theta / 2, \quad (57a)$$

$$|\langle a_1 | b_2 \rangle|^2 = |\langle a_2 | b_1 \rangle|^2 = \sin^2 \theta / 2. \quad (57b)$$

The optimal bound on the sum of uncertainties of $\boldsymbol{\sigma} \cdot \mathbf{n}_1$ and $\boldsymbol{\sigma} \cdot \mathbf{n}_2$, when they are sequentially measured on the same ensemble of systems, as given by eq. (40c), becomes

$$\Lambda_S(\boldsymbol{\sigma} \cdot \mathbf{n}_1, \boldsymbol{\sigma} \cdot \mathbf{n}_2) = \Lambda_S(\theta) = -\cos^2 \theta / 2 \log \cos^2 \theta / 2 - \sin^2 \theta / 2 \log \sin^2 \theta / 2, \quad (58)$$

which is nothing but the entropic uncertainty of $\boldsymbol{\sigma} \cdot \mathbf{n}_2$ in the eigenstates of $\boldsymbol{\sigma} \cdot \mathbf{n}_1$.

When $\boldsymbol{\sigma} \cdot \mathbf{n}_1$ and $\boldsymbol{\sigma} \cdot \mathbf{n}_2$ are measured on distinct but identically prepared ensembles of systems, the uncertainty bound (17) obtained by Deutsch becomes

$$\Lambda_D(\boldsymbol{\sigma} \cdot \mathbf{n}_1, \boldsymbol{\sigma} \cdot \mathbf{n}_2) \geq \Lambda_{D1}(\theta) = 2 \log(2 / [1 + \max\{|\cos \theta / 2|, |\sin \theta / 2|\}]). \quad (59)$$

The stronger bound (22) obtained by Maassen and Uffink becomes

$$\Lambda_D(\boldsymbol{\sigma} \cdot \mathbf{n}_1, \boldsymbol{\sigma} \cdot \mathbf{n}_2) \geq \Lambda_{D2}(\theta) = 2 \log(1 / \max\{\cos^2 \theta / 2, \sin^2 \theta / 2\}). \quad (60)$$

Recently, Sanchez-Ruiz [20] has shown that the optimum bound on the sum of uncertainties of $\boldsymbol{\sigma} \cdot \mathbf{n}_1$ and $\boldsymbol{\sigma} \cdot \mathbf{n}_2$, when they are measured on distinct but identically prepared ensembles of systems is given by the following:

$$\Lambda_D(\boldsymbol{\sigma} \cdot \mathbf{n}_1, \boldsymbol{\sigma} \cdot \mathbf{n}_2) = \Lambda_D(\theta) = -2 \cos^2 \theta / 4 \log \cos^2 \theta / 4 - 2 \sin^2 \theta / 4 \log \sin^2 \theta / 4, \quad (61)$$

when $0 \leq \theta \leq \theta^*$, where θ^* is given by the transcendental equation

$$\cos(\theta^* / 2) \log[(1 + \cos \theta^* / 2) / (1 - \cos \theta^* / 2)] = 2, \quad (62)$$

Table 1. Numerical estimates of different uncertainty bounds.

θ°	$\Lambda_S(\theta)$	$\Lambda_D(\theta)$	$\Lambda_{D2}(\theta)$	$\Lambda_{D1}(\theta)$
0	0.000	0.000	0.000	0.000
10	0.045	0.028	0.008	0.004
20	0.135	0.089	0.031	0.015
30	0.246	0.173	0.069	0.034
40	0.361	0.271	0.124	0.061
50	0.469	0.378	0.197	0.096
60	0.562	0.492	0.288	0.139
70	0.633	0.604	0.399	0.190
80	0.678	0.673	0.533	0.249
90	0.693	0.693	0.693	0.317

Note: All logarithms are calculated to base e .

which implies $\theta^* \approx 67^\circ$. Equation (61) shows that the sum of uncertainties in $\boldsymbol{\sigma} \cdot \mathbf{n}_1$ and $\boldsymbol{\sigma} \cdot \mathbf{n}_2$, takes on the infimum value in the eigenstates of $\boldsymbol{\sigma} \cdot (\mathbf{n}_1 + \mathbf{n}_2)$ when $0 \leq \theta \leq \theta^*$. When $\pi - \theta^* \leq \theta \leq \pi$, the optimum bound is

$$\begin{aligned} \Lambda_D(\boldsymbol{\sigma} \cdot \mathbf{n}_1, \boldsymbol{\sigma} \cdot \mathbf{n}_2) = \Lambda_D(\theta) = & -\cos^2(\pi/4 + \theta/4) \log \cos^2(\pi/4 + \theta/4) \\ & -\sin^2(\pi/4 + \theta/4) \log \sin^2(\pi/4 + \theta/4) \\ & -\cos^2(\pi/4 - \theta/4) \log \cos^2(\pi/4 - \theta/4) \\ & -\sin^2(\pi/4 - \theta/4) \log \sin^2(\pi/4 - \theta/4). \end{aligned} \quad (63)$$

In this case, the optimum bound is attained in the eigenstates of $\boldsymbol{\sigma} \cdot (\mathbf{n}_1 - \mathbf{n}_2)$. When $\theta^* \leq \theta \leq \pi - \theta^*$, the optimal bound $\Lambda_D(\theta)$ seems to have no analytical expression and has to be evaluated numerically.

Numerical values of $\Lambda_{D1}(\theta)$ (the Deutsch bound), $\Lambda_{D2}(\theta)$ (the Maassen–Uffink bound) and the optimum bound $\Lambda_D(\theta)$, for the case of distinct measurements, have been tabulated by Garrett and Gull [19]. In table 1, we give these values along with the numerical values of the optimum bound $\Lambda_S(\theta)$ for successive measurements given by eq. (58). These values are also plotted in figure 1.

From table 1, we see that

$$\Lambda_S(\theta) \geq \Lambda_D(\theta) \geq \Lambda_{D2}(\theta) \geq 2\Lambda_{D1}(\theta), \quad (64)$$

where, equality holds only when $\theta = 0, \pi/2$. Hence, in a two-dimensional Hilbert space, the optimal bound $\Lambda_S(\boldsymbol{\sigma} \cdot \mathbf{n}_1, \boldsymbol{\sigma} \cdot \mathbf{n}_2)$ for successive measurements of $\boldsymbol{\sigma} \cdot \mathbf{n}_1$ and $\boldsymbol{\sigma} \cdot \mathbf{n}_2$, is strictly greater than the optimal bound $\Lambda_D(\boldsymbol{\sigma} \cdot \mathbf{n}_1, \boldsymbol{\sigma} \cdot \mathbf{n}_2)$ when the observables $\boldsymbol{\sigma} \cdot \mathbf{n}_1, \boldsymbol{\sigma} \cdot \mathbf{n}_2$ are measured on distinct ensembles, except when $\mathbf{n}_1, \mathbf{n}_2$ are either parallel or perpendicular, i.e., when the observables $\boldsymbol{\sigma} \cdot \mathbf{n}_1, \boldsymbol{\sigma} \cdot \mathbf{n}_2$ either commute or are complementary.

5. Conclusion

We have seen that there are actually two formulations of the uncertainty principle in quantum theory: one for the case of distinct measurements performed on two distinct but identically prepared ensembles of systems; and another for the case of successive measurements

Influence of measurements on uncertainties

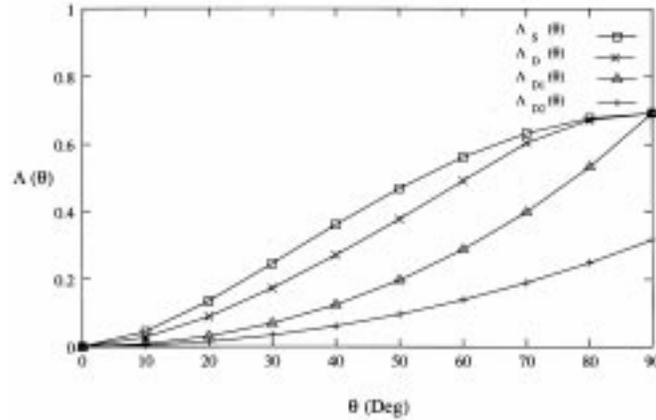


Figure 1. Numerical estimates of different uncertainty bounds.

performed on the same ensemble of systems. Consequently there are two classes of uncertainty relations, which correspond to these two different experimental situations.

The standard variance form of the uncertainty relation due to Robertson, or the entropic form uncertainty relations due to Everett and Deutsch and improved upon by later researchers, all deal with measurements performed on distinct but identically prepared ensembles of systems. It is now well-understood that these uncertainty relations essentially reflect the limitations imposed by quantum theory on the preparation of ensembles of systems in identical states. They do not in any way express the influence of measurement of one observable on the uncertainty in the outcome of another, contrary to what seems to have been envisaged by Heisenberg in his pioneering work on the uncertainty principle.

On the other hand, a different class of uncertainty relations can be formulated for the case when a set of observables is sequentially measured on the same ensemble of systems. These uncertainty relations clearly reveal the influence of measurement of one observable on the uncertainties in the outcomes of later measurements. The fact that the optimal lower bound on the sum of uncertainties of successive measurements turns out to be in general not less than and often strictly greater than the optimal bound on the corresponding sum of uncertainties for measurements performed on distinct ensembles of systems, shows that in quantum theory, measurements do have a profound influence on the uncertainties in the outcomes of all subsequent observations.

As regards, the particular physical situation discussed by Heisenberg in his pioneering paper, viz. the measurement of position followed by that of momentum or vice versa, this has remained outside the purview of the standard mathematical formulation of quantum theory because of the difficulties associated with the extension of the conventional collapse postulate for observables such as position and momentum, which have continuous spectra. Various investigations ([6], chapters 7,8) have, however, shown that any meaningful generalization of the collapse postulate would necessarily imply that the state of a system immediately after a precise measurement of position would be such that its momentum distribution is entirely concentrated at $\pm\infty$, and vice versa. This is perhaps the extent to which Heisenberg's intuitive understanding of the quantum uncertainty principle can be vindicated on the basis of the mathematical formalism of quantum mechanics.

Acknowledgement

The author is grateful to K Ramasubramanian for helping him with the graph.

References

- [1] W Heisenberg, *Z. Phys.* **43**, 172 (1927)
English translation cited from *Quantum theory and measurement* edited by J A Wheeler and W H Zurek (University Press, Princeton, 1983) p. 64
For another English version of the same passage, see, M Jammer, *The philosophy of quantum mechanics* (Wiley, New York, 1974) p. 63
- [2] H P Robertson, *Phys. Rev.* **34**, 163 (1929); reprinted in Wheeler and Zurek cited in [1], pp. 127–8
- [3] M A Nielsen and I L Chuang, *Quantum computation and quantum information* (University Press, Cambridge, 2000) p. 89
- [4] D Home, *Conceptual foundations of quantum physics* (Plenum, New York, 1997) pp. 13–14
- [5] E P Wigner, *Am. J. Phys.* **31**, 6 (1963)
See also E P Wigner, in *Foundations of quantum mechanics* edited by B d'Espagnat (Academic, New York, 1971) pp. 1–16
- [6] M D Srinivas, *Measurements and quantum probabilities* (Universities Press, Hyderabad, 2001)
- [7] B Gnanapragasam and M D Srinivas, *Pramana – J. Phys.* **12**, 699 (1979)
- [8] D Deutsch, *Phys. Rev. Lett.* **50**, 631 (1983)
- [9] R J McElice, *The theory of information and coding* (Addison Wesley, Boston, 1977)
T M Cover and J A Thomas, *Elements of information theory* (Wiley, New York, 1991)
- [10] M D Srinivas, *Pramana – J. Phys.* **24**, 673 (1985)
- [11] H Everett, *The many world interpretation of quantum mechanics* edited by B S De Witt and N Graham (University Press, Princeton, 1973)
- [12] I Hirschman, *Am. J. Math.* **79**, 152 (1957)
- [13] W Beckner, *Ann. Math.* **102**, 159 (1975)
- [14] I Bialynicki-Birula and J Mycielski, *Commun. Math. Phys.* **44**, 129 (1975)
- [15] M H Partovi, *Phys. Rev. Lett.* **50**, 1883 (1983)
- [16] K Kraus, *Phys. Rev.* **D35**, 3070 (1987)
- [17] H Maassen and J B M Uffink, *Phys. Rev. Lett.* **60**, 1103 (1988)
- [18] M Krishna and K R Parthasarathy, *Sankhya* **A64**, 842 (2002)
This paper also derives an entropic uncertainty relation for more general class of observables given by discrete positive-operator-valued measures
- [19] A J M Garrett and S F Gull, *Phys. Lett.* **A151**, 453 (1990)
- [20] J Sanches-Ruiz, *Phys. Lett.* **A244**, 189 (1998)
- [21] L de Broglie, *Rev. Sci.* (March 1948) p. 259
L de Broglie, *Physics and microphysics* (Hutchinson, London, 1955)
- [22] J J Sakurai, *Modern quantum mechanics*, 2nd ed. (Addison-Wesley, Reading, Massachusetts, 1994) pp. 32–4
- [23] N J Cerf and C Adami, *Quantum mechanics of measurement*, e-print: quant-ph/9605002 v2 (1997)