Investigation into Variance Based Uncertainty Relations in the Presence of Quantum Memory

Mannathu Gopikrishnan

A dissertation submitted for the partial fulfilment of

BS-MS dual degree in Science

Indian Institute of Science Education and Research, Mohali

June 2020

Certificate of Examination

This is to certify that the dissertation titled "**Investigation into Variance Based Uncertainty Relations in the Presence of Quantum Memory**" submitted by Mr. Mannathu Gopikrishnan (Reg No. MS15074) for the partial fulfilment of BS-MS dual degree programme of the Institute, has been examined by the thesis committee duly appointed by the Institute. The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Dr Sandjib Dey   Dr Anosh Joseph   Dr Manabendra Nath Bera
(Supervisor)

Dated: 15 June 2020

Declaration

The work presented in this dissertation has been carried out by me under the guidance of Dr Manabendra Nath Bera at the Indian Institute of Science Education and Research Mohali. This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in the bibliography.

Mannathu Gopikrishnan
(Candidate)

Dated: 15 June 2020

In my capacity as supervisor of the candidate's project work, I certify that the above statements by the candidate are true to the best of my knowledge.

Dr Manabendra Nath Bera
(Supervisor)

Acknowledgement

Working on my thesis has been a most enlightening experience and one I definitely could not have completed without generous assistance from various quarters.

First and foremost, I thank my guide Dr Manabendra Nath Bera for his infinite patience and willingness to provide help with all aspects of the project. His insights are second to none. I also thank my thesis committee members Dr Anosh Joseph and Dr Sanjib Dey for evaluating my work and providing valuable feedback.

I am most grateful to my parents for funding my time in college, and all the wonderful people at IISER Mohali and around it, who have made the last 5 years lively and pass by in the blink of an eye.

List of Figures

Contents

Abstract

The following work is an attempt to formulate a variance based uncertainty relation that aims to capture a reduction in uncertainty associated with the action of non-commuting observables on a quantum system. This reduction in uncertainty (for bipartite entangled systems) is brought about due to the conditioning provided by correlations between the sub-component of the system on which the observables act, and the other sub-component entangled with it, which we shall refer to as the 'memory'.

Although we will be attempting to formulate a relation based on variance, as preliminary material, we shall develop the theory behind uncertainty relations and entropy, both of which play a pivotal role in understanding our work.

The existing relation that explores the domain of our work has been formulated in terms of conditional entropies, but since variance is a much more user friendly quantity to compute, we believe that if successful, this work will add considerable value.

Chapter 1

Quantum Mechanics

1.1 The Basics

We need to start somewhere on the journey to a new uncertainty relation, and the best (and only) place to start is with the postulates of quantum mechanics, then move on to the density matrix formulation of these postulates, which would enable us to get a much better vantage point over quantum mechanics.
The basic postulates are as follows [11]:

Postulate 1

All the information about a quantum mechanical system is contained in its *State*, which is represented by a vector of unit length in *Hilbert Space*.
The Hilbert Space needs to satisfy the following criteria [20]:

(i) It is a vector space over the set of *Complex Numbers*.

(ii) There must be defined, an inner product $\langle\psi|\phi\rangle$ which can map vectors to the set of Complex Numbers.

(iii) It is complete in the norm.

Postulate 2

The time evolution of a closed quantum system is given by a unitary transformation, which arises as a solution to the Time Dependent Schrödinger Equation.

$$i\hbar\frac{d\,|\psi\rangle}{dt} = \hat{H}\,|\psi\rangle\,, \tag{1.1}$$

solving which, we get:

$$|\psi(t)\rangle = e^{-iHt/\hbar} |\psi(0)\rangle . \tag{1.2}$$

Thus, the evolution of a closed quantum system is given by the operator $e^{-iHt/\hbar}$. Now, if we take the product of the time evolution operator with its Hermitian conjugate like so:

$$e^{-iHt/\hbar} e^{iHt/\hbar} = \mathbb{1}, \tag{1.3}$$

then such an operator, which conforms to the condition $UU^\dagger = U^\dagger U = \mathbb{1}$ is termed a *unitary operator*. These are essentially transformations in the Hilbert space which preserve the lengths (the probabilities) of vectors and hence describe the evolution of all such closed quantum systems.

Postulate 3

Measurements on quantum systems are described by a set of *measurement operators* $\{M_m\}$. (The *m* refers to the many outcomes which may occur during the measurement process).

The probability of outcome *m* occurring is given by:

$$p(m) = \langle\psi| M_m^\dagger M_m |\psi\rangle . \tag{1.4}$$

The state of the system after performing this measurement is given by:

$$|\psi'\rangle = \frac{M_m |\psi\rangle}{\sqrt{\langle\psi| M_m^\dagger M_m |\psi\rangle}}. \tag{1.5}$$

The probabilities of all possible outcomes must sum to 1.

$$\sum_m p(m) = \sum_m \langle\psi| M_m^\dagger M_m |\psi\rangle = 1. \tag{1.6}$$

This leads us to the *completeness relation*

$$\sum_m M_m^\dagger M_m = \mathbb{1}, \tag{1.7}$$

which every set of measurement operators must satisfy in order to qualify as one.

We may note at this point, that *projective operators* $\{P_m\}$, describing *observables*, which are physical quantities that can be measured, are a special class of measurement operators. These, in addition to satisfying the above criteria for measurement operators in general, must also satisfy the following criterion:

$$P_m P_{m'} = \delta_{m'm} P_m. \tag{1.8}$$

Upon acting on the state $|\psi\rangle$, the probability of getting the outcome *m* is given by

$$p(m) = \langle\psi| P_m |\psi\rangle . \tag{1.9}$$

The state immediately after measurement becomes:

$$|\psi'\rangle = \frac{P_m |\psi\rangle}{\sqrt{p(m)}}. \tag{1.10}$$

These observables have a spectral decomposition given by

$$M = \sum_m m P_m, \tag{1.11}$$

where P_m is the projector onto the eigenspace of M with eigenvalues m, and the possible *outcomes* of the measurement correspond to these eigenvalues m.

1.2 Prelude to the Density Matrix: The Qubit and Composite Systems

1.2.1 The Qubit

The qubit is the simplest quantum system, and the most important one, as it is the testbed upon which we shall illustrate concepts which underlie all of quantum mechanics.

A qubit is comparable to the classical *binary bit* in that, just as the bit is a 2 level classical system, the qubit also represents a 2 level system in quantum mechanics.

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle . \tag{1.12}$$

But that is where the similarity ends, because while the classical bit is *always* in one of the two states it can occupy, the same cannot be said of the qubit. The qubit, as long as it is isolated (protected from any sort of interaction with its environment), lives in a *superposition* (i.e., a convex linear combination) of its basis states.

Once a qubit is subjected to a measurement though, it collapses to occupy one of its two basis states, which here happens to be either $|0\rangle$ or $|1\rangle$. Upon repeating this collapse for a large number of identical qubits, it may be seen that the statistical probabilities of the basis states to which the qubits collapse tend to the following values:

$$p(|0\rangle) = |\alpha|^2 \qquad\qquad p(|1\rangle) = |\beta|^2 \tag{1.13}$$

1.2.2 Composite Systems

To illustrate this, we consider the simplest case involving two qubits. The first qubit belongs to Hilbert Space A ($\mathcal{H}_A$), and the second belongs to Hilbert Space B ($\mathcal{H}_B$).
We represent the qubit in the space $\mathcal{H}_A$ as:

$$|\psi\rangle_A = a_1 |0\rangle_A + a_2 |1\rangle_A \,,$$

and in the matrix form, represent it as:

$$|\psi\rangle_A = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

$$since \qquad |0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad and \qquad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Similarly, the qubit in the space $\mathcal{H}_B$ can be written as:

$$|\psi\rangle_B = b_1 |0\rangle_B + b_2 |1\rangle_B \,,$$

and in the matrix form, as:

$$|\psi\rangle_B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Now, the composite system will belong to the space $\mathcal{H}_{AB}$, where

$$\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$$

and the *composite state* $|\psi\rangle_{AB}$, will be given by

$$|\psi\rangle_{AB} = |\psi\rangle_A \otimes |\psi\rangle_B \,.$$

It may be represented in the matrix form as:

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \otimes \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} a_1 b_1 \\ a_1 b_2 \\ a_2 b_1 \\ a_2 b_2 \end{bmatrix}$$

If we prefer long equations, it may also be represented as:

$$|\psi\rangle_{AB} = a_1 b_1 |0\rangle_A \otimes |0\rangle_B + a_1 b_2 |0\rangle_A \otimes |1\rangle_B + a_2 b_1 |1\rangle_A \otimes |0\rangle_B + a_2 b_2 |1\rangle_A \otimes |1\rangle_B ,$$

or less tediously, as:

$$|\psi\rangle_{AB} = a_1 b_1 |00\rangle + a_1 b_2 |01\rangle + a_2 b_1 |10\rangle + a_2 b_2 |11\rangle .$$

1.3 The Density Operator

The density operator is an alternative approach to quantum mechanics which, while being equivalent to the state vector formulation, allows for a convenient representation of multi-qubit states, ensembles of quantum systems, and interactions between them. This gives a powerful framework to study *open quantum systems*, which are systems not isolated from their surroundings.

We formulate the density matrix by considering a quantum system, which can be in any one of a number of states $|\psi_i\rangle$ with classical probabilities p_i, and call it an *ensemble* of pure states $\{p_i, |\psi_i\rangle\}$.

The density operator for this system is defined in the following manner:

$$\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i| . \tag{1.14}$$

The time evolution of a state represented by the density operator under application of a unitary transformation will be :

$$\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i| \xrightarrow{\text{U}} \rho = \sum_i p_i U |\psi_i\rangle \langle\psi_i| U^\dagger = U\rho U^\dagger .$$

Note that this relies solely on material developed in the state vector approach (i.e., the unitary operator U acting on a ket: $|\psi\rangle \xrightarrow{\text{U}} U|\psi\rangle$). This combination approach blending the new and the old formalism helps to tackle much more advanced problems conveniently.

1.3.1 Properties of The Density Operator

(a) **Trace Condition :** The density operator has a trace equal to one. What this means is that its eigenvalues sum to one. This is a fundamental consequence of the classical probabilities corresponding to the states which comprise the mixed quantum state summing to one.

(b) **Positivity Condition :** The second condition we impose on the density operator is that the operator ρ be a *positive* operator.

We elaborate on the above points below:

The Trace

The trace of a matrix is the sum of the elements on the main diagonal of the matrix. i.e., if we have a matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

the trace will be defined by:

$$Tr(A) = \sum_k a_{kk} = a_{11} + a_{22} + a_{33}$$

But the trace also has a deeper meaning:

$$The\ trace\ of\ a\ matrix\ is\ the\ sum\ of\ its\ eigenvalues.$$

Thus, if $\{\lambda_i\}$ is the set of eigenvalues of the matrix A, then:

$$Tr[A] = \sum_i \lambda_i$$

Properties of the Trace

The trace is a linear function. Hence,

1. $Tr[A + B] = Tr[A] + Tr[B]$.

2. $Tr[kA] = k\,Tr[A]$, where k is a constant.

3. The trace is cyclic:

$$Tr[ABC] = Tr[BCA] = Tr[CAB].$$

   It should be noted, however, that cyclicity of the trace does not hold for infinite dimensional matrices.

Thus, if we have a density matrix $\rho = \sum_i p_i \left|\psi_i\right\rangle \left\langle\psi_i\right|$, then the trace equals unity condition merely implies that all eigenvalues of the state must sum to one.
This makes intuitive sense as well, since the eigenvalues of ρ are nothing but probabilities of collapse to the individual basis states, which must indeed sum to 1.

$$\begin{aligned} Tr[\rho] &= \sum_i p_i Tr[\left|\psi_i\right\rangle \left\langle\psi_i\right|] \\ &= \sum_i p_i \\ &= 1 \end{aligned}$$

Positive Operators

Suppose $|\phi\rangle$ is any arbitrary vector in state space. Then, for the positivity condition to be satisfied, we must have

$$\langle\phi|\rho|\phi\rangle = \sum_i p_i \langle\phi|\psi_i\rangle\langle\psi_i|\phi\rangle \tag{1.15}$$

$$= \sum_i p_i |\langle\phi|\psi\rangle|^2 \tag{1.16}$$

$$\geq 0 \tag{1.17}$$

Thus, the Density Operators are regarded as a *Positive Operator* with a trace of 1.

With these preliminaries done, the stage is now set for us to delve into the density operator formalism.

1.4 The Density Operator Formulation of Quantum Mechanics

Postulate 1

The state of any quantum mechanical system (which resides vector space $\mathcal{H}$) can be described by a *density operator*, which is a positive operator with trace equal to 1.
In addition, if the system is in a set of states $\{\rho_i\}$ with probabilities $\{p_i\}$, then the density operator for the system will be given by;

$$\sum_i p_i \rho_i. \tag{1.18}$$

Postulate 2

Like in the state vector formulation, the evolution of a closed quantum system is described by a *Unitary Transformation*.
i.e., The state of the system at time t_1 is related to the state of the system at time t_2 by a *Unitary Operation* U, which depends only on t_1 and t_2.

$$\rho^{'} = U \rho U^{\dagger} \tag{1.19}$$

Postulate 3

Measurements are described just as they were in state space formulation, with a set of *Measurement Operators* $\{M_m\}$.
Probability of outcome m occurring is;

$$p(m) = Tr(M_m^{\dagger} M_m \rho). \tag{1.20}$$

and the state after measurement is given by:

$$\frac{M_m \rho M_m^{\dagger}}{tr(M_m^{\dagger} M_m \rho)}. \tag{1.21}$$

The measurement operators also satisfy the completeness relation:

$$\sum_m M_m^{\dagger} M_m = \mathbb{1}. \tag{1.22}$$

Postulate 4

The state space of a composite physical system is the *tensor product* of the component physical systems. i.e., If we have systems numbered 1 through n, and system i is prepared in the state ρ_i, then the joint state of the system is given by:

$$\rho = \rho_1 \otimes \rho_2 \otimes ....\rho_n. \tag{1.23}$$

1.4.1 Some Applications of the Density Matrix

Testing for Pure States

One of the many uses of the density matrix, apart from being a very convenient means of representation of states, is to determine whether a given state is *mixed* or *pure*. The means to check this is to see if $Tr(\rho^2) < 1$.
If this quantity is indeed less than 1, then the state is mixed, with equality occurring if and only if the given state is pure.

The Reduced Density Operator

The density matrix is extremely handy as a tool to study the *subsystems* of a composite quantum system.
Bipartite states (systems consisting of two distinct sub-systems in different Hilbert spaces) are of great interest in quantum mechanics, and it may be the case that we need to study not the combined dynamics of the entire entangled system, but that of just one of its component subsystems.
The partial tracing operation essentially *traces out* or ignores one of the subsystems and hands over a *reduced density matrix*, which enables us to look into the component of interest[17]. (We note that this process is also accompanied by a loss of information regarding correlations between states).
For a given bipartite state ρ_{AB}, we define the reduced density operators ρ_A and ρ_B as:

$$\rho_A = Tr_B(\rho_{AB}) = \sum_{i=1}^{n_B} (\mathbb{1}_A \otimes \langle\phi_i|)\rho_{AB}(\mathbb{1}_A \otimes |\phi_i\rangle) \tag{1.24}$$

and likewise, for ρ_B,

$$\rho_B = Tr_A(\rho_{AB}) = \sum_{i=1}^{n_A} (\langle\psi_i| \otimes \mathbb{1}_B)\rho_{AB}(|\psi_i\rangle \otimes \mathbb{1}_B) \tag{1.25}$$

Here, $\{|\phi_i\rangle\}$ and $\{|\psi_i\rangle\}$ form a complete set of orthonormal basis vectors in the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively.
This concludes our rudimentary introduction to the toolkit of quantum mechanics. The next step is to try and understand uncertainty relations, how they arise and what their implications are, as these form the conceptual foundation of this project.

Chapter 2

Uncertainty Relations

Uncertainty relations, to sum it up, are at the heart of both quantum mechanics, and of this project. It is essential that we study existing relations so that we may gain an understanding of how they came about, how they were improved, what they are trying to convey, and finally, what may be done from our side to improve them.

We thus start this chapter by discussing the very genesis of uncertainty relations, which did nothing less than start off a whole new chapter in Physics.

2.1 The Heisenberg-Robertson Relation

The Uncertainty Principle was formulated by Werner Heisenberg in 1927 [6], stating the impossibility of a joint *sharp* measurement of both the position and momentum of a quantum state.
This was expressed in the relation:

$$\Delta p \Delta q \sim h, \tag{2.1}$$

where Δp and Δq denote the standard deviations in the measurement of position and momentum. This arose from the non-commutativity of the position and momentum operators in quantum mechanics. The matrix relation said:

$$\boldsymbol{p}\boldsymbol{q} - \boldsymbol{q}\boldsymbol{p} = -i\hbar\mathbb{1}, \tag{2.2}$$

where $\boldsymbol{p}$ and $\boldsymbol{q}$ denote the position and momentum operators. What Heisenberg put forth was that the more accurately / sharply we try to determine the position of a system, say

that of an electron (using γ rays, for highest resolution), the more we would disturb the system's momentum (by the γ rays transferring their own momentum onto the electron), thereby eliminating the possibility of a sharp momentum measurement.
A formal mathematical statement to the concept was given by E.H. Kennard (1927) [7] as:

$$\Delta\hat{Q}\Delta\hat{P} \geq \frac{\hbar}{2} \tag{2.3}$$

A formulation for general non-commuting operators $\hat{X}$ and $\hat{Y}$ though, was provided by H.P. Robertson [13] as:

$$\Delta\hat{X}\Delta\hat{Y} \geq \frac{1}{2}|\langle[\hat{X},\hat{Y}]\rangle| \tag{2.4}$$

2.1.1 Arriving at the Heisenberg-Robertson Uncertainty Relation

We start by defining the very quantity on which this relation rests: the variance.
The variance related to an observable $\hat{A}$ acting on a state $|\psi\rangle$ is defined by:

$$\begin{aligned}\Delta A^2 &= \langle(\hat{A} - \langle\hat{A}\rangle)^2\rangle \\ &= \langle A^2 - 2\hat{A}\langle\hat{A}\rangle + \langle\hat{A}\rangle^2\rangle \\ &= \langle\hat{A}^2\rangle - 2\langle\hat{A}\rangle^2 + \langle\hat{A}\rangle^2 \\ &= \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2\end{aligned}$$

Where $\langle A\rangle$ is the expectation value of the measurement, and is given by:

$$\langle\hat{A}\rangle = \frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle} \tag{2.5}$$

We now define the following states:

$$|f\rangle = \Delta\hat{A}\,|\psi\rangle\,, \tag{2.6}$$

and

$$|g\rangle = \Delta\hat{B}\,|\psi\rangle\,, \tag{2.7}$$

where

$$\Delta\hat{A}\psi = (\hat{A} - \langle A\rangle\mathbb{1})\,|\psi\rangle\,,$$

and

$$\Delta\hat{B}\psi = (\hat{B} - \langle B\rangle\mathbb{1})\,|\psi\rangle\,.$$

We now introduce the Cauchy-Schwarz inequality (proven in Appendix A):

$$\langle f|f\rangle\langle g|g\rangle \;\geq\; |\langle f|g\rangle|^2, \tag{2.8}$$

and apply it using the states defined in eqs (2.6) and (2.7), which gives us:

$$\langle\psi|\,\Delta\hat{A}^\dagger\Delta\hat{A}\,|\psi\rangle\,\langle\psi|\,\Delta\hat{B}^\dagger\Delta\hat{B}\,|\psi\rangle \;\geq\; |\,\langle\psi|\,\Delta\hat{A}^\dagger\Delta\hat{B}\,|\psi\rangle\,|^2. \tag{2.9}$$

Since $\hat{A}$ and $\hat{B}$ are both Hermitian, we have:

$$\begin{aligned}\Delta\hat{A}^\dagger &= \hat{A}^\dagger - \langle\hat{A}\rangle^\dagger\\ &= \hat{A} - \langle\hat{A}\rangle\\ &= \Delta\hat{A}\end{aligned}$$

The same holds for $\Delta\hat{B}^\dagger$. So, eq (2.9) becomes;

$$\langle\psi|\,\Delta\hat{A}^2\,|\psi\rangle\,\langle\psi|\,\Delta\hat{B}^2\,|\psi\rangle \;\geq\; |\,\langle\psi|\,\Delta\hat{A}\Delta\hat{B}\,|\psi\rangle\,|^2 \tag{2.10}$$

We note that:

$$\Delta\hat{A}\Delta\hat{B} = \frac{1}{2}[\Delta\hat{A},\ \Delta\hat{B}] + \frac{1}{2}\{\Delta\hat{A},\ \Delta\hat{B}\}.$$

But,

$$[\Delta\hat{A}, \Delta\hat{B}] = [\hat{A}, \hat{B}].$$

Hence,

$$\langle\psi|\,\Delta\hat{A}^2\,|\psi\rangle\,\langle\psi|\,\Delta\hat{B}^2\,|\psi\rangle \;\geq\; \frac{1}{2}\langle[\hat{A}, \hat{B}]\rangle + \frac{1}{2}\langle\{\Delta\hat{A}, \Delta\hat{B}\}\rangle,$$

Now, $[\hat{A}, \hat{B}]$ is an Anti-Hermitian operator. Its expectation value will therefore be imaginary. Expectation of $\{\hat{A}, \hat{B}\}$, on the other hand, will be real, as it is a Hermitian operator.
This means the above equation is of the form $z = a+ib$.In such a case, to find the magnitude of the complex number z, we simply do;

$$|z|^2 = |a|^2 + |b|^2$$

So,

$$\begin{aligned}|\langle\Delta\hat{A}\Delta\hat{B}\rangle|^2 \;&=\; \frac{1}{4}|\langle[\Delta\hat{A}, \Delta\hat{B}]\rangle|^2 + \frac{1}{4}|\langle\{\Delta\hat{A},\ \Delta\hat{B}\}\rangle|^2\\ |\langle\Delta\hat{A}\Delta\hat{B}\rangle|^2 \;&\geq\; \frac{1}{4}|\langle[\Delta\hat{A}, \Delta\hat{B}]\rangle|^2\end{aligned}$$

Now, on going back to (2.10),

$$\begin{aligned}\langle\psi|\,\Delta\hat{A}^2\,|\psi\rangle\,\langle\psi|\,\Delta\hat{B}^2\,|\psi\rangle &\geq |\langle\Delta\hat{A}\Delta\hat{B}\rangle|^2\\ \langle\Delta\hat{A}^2\rangle\langle\Delta\hat{B}^2\rangle &\geq |\langle\Delta\hat{A}\Delta\hat{B}\rangle|^2\\ &\geq \frac{1}{4}|\langle[\hat{A},\hat{B}]\rangle|^2\end{aligned}$$

Thus, we have:

$$\Delta\hat{A}^2\Delta\hat{B}^2 \geq \frac{1}{4}|\langle[\hat{A},\hat{B}]\rangle|^2 \tag{2.11}$$

on taking the square root of which, we get the Robertson-Heisenberg uncertainty relation:

$$\Delta\hat{A}\Delta\hat{B} \geq \frac{1}{2}|\langle[\hat{A},\hat{B}]\rangle|. \tag{2.12}$$

2.2 The Schrödinger Uncertainty Relation

The Schrödinger uncertainty relation [15] tightens the bound provided by the Robertson relation by adding an anti-commutator term on the RHS, and is a good illustration of tightening of an existing bound by a stronger one.

We start off from the following part of the proof for the Robertson relation:

$$|f\rangle = \Delta\hat{A}\psi, \tag{2.13}$$

and

$$|g\rangle = \Delta\hat{B}\psi, \tag{2.14}$$

where

$$\Delta\hat{A}\psi = (\hat{A} - \langle A\rangle\mathbb{1})\,|\psi\rangle\,,$$

and

$$\Delta\hat{B}\psi = (\hat{B} - \langle B\rangle\mathbb{1})\,|\psi\rangle\,,$$

We recall the Cauchy-Schwarz inequality,

$$\langle f|f\rangle\langle g|g\rangle \;\geq\; |\langle f|g\rangle|^2. \tag{2.15}$$

Now, $\langle f|g\rangle$ is a complex number. A complex number can be written in the form $z = a + ib$, which in turn can be expressed in the polar form as: $z = cos\theta + i\ sin\theta$, The magnitude of z will be given by: $|z|^2 = |a|^2 + |b|^2 = cos^2\theta + sin^2\theta$.

But,

$$cos\theta = \left(\frac{z+z^*}{2}\right) \qquad\qquad sin\theta = \left(\frac{z-z^*}{2i}\right)$$

Thus,

$$\begin{aligned}|z|^2 &= [Re(z)]^2 + [Im(z)]^2 \\ &= \left(\frac{z+z^*}{2}\right)^2 + \left(\frac{z-z^*}{2i}\right)^2\end{aligned}$$

So,

$$|\langle f|g\rangle|^2 = \left(\frac{\langle f|g\rangle + \langle f|g\rangle^*}{2}\right)^2 + \left(\frac{\langle f|g\rangle - \langle f|g\rangle^*}{2i}\right)^2 \tag{2.16}$$

Now we work out the terms needed for the above expression. We start with finding $\langle f|g\rangle$:

$$\begin{aligned}\langle f|g\rangle &= \langle\psi|\,\Delta\hat{A}^\dagger\Delta\hat{B}\,|\psi\rangle \\ &= \langle\psi|\,\Delta\hat{A}\Delta\hat{B}\,|\psi\rangle \\ &= \langle(\hat{A}-\langle\hat{A}\rangle)(\hat{B}-\langle\hat{B}\rangle)\rangle \\ &= \langle\hat{A}\hat{B} - \langle\hat{B}\rangle\hat{A} - \langle\hat{A}\rangle\hat{B} + \langle\hat{A}\rangle\langle\hat{B}\rangle\rangle \\ &= \langle\hat{A}\hat{B}\rangle - \langle\hat{B}\rangle\langle\hat{A}\rangle,\end{aligned}$$

and upon taking the complex conjugate of this,we obtain:

$$\begin{aligned}\langle f|g\rangle^* &= \langle\psi|\,\Delta\hat{B}^\dagger\Delta\hat{A}\,|\psi\rangle \\ &= \langle(\hat{B}-\langle\hat{B}\rangle)(\hat{A}-\langle\hat{A}\rangle)\rangle \\ &= \langle\hat{B}\hat{A} - \langle\hat{A}\rangle\hat{B} - \langle\hat{B}\rangle\hat{A} + \langle\hat{B}\rangle\langle\hat{A}\rangle\rangle \\ &= \langle\hat{B}\hat{A}\rangle - \langle\hat{B}\rangle\langle\hat{A}\rangle\end{aligned}$$

Adding the above two expressions together, we may write:

$$\begin{aligned}\frac{\langle f|g\rangle + \langle f|g\rangle^*}{2} &= \frac{\langle\hat{A}\hat{B}\rangle + \langle\hat{B}\hat{A}\rangle - 2\langle\hat{A}\rangle\langle\hat{B}\rangle}{2} \\ &= \frac{1}{2}\langle\{\hat{A},\hat{B}\}\rangle - \langle\hat{A}\rangle\langle\hat{B}\rangle\end{aligned}$$

and

$$\begin{aligned}\frac{\langle f|g\rangle - \langle f|g\rangle^*}{2i} &= \frac{\langle\hat{A}\hat{B}\rangle - \langle\hat{B}\hat{A}\rangle}{2i} \\ &= \frac{1}{2i}\langle[\hat{A},\hat{B}]\rangle\end{aligned}$$

Putting the above 2 results in eq(2.16), we get:

$$\Delta A^2 \Delta B^2 \geq \left|\frac{1}{2i}\langle[\hat{A},\hat{B}]\rangle\right|^2 + \left|\frac{1}{2}\langle\{\hat{A},\hat{B}\}\rangle - \langle\hat{A}\rangle\langle\hat{B}\rangle\right|^2 \tag{2.17}$$

This is the Robertson-Schrödinger uncertainty relation. The additional anticommutator term on the RHS is what makes the already existing bound tighter.

2.3 The Limitation with 'Product' Uncertainty Relations

The Robertson-Heisenberg relation gives a generalized treatment to uncertainties involved in quantum measurements, and the Robertson-Schrödinger relation tightens that bound. What both these relations do is to express a limitation in the possible preparations of a quantum system by giving a lower bound to the product of variances of two observables in terms of their commutators.

What they do not do is capture fully the notion of incompatible observables because, in certain situations, the LHS often tends to be trivial, as is illustrated below with the Heisenberg - Robertson relation:

$$\Delta\hat{A}^2\Delta\hat{B}^2 \geq \frac{1}{4}|\langle[\hat{A},\hat{B}]\rangle|^2 \tag{2.18}$$

We understand that variance is given by:

$$\Delta A^2 = \langle A^2\rangle - \langle A\rangle^2.$$

Now, if we chose to make a measurement in an eigenstate of *one of* the observables, then,

$$\begin{aligned}
\Delta A^2 &= \langle A^2\rangle - \langle A\rangle^2 \\
&= \langle\psi|A^2|\psi\rangle - \langle\psi|A|\psi\rangle\langle\psi|A|\psi\rangle \\
&= a_i^2\langle\psi|\psi\rangle - a_i^2\langle\psi|\psi\rangle\langle\psi|\psi\rangle \\
&= a_i^2 - a_i^2 \\
&= 0.
\end{aligned}$$

In such a case, we see that the LHS goes to zero. The relation has thus become *trivial*. As a remedy for this, we have uncertainty relations expressed in the form of a sum of variances. Thus, even in cases where measurement is carried out in an eigenstate of one of the observables, the relation does not become trivial. One such relation is the Maccone - Pati uncertainty relation.

2.4 The Maccone - Pati Uncertainty Relation

Lorenzo Maccone and Arun K Pati in their 2015 paper detail an uncertainty relation defined in terms of the sum of variances[9]. Studying this relation was a crucial step, as this work is basically an attempt to formulate such a 'sum' based uncertainty in its course.

The relation reads as follows:

$$\Delta A^2 + \Delta B^2 \geq \text{max}\,(\mathcal{L}_1, \mathcal{L}_2) \tag{2.19}$$

with

$$\mathcal{L}_1 = \frac{1}{2} |\langle \psi^{\perp}_{A+B}| A + B |\psi\rangle|^2$$
$$\mathcal{L}_2 = \pm i[\hat{A}, \hat{B}] + |\langle \psi| \hat{A} \pm i\hat{B} |\psi^{\perp}\rangle|^2$$

Maccone-Pati give us potential lower bounds, with the choice of the lower bound being up to the user, and which may be selected to give the tightest bound on uncertainty. Another deciding factor for tightness is the choice of the orthogonal state $|\psi\rangle$, which we shall discuss in more detail. The proofs which we shall use to derive the above given bounds are based on the parallelogram law of vector addition, which has been derived in Appendix C[10].

Proof. **Deriving the bound $\mathcal{L}_1 = \frac{1}{2} |\langle \psi^{\perp}_{A+B}| A + B |\psi\rangle|^2$**

We define two states:

$$|\eta\rangle \equiv (\hat{A} - \langle \hat{A}\rangle \mathbb{I}) |\psi\rangle \qquad\qquad |\xi\rangle \equiv (\hat{B} - \langle \hat{B}\rangle \mathbb{I}) |\psi\rangle$$

Now, we apply the parallelogram law for vector spaces to the states $|\eta\rangle$ and $|\xi\rangle$, which reads :

$$2(\||\eta\rangle\|^2 + \||\xi\rangle\|^2) = \||\eta\rangle + |\xi\rangle\|^2 + \||\eta\rangle - |\xi\rangle\|^2 \tag{2.20}$$

Thus giving us:

$$\||\eta\rangle\|^2 + \||\xi\rangle\|^2 = \frac{1}{2}(\||\eta\rangle + |\xi\rangle\|^2 + \||\eta\rangle - |\xi\rangle\|^2)$$

The norm is positive. Hence, both quantities on the left side are positive. Which enables us to write:

$$\begin{aligned}\||\eta\rangle\|^2 + \||\xi\rangle\|^2 &\geq \frac{1}{2}(\||\eta\rangle \pm |\xi\rangle\|^2 \\ &= \frac{1}{2}(\langle\eta| \pm \langle\xi|)(|\eta\rangle \pm |\xi\rangle)\langle\psi^{\perp}|\psi^{\perp}\rangle\end{aligned}$$

where the state $|\psi^{\perp}\rangle$, is orthonormal to the state $|\psi\rangle$

Now, we make use of the *Cauchy - Schwarz* inequality on the RHS, and obtain:

$$\begin{aligned}\||\eta\rangle\|^2 + \||\xi\rangle\|^2 &\geq \frac{1}{2}(\langle\eta| \pm \langle\xi|)\,|\psi^{\perp}\rangle \\ &= \frac{1}{2}|\,\langle\psi|\,\hat{A} \pm \hat{B}\,|\psi^{\perp}\rangle\,|^2\end{aligned}$$

Now, since $|\eta\rangle \equiv (\hat{A} - \langle\hat{A}\rangle\mathbb{I})\,|\psi\rangle$, and $|\xi\rangle \equiv (\hat{B} - \langle\hat{B}\rangle\mathbb{I})\,|\psi\rangle$

$$\||\eta\rangle\|^2 = \Delta\hat{A}^2$$

and

$$\||\xi\rangle\|^2 = \Delta\hat{B}^2$$

Thus, we obtain the first of the two bounds :

$$\Delta\hat{A}^2 + \Delta\hat{B}^2 \geq \frac{1}{2}|\,\langle\psi|\,\hat{A} \pm \hat{B}\,|\psi^{\perp}\rangle\,|^2$$

Now we move to proving the second bound.

Deriving the bound $\mathcal{L}_2 = \pm i[\hat{A}, \hat{B}] + |\,\langle\psi|\,\hat{A} \pm i\hat{B}\,|\psi^{\perp}\rangle\,|^2$

As in the previous proof, we first define the following two states:

$$|\eta\rangle \equiv (\hat{A} - \langle\hat{A}\rangle\mathbb{I})\,|\psi\rangle \qquad\qquad |\xi\rangle \equiv (\hat{B} - \langle\hat{B}\rangle\mathbb{I})\,|\psi\rangle$$

and we note that:

$$\||\xi\rangle\| = \|i\,|\xi\rangle\| \tag{2.21}$$

So now, we may have:

$$\begin{aligned}\||\eta\rangle \pm i\,|\xi\rangle\|^2 &= (\langle\eta| \mp i\,\langle\xi|)(|\eta\rangle \pm i\,|\xi\rangle) \\ &= \||\eta\rangle\|^2 + \||\xi\rangle\|^2 \mp i\langle\xi|\eta\rangle \pm i\langle\eta|\xi\rangle \\ &= \||\eta\rangle\|^2 + \||\eta\rangle\|^2 \pm i(\langle\eta|\xi\rangle - \langle\xi|\eta\rangle)\end{aligned}$$

But on carrying out the expansion of the last term on the RHS, we see that:

$$\langle\eta|\xi\rangle - \langle\xi|\eta\rangle = [\hat{A}, \hat{B}]$$

Thus,

$$\||\eta\rangle \pm i\,|\xi\rangle\|^2 = \||\eta\rangle\|^2 + \||\eta\rangle\|^2 \pm i[\hat{A}, \hat{B}] \tag{2.22}$$

and,

$$\||\eta\rangle \pm i\,|\xi\rangle\|^2 = (\langle\eta| \mp i\,\langle\xi|)(|\eta\rangle \pm i\,|\xi\rangle) \tag{2.23}$$

which, using the state $|\psi^{\perp}\rangle$, can be written as:

$$\begin{aligned}
\||\eta\rangle \pm i\,|\xi\rangle\|^2 &= (\langle\eta| \mp i\,\langle\xi|)(|\eta\rangle \pm i\,|\xi\rangle)\langle\psi^{\perp}|\psi^{\perp}\rangle \\
&\geq |(\langle\eta| \mp i\,\langle\xi|)\,|\psi^{\perp}\rangle\,|^2 \\
&= |\,\langle\psi|\,[\hat{A} - \langle\hat{A}\rangle \mp i(\hat{B} - \langle\hat{B}\rangle)]\,|\psi^{\perp}\rangle\,|^2 \\
&= |\,\langle\psi|\,\hat{A} \mp i\hat{B}\,|\psi^{\perp}\rangle - (\langle\hat{A}\rangle \mp i\langle\hat{B}\rangle)\langle\psi|\psi^{\perp}\rangle|^2 \\
&= |\,\langle\psi|\,\hat{A} \mp i\hat{B}\,|\psi^{\perp}\rangle\,|^2
\end{aligned}$$

Now, making use of the parallelogram law for vector spaces, and incorporating results from the above steps and the fact that $\||\xi\rangle\| = \|i\,|\xi\rangle\|$, we get:

$$2(\Delta\hat{A}^2 + \Delta\hat{B}^2) = \Delta\hat{A}^2 + \Delta\hat{B}^2 \pm i[\hat{A}, \hat{B}] + |\,\langle\psi|\,\hat{A} + i\hat{B}\,|\psi^{\perp}\rangle\,|^2$$

Leaving us with the second bound:

$$\Delta\hat{A}^2 + \Delta\hat{B}^2 \;\geq\; \pm i[\hat{A}, \hat{B}] + |\,\langle\psi|\,\hat{A} \pm i\hat{B}\,|\psi^{\perp}\rangle\,|^2 \tag{2.24}$$

■

A problem of interest lies in choosing an appropriate value for the state $|\psi^{\perp}\rangle$ orthogonal to $|\psi\rangle$, because as is evident, the tightness of the bound is invariably linked to this choice. A particularly straightforward illustration on choosing a suitable state, based on L. Vaidman's 1992 paper [18] is as follows:

$$|\psi^{\perp}\rangle = \frac{(\hat{A} - \langle\hat{A}\rangle)\,|\psi\rangle}{\Delta\hat{A}}, \tag{2.25}$$

which Vaidman arrived at as follows:
Any state $|\psi\rangle$ can be expressed as a linear combination of itself and an orthogonal state. So we may write:

$$\hat{A}\,|\psi\rangle = a\,|\psi\rangle + b\,|\psi^{\perp}\rangle\,, \tag{2.26}$$

where b is a real number. It is evident from the above expression that $a = \langle\psi|\,\hat{A}\,|\psi\rangle = \langle\hat{A}\rangle$. Now, if $\hat{A}$ is Hermitian, we shall have:

$$\langle\psi|\,\hat{A}^\dagger\hat{A}\,|\psi\rangle = (a^*\,\langle\psi| + b\,\langle\psi|^\dagger)(a\,|\psi\rangle + b\,|\psi^\perp\rangle)$$
$$\langle\psi|\,\hat{A}^2\,|\psi\rangle = |a|^2 + b^2.$$

$|a|^2$ is merely the square of the expectation value of $\hat{A}$, so that:

$$b^2 = \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2$$

This means $b = \Delta\hat{A}$. Putting this back in the linear combination, we obtain an expression for an orthogonal state $|\psi\rangle^\perp$

$$|\psi^\perp\rangle = \frac{(\hat{A} - \langle\hat{A}\rangle)\,|\psi\rangle}{\Delta\hat{A}}$$

This is one of the means by which we can choose an orthogonal state $|\psi^\perp\rangle$.

Having explored the development of a sum based uncertainty relation, we move on and explore the information theory behind uncertainty, taking special care to understand conditional entropies, as we shall eventually be dealing with conditioning of uncertainties in the form of variances for our own work.

Chapter 3

Entropy

The objective of this project was to find an expression (if one exists) providing a lower bound for uncertainty in the measurement of two complementary observables in the presence of a *memory*, while using variance as a quantifier for uncertainty.

Existing literature does prove that such a relation exists, and in fact uses **entropy** to quantify the uncertainties [2]. If we are to see further, we must stand on the shoulders of giants, and if we are to attempt to formulate a relation of our own, an understanding of the existing relations is of paramount importance. A section dedicated to entropy is therefore inevitable, as knowledge gained translated to getting closer to a relation of our own.

3.1 Shannon Entropy

The concept of entropy was developed by C.E. Shannon in his seminal 1948 paper [16], out of the need to quantify exactly how much information was associated with each possible outcome of a random variable in a communication channel, since the initial theory, despite the universal significance it would later have, was developed keeping communication systems in mind.

To each outcome of a random variable was allotted a quantity called the *surprisal*, denoted by $I(x)$, which was essentially a measure of how *unexpected* that particular outcome was and therefore, how much information it carried.

It was assigned the following mathematical form:

$$I(x) = \log_2 \frac{1}{p(x)} \tag{3.1}$$

$$= -\log_2 p(x). \tag{3.2}$$

Events which have a lower possibility of occurrence have higher values of surprisal (there is always more interest in knowing how probable a rare event will be as compared to an everyday vanilla occurrence), whereas more common occurrences had lower surprisals (as they revealed much less information).

Entropy was then formulated as the *expectation* of these surprisals over all values of the random variable of interest.

$$H(X) = E(I(X)) \tag{3.3}$$

Which is represented more explicitly as:

$$H(X) = H(p_1, p_2, ..., p_n) \tag{3.4}$$

$$= \sum_x p_x I(x) \tag{3.5}$$

$$= -\sum_x p_x \log p_x \tag{3.6}$$

Here, the different p_x are the individual probabilities for each outcome of the random variable X.

We note that the Shannon Entropy is *always positive* in the classical domain.

3.1.1 Two views of Entropy

There are two distinct ways of looking at entropy, which also happen to be complementary points of view.

(***a***) As a measure of uncertainty about the random variable X, *before* we actually learn the value of X.

(***b***) As a measure of how much information we have gained / how much uncertainty has been reduced *after* we learn the value of X.

Entropy can thus act as a measure of either the uncertainty in the value of a random variable, or of the information contained in it.

3.1.2 Relative Entropy

The relative entropy measures *how close* two probability distributions $p(x)$ and $q(x)$, over the same index set, are to each other.
It is defined as:

$$H(p(x)||q(x)) = \sum_x p(x) \log \frac{p(x)}{q(x)}. \tag{3.7}$$

The relative entropy is a **non - negative** quantity i.e., $H(p(x)||q(x)) \geq 0$ always, the proof of which is detailed below.

$$\begin{aligned} H(p(x)||q(x)) &= \sum_x p(x) \log \frac{p(x)}{q(x)} \\ &= -\sum_x p(x) \log \frac{q(x)}{p(x)} \end{aligned}$$

Now, to proceed further, we use the identity $\log x \ln 2 = \ln x \leq x - 1$. Hence giving us: $-\log x \geq \frac{1}{\ln 2}(1 - x)$ Using this inequality, we now have:

$$\begin{aligned} H(p(x)||q(x)) &\geq \frac{1}{\ln 2} \sum_x p(x) \left(1 - \frac{q(x)}{p(x)}\right) \\ &= \frac{1}{\ln 2} \sum_x (p(x) - q(x)) \\ &= \frac{1}{\ln 2} \left(\sum_x p(x) - \sum_x q(x) \right) \\ &= 0. \end{aligned}$$

Thus proving that $H(p(x)||q(x)) \geq 0$ always.

3.1.3 Conditional Entropy

The entropy of X conditioned on Y is the uncertainty associated with the random variable X *after* we know the value of Y. It is given the expression:

$$H(X|Y) = H(X,Y) - H(X) \tag{3.8}$$

It can also be expressed as:

$$H(X|Y) \;=\; -\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p(x,y) \log \frac{p(x,y)}{p(y)} \tag{3.9}$$

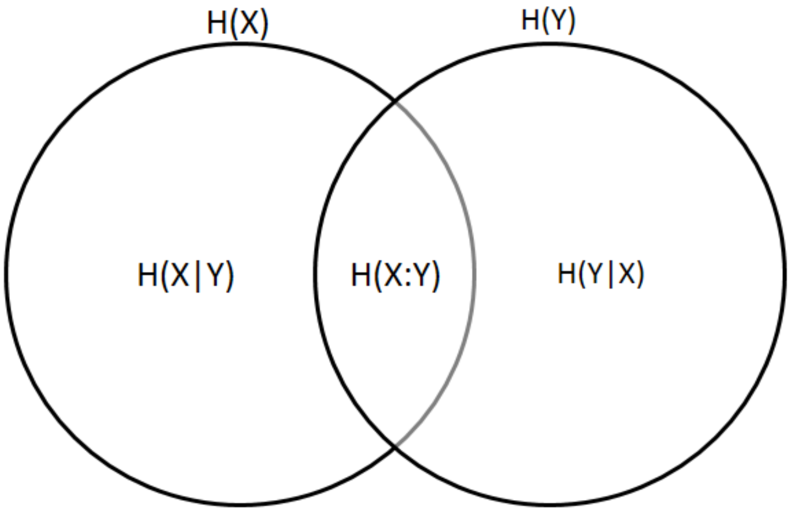


Figure 3.1: The various Shannon Entropies

3.1.4 The motivation for Conditional Entropy

Let $H(X|Y = y)$ be the entropy of the discrete random variable X when conditioned on the random variable Y taking on a single value y.

We know that the unconditioned entropy of the random variable X is given by:

$$\begin{aligned} H(X) &= E[I(X)] \\ &= -\sum_{x \in \mathcal{X}} p(X = x) \log p(X = x) \end{aligned}$$

Now, upon conditioning over a single value of the random variable Y, we get:

$$H(X|Y = y) = -\sum_{x} p(X = x|Y = y) \log_2 p(X = x|Y = y) \tag{3.10}$$

The key concept here is that $H(X|Y)$ is obtained by merely averaging the above given expression for $H(X|Y = y)$ over all values of $y \in Y$.

$$\begin{aligned} H(X|Y) &= \sum_{y \in \mathcal{Y}} p(y) H(X|Y) \\ &= -\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p(y) p(x|y) \log_2 p(x|y) \\ &= -\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p(x, y) \log_2 p(x|y) \\ &= -\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p(x, y) \log_2 \frac{p(x, y)}{p(y)} \end{aligned}$$

Thus, we have an expression for the classical version of a quantity we will be building this project on; that of the conditional entropy, given as:

$$H(X|Y) = -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log_2\frac{p(x,y)}{p(y)} \tag{3.11}$$

3.1.5 Mutual Information

This quantity measures how much information is *common* to both X and Y. It is obtained by subtracting the joint entropy $H(X,Y)$ from the information content of X and Y taken separately.

$$H(X:Y) = H(X) + H(Y) - H(X,Y) \tag{3.12}$$

3.1.6 Properties of Shannon Entropy

(i) $H(X,Y) = H(Y,X)$ *and* $H(X:Y) = H(Y:X)$

(ii) $H(X|Y) \geq 0$

i.e., The conditional entropy is positive.

We note that this holds true only for classical random variables, and shall see that in the quantum domain, we can have negative conditional entropies.

Proof. The joint entropy of a system of two random variables X and Y is expressed as:

$$\begin{aligned} H(X,Y) &= -\sum_{x,y} p(x,y)\log p(x)p(x|y) \\ &= -\sum_{x} p(x)\log p(x) - \sum_{x,y} p(x,y)\log p(x|y) \end{aligned}$$

The first term on the RHS amounts to $H(X)$. The second term is the expression for $H(X|Y)$ from eq (3.11) and it basically says:

$$H(X|Y) = -\sum_{x,y} p(x,y)\log p(x|y)$$

Since all probabilities are less than or equal to 1, we will have $-\log p(x|y) \geq 0$, thus giving us $-\sum_{x,y} p(x,y)\log p(x|y) \geq 0$, which in turn means:

$$H(X|Y) \geq 0$$

■

(iii) $H(X,Y) \geq H(X)$

Proof. From the previous proof, we obtained:

$$H(X,Y) = H(X) + H(X|Y)$$

we also proved $H(X|Y) \geq 0$ Hence $H(X,Y) \geq H(X)$ ■

(iv) The maximum entropy of a random variable with d outcomes is when all outcomes have an equally likely chance of occurring. This makes intuitive sense as well, since we have maximum confusion/uncertainty about *which* event will occur, when any event can occur with the same probability.

Proof. Relative entropy always being positive, we have:

$$H(p(x) \,||\, q(x)) \geq 0$$

Since all d events $q(x)$ have equal probability $\frac{1}{d}$,

$$H(p(x) \,||\, 1/d) \geq 0 \tag{3.13}$$

In this case, the expression for relative entropy is $\sum_x p(x) \log \frac{p(x,y)}{1/d}$. Thus giving us:

$$\begin{aligned} \sum_x p(x) \log \frac{p(x,y)}{1/d} \;&\geq\; 0 \\ \sum_x p(x) \log p(x) - \sum_x p(x) \log 1/d \;&\geq\; 0 \end{aligned}$$

The $p(x)$ on the second term sums to 1, leaving us with:

$$-H(X) + \log d \;\geq\; 0 \tag{3.14}$$

This gives us the bound on the maximum possible entropy of a d dimensional random variable as:

$$H(X) \leq \log d \tag{3.15}$$

■

(v) $H(X) + H(Y) \geq H(X,Y)$

Proof. Taking all terms to one side, we have:

$$H(X,Y) - H(X) - H(Y) = -\sum_{x,y} p(x,y) \log p(x,y) + \sum_{x} p(x) \log p(x) + \sum_{y} p(y) \log p(y)$$

This condenses to give:

$$\sum_{x,y} p(x,y) \log \frac{p(x)p(y)}{p(x,y)} \tag{3.16}$$

We have previously shown that $\log x \leq \frac{1}{\ln 2}(x-1)$.

Now, using this, we have:

$$\begin{aligned} H(X,Y) - H(X) - H(Y) &= \sum_{x,y} p(x,y) \log \frac{p(x)p(y)}{p(x,y)} \\ &\leq \frac{1}{\ln 2} \sum_{x,y} p(x,y) \left(\frac{p(x)p(y)}{p(x,y)} - 1 \right) \\ &= \frac{1}{\ln 2} \sum_{x,y} (p(x)p(y) - p(x,y)) \end{aligned}$$

The probabilities sum to 1, leaving us with:

$$H(X,Y) - H(X) - H(Y) \leq 0$$

and therefore:

$$H(X) + H(Y) \geq H(X,Y) \tag{3.17}$$

This gives us the subadditivity principle for Shannon entropies. ■

3.1.7 The Fano Inequality

Conditioning of random variables is not a magic bullet to cure all uncertainty. Suppose we use knowledge gained from inferring the value of a random variable Y to infer the value of another random variable X. It is important to mention that there is a limit to *how much* information we can gather by such a method. This bound is given by the Fano inequality.

$$H(p_e) + p_e \log(|X| - 1) \geq H(X|Y) \tag{3.18}$$

Where we define $\widetilde{X} \equiv f(Y)$ as the subset of values of X that arise as a result of knowing Y, and $p_e = p(X \neq \widetilde{X})$ as the probability of getting an *error* i.e., guessing a value of X that is *not a function of* Y.

Proof of this theorem is given in Appendix A.

3.2 The von Neumann Entropy

The von Neumann entropy is the quantum equivalent of the Shannon entropy. Here, instead of using classical probability distributions, we use the density matrices of the states involved.

The von Neumann entropy is calculated as:

$$S(\rho) \;=\; -Tr[\rho \log_2 \rho] \tag{3.19}$$

Thus, if k are the eigenvalues of the state ρ, then

$$S(\rho) \;=\; -\sum_k k \log_2 k \tag{3.20}$$

3.2.1 Relative von Neumann Entropy

Just as we defined a relative Shannon entropy, so we define the relative von Neumann entropy:

$$S(\rho||\sigma) = Tr[\rho \log_2 \rho] - Tr[\rho \log_2 \sigma] \tag{3.21}$$

For the value of $S(\rho||\sigma)$ to be finite, the support of ρ must have a null intersection with the kernel of σ.

The quantum relative entropy is a very useful quantity, which like its classical analogue, also always assumes a positive value.

$$S(\rho||\sigma) \geq 0 \tag{3.22}$$

3.2.2 Properties of Interest of the von Neumann Entropy

(i) The von Neumann entropy is non-negative, apart from the case where we calculate the *conditional* von Neumann entropy of an entangled state, in which case it can be negative.

(ii) The von Neumann entropy of a pure state is *zero*.
This can be inferred from the fact that a pure state has only one eigenvalue, which is

1. Hence:

$$
\begin{aligned}
S(\rho) &= -\sum_k k \log_2 k \\
&= -(1 \log_2 1) \\
&= 0
\end{aligned}
$$

(iii) The maximum entropy possible for a system in n-dimensional Hilbert space is given by $\log n$.

The state of a system which entails the maximum uncertainty is a *maximally mixed* state: where its density matrix is of the form $\frac{\mathbb{1}}{n}$. In such a case, the relative entropy will be expressed as : $S(\rho||\sigma) = S(\rho \,||\, \mathbb{1}/n)$
Since the quantum relative entropy is positive, we will have:

$$
\begin{aligned}
S(\rho||\mathbb{1}/n) &\geq 0 \\
Tr[\rho \log_2 \rho] - Tr[\rho \log_2 \mathbb{1}/n] &\geq 0 \\
-S(\rho) + \log n \sum_k k &\geq 0
\end{aligned}
$$

We thus end up with $S(\rho) \leq \log n$

(iv) The entropy of a mixed state is given by the expression:

$$
S\left(\sum_i p_i \rho_i\right) = H(p_i) + p_i S(\rho_i) \tag{3.23}
$$

The extra Shannon entropy term can be understood as arising from the presence of classical states which make up a mixed state.

Proof. We have a mixed state comprising of classical states ρ_i, whose eigenvalues we take to be the set $\{\lambda_i^j\}$, with the index j running from 1 to n (depending on the dimension of the state). The set of eigenvectors we assign as $\{|\epsilon_i^j\rangle\}$. Thus, if ρ_i has eigenvalues $\{\lambda_i^j\}$, $p_i\rho_i$ will have eigenvalues $\{p_i\lambda_i^j\}$, leading to the von Neumann entropy being given by:

$$
\begin{aligned}
S\left(\sum_i p_i\rho_i\right) &= -\sum_{i,j} p_i\lambda_i^j \log_2\left(p_i\lambda_i^j\right) \\
&= -\sum_{i,j} p_i\lambda_i^j \left(\log_2 p_i + \log_2 \lambda_i^j\right) \\
&= -\sum_{i,j} p_i\lambda_i^j \log_2 p_i - \sum_{i,j} p_i\lambda_i^j \log_2 \lambda_i^j
\end{aligned}
$$

In case of the first term on the RHS, the eigenvalues will sum to 1 as we run over the index j for all i (since trace of all the states comprising the mixed states *must* be 1). We can thus remove it from the term without consequence.

$$
\begin{aligned}
S\left(\sum_i p_i\rho_i\right) &= -\sum_{i,j} p_i \log_2 p_i - \sum_i \left(p_i\left(\sum_j \lambda_i^j \log_2 \lambda_i^j\right)\right) \\
&= H(p_i) + \sum_i p_i S(\rho_i)
\end{aligned}
$$

■

3.3 Entropic Uncertainty Relations

Uncertainty relations for a long time were expressed almost exclusively in terms of standard deviation/variance. But variance in itself is often times a rather flawed measure of uncertainty, owing to its insensitivity to a number of nuances which, in many cases, had rather serious implications.

The following were the shortcomings with variance which prompted the development of alternative formulations of uncertainty relations:

(a) Variance tends to **'misbehave'**, giving strikingly counter-intuitive results.A case in point being measurement of s_z for a spin-1 particle, giving a higher standard deviation when one component of spin is *known*, compared to when *none* of the three components are known.

(b) A relabelling of obsesrvables can at times cause the variance/standard deviation to change when no such change is warranted. This is unacceptable, since uncertainty/lack of information regarding a quantity must be independent of what label is used to define the quantity [4]

Owing to these shortcomings with the standard deviation, the Entropy is considered a much more fundamental and objective means of measuring the uncertainty associated with measurement of an observable.
Now that we have got down the formalism for understanding both uncertainty relations as well as entropy, the natural extension is to investigate Entropic Uncertainty relations, which will lead us to the next step in our work.

3.3.1 The Maassen-Ufflink Entropic Uncertainty Relation

It was Maassen and Ufflink who first prepared an entropic uncertainty relation for conjugate variables[8].

They formulated the final form of the relation in terms of Shannon entropy,and the bound was set by a 'complementarity measure'. The relation was as follows:

$$H(X) + H(Y) \geq \log \frac{1}{c} \tag{3.24}$$

where c is a measure of complementarity or *overlap* between the eigenvectors of the two observables, and is given by:

$$c = \max_{x',z'} |\langle X^x | Z^z \rangle|^2 \tag{3.25}$$

We here prove the stronger relation, which was provided by Coles *et al*[4][5]

$$H(X) + H(Z) \geq \log \frac{1}{c} + H(\rho_A) \tag{3.26}$$

Proof. First, we consider the measurement map X

$$\mathcal{X} = \sum_x |X^x\rangle \langle X^x| (\cdot) |X^x\rangle \langle X^x|$$

which, when acted on the classical state ρ_A gives the state ρ_X.
Note that $(\cdot)$ refers to any state on wich the measurement operator may act.

$$\mathcal{X}_{A\to X}(\rho_A) = \sum_x |X^x\rangle \langle X^x| \rho_A |X^x\rangle \langle X^x| \tag{3.27}$$

We can show that the Shannon entropy of ρ_X is equal to its von-Neumann entropy.

Thus,

$$\begin{aligned} H(X) &= -Tr[\rho_X \log \rho_X] \\ &= -Tr\left(\sum_x |X^x\rangle \langle X^x| \rho_A |X^x\rangle \langle X^x| \log \mathcal{X}(\rho_A)\right) \\ &= \sum_{k,x} \langle k|X^x\rangle \langle X^x| \rho_A |X^x\rangle \langle X^x| \log \mathcal{X}(\rho_A) |k\rangle \\ &= -\sum_{k,x} \langle X^x| \log \mathcal{X}(\rho_A) |k\rangle \langle k|X^x\rangle \langle X^x| \rho_A |X^x\rangle \\ &= -\sum_{,x} \langle X^x| \log \mathcal{X}(\rho_A) |X^x\rangle \langle X^x| \rho_A |X^x\rangle \\ &= -\sum_{x} \langle X^x| \rho_A |X^x\rangle \langle X^x| \log \mathcal{X}(\rho_A) |X^x\rangle \\ &= -Tr(\rho_A \log \rho_A) \end{aligned}$$

Thus leaving us with:

$$H(X) = -Tr(\rho_A \log \rho_A) \tag{3.28}$$

We now formulate $H(X)$ in terms of relative entropy as follows:

$$\begin{aligned} D(\rho_A||\mathcal{X}(\rho_A)) + H(\rho_A) &= Tr[\rho_A \log \rho_A - \rho_A \log \mathcal{X}(\rho_A)] - Tr[\rho_A \log \mathcal{X}(\rho_A)] && (3.29) \\ &= -Tr[\rho_A \log \mathcal{X}(\rho_A)] && (3.30) \\ &= H(X) && (3.31) \end{aligned}$$

Thus,

$$H(X) = D(\rho_A||\mathcal{X}(\rho_A)) + H(\rho_A)$$

Now, we apply the map $\mathcal{Z}_{A\to Z(\cdot)} = \sum_z |Z^z\rangle \langle Z^z| (\cdot) |Z^z\rangle \langle Z^z|$ to both arguments in the relative entropy term of eq (3.29) ($(\cdot)$ being the state that we are making the measurements on), and apply the Data Processing inequality to obtain:

$$\begin{aligned} D(\rho_A||\mathcal{X}(\rho_A)) &\geq D(\mathcal{Z}(\rho_A)||\mathcal{Z} \circ \mathcal{X}(\rho_A)) \\ &= D(\rho_Z||\mathcal{Z} \circ \mathcal{X}(\rho_A)) \end{aligned}$$

And as for $\mathcal{Z} \circ \mathcal{X}(\rho_A)$,

$$\begin{aligned}\mathcal{Z} \circ \mathcal{X}(\rho_A) &= \sum_z |Z^z\rangle \langle Z^z| \sum_x |X^x\rangle \langle X^x| \rho_A |X^x\rangle \langle X^x|Z^z\rangle \langle Z^z| \\ &= \sum_{z,x} |Z^z\rangle \langle Z^z|X^x\rangle \langle X^x| \rho_A |X^x\rangle \langle X^x|Z^z\rangle \langle Z^z| \\ &= \sum_z |Z^z\rangle \langle Z^z| \sum_x |\langle X^x|Z^z\rangle|^2 \langle X^x| \rho_A |X^x\rangle\end{aligned}$$

Now, $D(\rho_Z||\mathcal{Z} \circ \mathcal{X}(\rho_A))$ becomes:

$$D(\rho_Z||\mathcal{Z} \circ \mathcal{X}(\rho_A)) = Tr[\rho_Z \log \rho_Z] - Tr[\rho_Z \log (\mathcal{Z} \circ \mathcal{X}(\rho_A))] \tag{3.32}$$

$$= -H(\rho_Z) - Tr\left[\left(\sum_z |Z^z\rangle \langle Z^z| \rho_A |Z^z\rangle \langle Z^z|\right) \log (\mathcal{Z} \circ \mathcal{X}(\rho_A))\right] \tag{3.33}$$

Simplifying the second term on the RHS leaves us with:

$$Tr\left[\left(\sum_z |Z^z\rangle \langle Z^z| \rho_A |Z^z\rangle \langle Z^z|\right) \log (\mathcal{Z} \circ \mathcal{X}(\rho_A))\right] = \sum_z \langle Z^z|\rho_A|Z^z\rangle \log \left(\sum_x |\langle X^x|Z^z\rangle|^2 \langle X^x|\rho_A|X^x\rangle\right)$$

Thus, eq (3.33) becomes:

$$D(\rho_Z||\mathcal{Z} \circ \mathcal{X}(\rho_A)) = -H(\rho_Z) - \sum_z \langle Z^z|\rho_A|Z^z\rangle \log \left(\sum_x |\langle X^x|Z^z\rangle|^2 \langle X^x|\rho_A|X^x\rangle\right) \tag{3.34}$$

We are aware that the logarithm is a **monotonic function**. i.e., *the value of the function increases with increase in its input.*

So,

$$-\sum_z \langle Z^z|\rho_A|Z^z\rangle \log \left(\sum_x |\langle X^x|Z^z\rangle|^2 \langle X^x|\rho_A|X^x\rangle\right) \geq -\sum_z \langle Z^z|\rho_A|Z^z\rangle \log \left(\max_{x',z'} |\langle X^x|Z^z\rangle|^2 \sum_x \langle X^x|\rho_A|X^x\rangle\right)$$

In the above equation, $-\sum_z \langle Z^z|\rho_A|Z^z\rangle$ and $\sum_x \langle X^x|\rho_A|X^x\rangle$ are basically just the trace of ρ_A and will sum to 1 , leaving us with:

$$-\sum_{z}\langle Z^z|\rho_A|Z^z\rangle \log\left(\sum_{x}|\langle X^x|Z^z\rangle|^2\langle X^x|\rho_A|X^x\rangle\right) \geq -\log\max_{x',z'}|\langle X^x|Z^z\rangle|^2 \quad (3.35)$$

Putting this in eq (3.34) gives us:

$$D(\rho_Z||\mathcal{Z}\circ\mathcal{X}(\rho_A)) \;\geq\; -H(\rho_Z)+\log\frac{1}{\max\limits_{x',z'}|\langle X^x|Z^z\rangle|^2} \quad (3.36)$$

We already know:

$$H(X) = D(\rho_A||\mathcal{X}\rho_A) + H(\rho_A)$$

Hence,

$$H(X) - H(\rho_A) = D(\rho_A||\mathcal{X}\rho_A)$$

But we already have:

$$D(\rho_A||\mathcal{X}\rho_A) \;\geq\; D(\rho_Z||\mathcal{Z}\circ\mathcal{X}\rho_A)$$

and

$$D(\rho_Z||\mathcal{Z}\circ\mathcal{X}(\rho_A)) \;\geq\; -H(\rho_Z)+\log_2\frac{1}{\max\limits_{x',z'}|\langle X^x|Z^z\rangle|^2}$$

So, we now have,

$$H(X) - H(\rho_A) \geq D(\rho_Z||\mathcal{Z}\circ\mathcal{X}\rho_A)$$

which in turn gives us

$$H(X) - H(\rho_A) \geq -H(\rho_Z) + \log_2\frac{1}{c}$$

Rearranging the terms, we end up at the Maassen-Ufflink relation.

$$H(X) + H(Z) \geq \log_2\frac{1}{c} + H(\rho_A)$$

■

We see that the above bound is stronger than the one given by Maassen-Ufflink, and that on dropping the $H(\rho_A)$ term (which is always positive), we may recover the original bound of $H(X) + H(Y) \geq \log\frac{1}{c}$.

We have now gained an understanding of uncertainty relations and entropy, and most importantly, the notion of conditional entropies (section 3.1.6). We can now move on to the next section where we will try to formulate a relation of our own based on the understanding gained in the past two sections.

Chapter 4

Towards a New Uncertainty Relation

The ultimate aim of the project carried out was to try and develop a relation putting a bound on the uncertainties involved when making measurements corresponding to a pair non-commuting observables in the presence of a quantum memory.
This idea was suggested by my guide based on my reading of the 2010 paper by Berta et al, which introduces an uncertainty relation expressed as a *sum* of conditional entropies. All the material built up till now helps to understand the germ of the idea contained in the paper, and hopefully transfer the insights gained to the project at hand.

4.1 The Entropic Uncertainty Relation with Memory

The paper by Berta et al [2] formulates an entropic uncertainty relation incorporating quantum memory as:

$$H(X|B) + H(Y|B) \geq \log_2 \frac{1}{c} + H(A|B). \tag{4.1}$$

As per this relation, the uncertainty associated with the measurement of an observable on one part of a bipartite system is dependent on the amount of entanglement between the two parts comprising it. Now, to understand this notion, we must first clear a common confusion that exists around what is actually meant by an uncertainty relation.

The concept of what an uncertainty relation actually tries to encapsulate is best conveyed through an 'uncertainty game' [2][4] detailed below and accompanied by the diagram above:

- Here, before the "game" begins, Bob and Alice agree on two measurements X and Y, of which one will be made on each particle Bob sends to Alice.

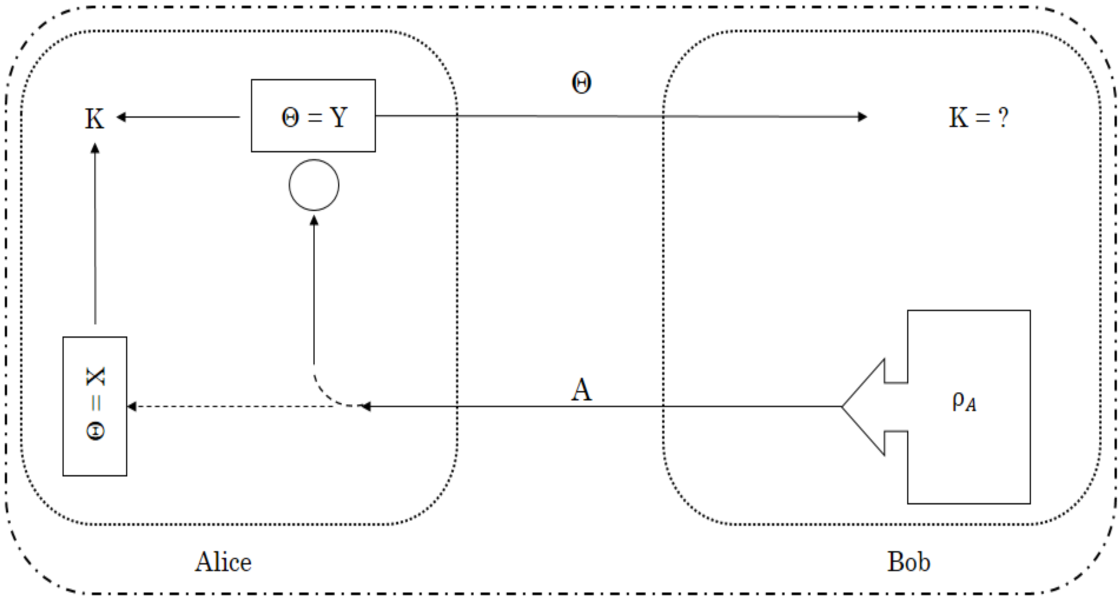


Figure 4.1: An uncertainty Game [4]

- Once the game begins, Bob will send a particle, which he prepared in a state of his own choice, to Alice.
- Alice will make one of the two possible measurements (either X or Y) on the system, and relay back to Bob which observable she measured.
- Bob then tries to guess what was the outcome of Alice's measurement as accurately as he can. The role an uncertainty relation plays here is to put a bound on how accurately Bob can guess the value of the measurement outcome from just having information about the *choice of measurement*.

$H(X|B)$ *and* $H(Y|B)$ here represent the uncertainty associated with measurement of the observables X and Y *given* that information about the measurement is stored in a quantum memory (*we note here that the measurements are carried out on system A and B acts as the quantum memory*)

This formulation gives rise to a number of interesting cases:

(a) If A and B are maximally entangled, then $H(A|B) = -\log_2 n$, n being the dimension of the system. We know that $\log_2 n$ is the maximum possible entropy for a n - dimensional system. Hence $\log_2 \frac{1}{c} - \log_2 n \leq 0$, which reduces the bound to:

$$H(X|B) + H(Y|B) \geq 0 \tag{4.2}$$

As we see, a maximally entangled memory can bring down uncertainty in the measurements to zero.

$(\boldsymbol{b})$ If A and B are not entangled, then on the RHS,

$$H(A|B) = H(A, B) - H(B)$$

Now, B being a pure state, $H(B) = 0$, and $H(A, B)$ will be positive. The right hand side can thus be reduced to $\log_2 \frac{1}{c}$.
Now, we know that conditioning reduces entropy. Hence,

$$H(X) \geq H(X|B) \qquad\qquad H(Y) \geq H(Y|B)$$

The relation now basically reduces to the Maassen-Ufflink bound.

$$H(X) + H(Y) \geq \log_2 \frac{1}{c}$$

$(\boldsymbol{c})$ In the case where the two subsystems are (not maximally) entangled, the term on the RHS $H(A|B)$ will be negative (as is the case with entanglement). This reduces the already existing bound by Maassen-Ufflink, and is where we get to witness conditioning actually reducing uncertainties

$$H(X|B) + H(Y|B) \geq \log_2 \frac{1}{c} - k \tag{4.3}$$

k being the negative quantity that is $H(A|B)$

Another thing we must note at this point is that, unlike what Heisenberg put forth in his paper, there is no notion of simultaneous measurement of conjugate variables, or even of successive measurements on the same system so to speak. Bob is creating a large number of *identical states*, and sending it to Alice, who performs *just* one operation (either X or Y). As explained by Maccone-Pati[9] the relation merely states that, once we tabulate the values Bob has guessed, find their means, and consequently the associated standard deviations, they will follow the given uncertainty relations. This is the purely statistical interpretation of an uncertainty *relation*, as opposed to the uncertainty *principle*, and is what all these relations, including the one we intend to formulate, try to say.

The concept behind conditioning grasped, what was needed next was to develop a toolkit to formulate an uncertainty relation in terms of variance/standard deviation.

While these quantities incorporate themselves in theories which does not involve conditioning without any hassle, things get a little difficult when memory and conditioning are thrown into the mix.
The main difficulty encountered is the inability to express conditioned variance/standard deviation corresponding to say, $\Delta(X|B)$ and $\Delta(Y|B)$, (if we are to look at the illustration above) in terms of an operator. This is possible quite easily for non conditioned standard deviations.

i.e., it is possible to express ΔA in terms of the operator A as $\Delta A = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2$, whereas there was no such operator $(\widehat{X|M})$ which would enable a similar expression for conditional variances / standard deviations.

It is for this reason that we had to look at research that formulates explicit relations for conditional uncertainties that circumvent the need for such conditional operators.

4.2 Standard deviation Based Uncertainty Measures

The paper by Sazim *et al* defines a conditional 'uncertainty' defined in terms of standard deviation as follows [14]:

$$\Delta(A|M) = \Delta(A + M) - \Delta M. \tag{4.4}$$

We note that this is not an operator, it is merely a scalar. It displays properties similar to those exhibited by conditional entropy [14]:

- Here too, conditioning one unknown using another 'known' reduces uncertainty,

$$\Delta(A|M) \leq \Delta A. \tag{4.5}$$

- Like entropy, this too obeys the chaining principle i.e.,

$$\Delta\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} (A_i|A_{i-1} + A_{i-2} + ... + A_1), \tag{4.6}$$

which will prove to be very convenient for us later on.

4.2.1 Sum Uncertainty Relations for Standard Deviation

A very useful relation that will be integral in going ahead with finding standard deviation/variance based uncertainty relations is one that reveals a fundamental property of the quantity of standard deviation, brought about by Pati and Sahu[12], which states:

"Quantum fluctuations in the sum of any 2 observables is always less than or equal to the sum of their individual fluctuations"

$$\Delta(A + B) \leq \Delta A + \Delta B \tag{4.7}$$

Proof. The proof begins by defining two un-normalized vectors as shown:

$$|\psi_1\rangle \equiv (\hat{A} - \langle\hat{A}\rangle\mathbb{I})\,|\psi\rangle \qquad\qquad |\psi_2\rangle \equiv (\hat{B} - \langle\hat{B}\rangle\mathbb{I})\,|\psi\rangle$$

Now, we take the norm of the sum of $|\psi_1\rangle + |\psi_2\rangle$:

$$\begin{aligned} \||\psi_1\rangle + |\psi_2\rangle\|^2 &= (\langle\psi_1| + \langle\psi_2|)(|\psi_1\rangle + |\psi_2\rangle) && (4.8)\\ &= \langle\psi_1|\psi_1\rangle + \langle\psi_1|\psi_2\rangle + \langle\psi_2|\psi_1\rangle + \langle\psi_2|\psi_2\rangle && (4.9)\\ &= \||\psi_1\rangle\|^2 + \||\psi_2\rangle\|^2 + 2Re\langle\psi_1|\psi_2\rangle && (4.10)\\ &= \Delta A^2 + \Delta B^2 + 2Re\langle\psi_1|\psi_2\rangle. && (4.11) \end{aligned}$$

We know:

$$Re\langle\psi_1|\psi_2\rangle \leq |\langle\psi_1|\psi_2\rangle|. \tag{4.12}$$

We invoke the Cauchy-Schwarz inequality and state:

$$\langle\psi_1|\psi_1\rangle\langle\psi_2|\psi_2\rangle \geq |\langle\psi_1|\psi_2\rangle|^2$$
$$\||\psi_1\rangle\|^2\,\||\psi_2\rangle\|^2 \geq |\langle\psi_1|\psi_2\rangle|^2,$$

which finally leaves us with:

$$\||\psi_1\rangle\|\,\||\psi_2\rangle\| \geq |\langle\psi_1|\psi_2\rangle|. \tag{4.13}$$

Now, eq (4.13) in eq (4.12) gives:

$$Re\langle\psi_1|\psi_2\rangle \leq \|\psi_1\|^2\,\|\psi_2\|^2\,, \tag{4.14}$$

and putting eq (4.14) in eq (4.10) gives:

$$\||\psi_1\rangle + |\psi_2\rangle\| \leq \Delta A^2 + \Delta B^2 + 2\Delta A\Delta B,$$

which will condense to:

$$\| |\psi_1\rangle + |\psi_2\rangle \| \leq (\Delta A + \Delta B)^2. \tag{4.15}$$

We now expand $\| |\psi_1\rangle + |\psi_2\rangle \|^2$ in terms of our un-normalized vectors $|\psi_1\rangle \equiv (\hat{A} - \langle \hat{A} \rangle \mathbb{I}) |\psi\rangle$ and $|\psi_2\rangle \equiv (\hat{B} - \langle \hat{B} \rangle \mathbb{I}) |\psi\rangle$:

$$\begin{aligned}
\| |\psi_1\rangle + |\psi_2\rangle \|^2 &= \left\| (\hat{A} - \langle \hat{A} \rangle + \hat{B} - \langle \hat{B} \rangle) |\psi\rangle \right\|^2 \\
&= \langle \psi | [(\hat{A} + \hat{B}) - (\langle \hat{A} \rangle + \langle \hat{B} \rangle)][(\hat{A} + \hat{B}) - (\langle \hat{A} \rangle + \langle \hat{B} \rangle)] |\psi\rangle \\
&= \langle (\hat{A} + \hat{B})^2 \rangle + \langle \hat{A} + \hat{B} \rangle^2 - 2\langle (\langle \hat{A} \rangle + \langle \hat{B} \rangle)(\hat{A} + \hat{B}) \rangle \\
&= \langle (\hat{A} + \hat{B})^2 \rangle + \langle \hat{A} + \hat{B} \rangle^2 - 2(\langle \hat{A} \rangle + \langle \hat{B} \rangle)^2 \\
&= \langle (\hat{A} + \hat{B})^2 \rangle - \langle A + B \rangle^2 \\
&= \Delta(A + B)^2.
\end{aligned}$$

Since $\| |\psi_1\rangle + |\psi_2\rangle \| \leq (\Delta A + \Delta B)^2$, and since $\| |\psi_1\rangle + |\psi_2\rangle \|^2 = \Delta(A + B)^2$, we obtain:

$$\Delta(A + B)^2 \leq (\Delta A + \Delta B)^2, \tag{4.16}$$

leaving us, upon taking the square root, with [12]:

$$\Delta(A + B) \leq \Delta A + \Delta B. \tag{4.17}$$

■

Now that we have developed all the tools necessary for devising a relation incorporating memory, we move towards our attempt at actually formulating one for ourselves.

4.3 Finding a Rudimentary Uncertainty Bound

As a first exercise towards formulating our uncertainty relation, we try to set a preliminary bound for the uncertainty in measuring two complementary observables X and Y such that $[X, Y] \neq 0$ when a memory is involved.
For this, we start with the definition of the expression $\Delta(X|M)$, which according to Sazim, Pati *et al*, is given by:

$$\Delta(X|M) = \Delta(X + M) - \Delta M. \tag{4.18}$$

Now, we proceed formulating this relation along the lines of the Maccone-Pati uncertainty relation i.e., *as a sum of variances*.

$$\Delta(X|M)^2 + \Delta(Y|M)^2 = [\Delta(X+M) - \Delta M]^2 + [\Delta(Y+M) - \Delta M]^2, \tag{4.19}$$

which expands to

$$\Delta(X|M)^2 + \Delta(Y|M)^2 = \Delta(X+M)^2 - 2\Delta M\Delta(X+M) + \Delta M^2 \\ + \Delta(Y+M)^2 - 2\Delta M\Delta(Y+M) + \Delta M^2. \tag{4.20}$$

Now, we know that $[\Delta(X+M) - \Delta(Y+M)]^2 \geq 0$. This means that:

$$\Delta(X+M)^2 + \Delta(Y+M)^2 \geq 2\Delta(X+M)\Delta(Y+M). \tag{4.21}$$

Using eq (4.21), we may now write eq (4.20) as:

$$\Delta(X|M)^2 + \Delta(Y|M)^2 \geq 2\Delta(X+M)\Delta(Y+M) + 2\Delta M\Delta M \\ - 2\Delta M\Delta(X+M) - 2\Delta M\Delta(Y+M), \tag{4.22}$$

and on grouping like terms, we get:

$$\Delta(X|M)^2 + \Delta(Y|M)^2 \geq 2\Delta(X+M)[\Delta(Y+M) - \Delta M] \\ - 2\Delta M[\Delta(Y+M) - \Delta M], \tag{4.23}$$

which condenses to:

$$\Delta(X|M)^2 + \Delta(Y|M)^2 \ \geq \ 2[\Delta(X+M) - \Delta M][\Delta(Y+M) - \Delta M] \tag{4.24}$$

$$= 2\Delta(X|M)\Delta(Y|M). \tag{4.25}$$

Thus leaving us with the rudimentary bound

$$\Delta(X|M)^2 + \Delta(Y|M)^2 \ \geq \ 2\Delta(X|M)\Delta(Y|M). \tag{4.26}$$

A bound of this form is one that is universal among standard deviation based uncertainty relations. The fact that it holds for our attempt at the relation is a promising sign that we may be on the right track, or at least, such a relation can exist.

4.4 The Connected Correlator

Before we venture out on attempting an uncertainty relation, we need to look into a quantity that can be used to quantify entanglement, since entanglement is, after all what provides the element of a 'memory' in our relation, and we must take care to introduce this element of entanglement in our relation as well.

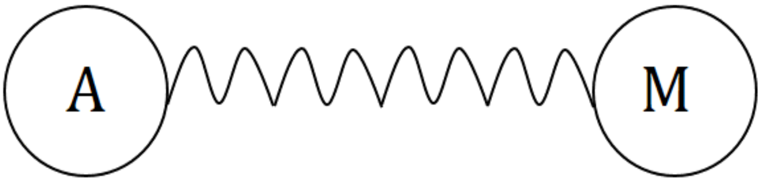


Figure 4.2: The Subsystems of Our Entangled System

We see that we have two subsystems A and M living in two different Hilbert spaces, and it is to be understood that we are acting the observables X and Y on subsystem A, belonging to the Hilbert Space $\mathcal{H}_A$. This would, undoubtedly affect its entangled partner M in the Hilbert Space $\mathcal{H}_B$, and these correlations are what will provide the conditioning effect, bringing down the uncertainty in the measurements of X and Y.
It is therefore necessary to incorporate the Connected Correlator [1] [19], a measure which quantifies the entanglement in the state ρ_{AM}. Now, the paper by Bagchi et al gives the connected correlator for the bipartite state ρ_{AB} with two observables A and B acting on each subsystem A and B in their respective spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ as:

$$CC_{AB} = Tr[(A \otimes B)\rho_{AB}] - Tr[A\rho_A]Tr[B\rho_B]. \tag{4.27}$$

A non-zero value for the function indicates entanglement in case of a pure state, although in case of mixed states, the function gives a non-zero value for both entangled and non-entangled states.

4.5 Attempt at an Uncertainty Relation

Using a motivation similar to the paper by Berta et al [2], we consider the case where a system consists of two parts, and we act a measurement operator on one of the two subsystems, while keeping the other as our 'memory'. In such a case, we need our relation to be of a form resembling:

$$\Delta(X|M)^2 + \Delta(Y|M)^2 \geq k(A, B, M),$$

where k is a bound, which should be a function of the operators $X,\ Y$ and the memory M. Now, we have expressions for $\Delta(X|M)$ and $\Delta(Y|M)$ in terms of the relations given in Pati, Sazim et al [14] as:

$$\Delta(X|M) = \Delta(X + M) - \Delta M \qquad \Delta(Y|M) = \Delta(Y + M) - \Delta M \tag{4.28}$$

Thus, we have a starting point to our relation in terms of conditional uncertainty

$$\Delta(X|M)^2 + \Delta(Y|M)^2 = [\Delta(X + M) - \Delta M]^2 + [\Delta(Y + M) - \Delta M]^2. \tag{4.29}$$

Now, going over to the RHS, we expand the term $[\Delta(X + M) - \Delta M]^2$.

$$\Delta(X|M)^2 = [\Delta(X + M) - \Delta M]^2 \tag{4.30}$$

$$= \Delta(X + M)^2 - 2\Delta M \Delta(X + M) + \Delta(M)^2 \tag{4.31}$$

We now turn our focus to the RHS of eq (4.34) and concern ourselves with the term $\Delta(X + M)^2$

$$\Delta(X + M)^2 = \langle[(X + M) - \langle X + M\rangle \mathbb{1}]\rangle \tag{4.32}$$

$$= \langle(X + M)^2 - 2(X + M)\langle X + M\rangle \mathbb{1} + \langle X + M\rangle^2\rangle \tag{4.33}$$

$$= \langle(X + M)^2\rangle - 2\langle X + M\rangle^2 + \langle X + M\rangle^2. \tag{4.34}$$

The next step is an integral one, and is what will make our relation incorporate the fact that we are dealing with two different systems in different *Hilbert spaces*, and because of this, the measurements X and M should also be carried out in the corresponding different Hilbert spaces. This consideration will require that we write:

$$(X + M)^2 = (X \otimes \mathbb{1} + \mathbb{1} \otimes M)^2 \tag{4.35}$$

$$= (X \otimes \mathbb{1})^2 + 2(X \otimes \mathbb{1})(\mathbb{1} \otimes M) + (\mathbb{1} \otimes M)^2 \tag{4.36}$$

$$= (X \otimes \mathbb{1})^2 + 2(X \otimes M) + (\mathbb{1} \otimes M)^2 \tag{4.37}$$

The expectation of this quantity will now be given by;

$$\langle(X + M)\rangle^2 = \langle(X \otimes \mathbb{1})^2\rangle + 2\langle(X \otimes M)\rangle + \langle(\mathbb{1} \otimes M)^2\rangle \tag{4.38}$$

$$= \langle X^2\rangle + 2\langle X \otimes M\rangle + \langle M^2\rangle \tag{4.39}$$

Putting eq (4.39) in eq (4.34) will give us:

$$\Delta(X+M)^2 = \langle X^2\rangle + 2\langle X\otimes M\rangle + \langle M^2\rangle - \langle X+M\rangle^2 \tag{4.40}$$

$$= \langle X^2\rangle + \langle M^2\rangle - \langle X\rangle^2 - \langle M\rangle^2 - 2\langle X\rangle\langle M\rangle + 2\langle X\otimes M\rangle \tag{4.41}$$

leaving us with:

$$\Delta(X+M)^2 = \Delta X^2 + \Delta M^2 + 2[\langle X\otimes M\rangle - \langle X\rangle\langle M\rangle] \tag{4.42}$$

We substitute the value for $\Delta(X+M)^2$ obtained from eq (4.42) in eq (4.31), to obtain an expression for $\Delta(X|M)^2$:

$$\Delta(X|M)^2 = \Delta X^2 + \Delta M^2 + 2[\langle X\otimes M\rangle - \langle X\rangle\langle M\rangle] + \Delta M^2 - 2\Delta M\Delta(X+M) \tag{4.43}$$

Following the same procedure for the expression $\Delta(Y+M)^2$, we get:

$$\Delta(Y+M)^2 = \Delta Y^2 + \Delta M^2 + 2[\langle Y\otimes M\rangle - \langle Y\rangle\langle M\rangle] + \Delta M^2 - 2\Delta M\Delta(Y+M), \tag{4.44}$$

and from there, we obtain $\Delta(Y|M)^2$ as:

$$\Delta(Y|M)^2 = \Delta Y^2 + \Delta M^2 + 2[\langle Y\otimes M\rangle - \langle Y\rangle\langle M\rangle] + \Delta M^2 - 2\Delta M\Delta(Y+M). \tag{4.45}$$

Now that we finally have the values of the conditional variances $\Delta(X|M)^2$ and $\Delta(Y|M)^2$, we substitute these results from eq (4.43) and eq (4.45) in eq (4.28), thereby obtaining:

$$\Delta(X|M)^2 + \Delta(Y|M)^2 = \Delta X^2 + \Delta Y^2 + 4\Delta M^2 + 2[\langle X\otimes M\rangle + \langle Y\otimes M\rangle - \langle X+Y\rangle\langle M\rangle] - 2\Delta M[\Delta(X+M) + \Delta(Y+M)]. \tag{4.46}$$

We have from Bagchi et al [1] that the quantity $[\langle X\otimes M\rangle + \langle Y\otimes M\rangle - \langle X+Y\rangle\langle M\rangle]$ is in fact the *connected correlator* $CC_{(X\otimes M, Y\otimes M)}$.

Thus, eq (4.46) becomes:

$$\Delta(X|M)^2 + \Delta(Y|M)^2 = \Delta X^2 + \Delta Y^2 + 4\Delta M^2 + 2CC_{X\otimes M, Y\otimes M} - 2\Delta M[\Delta(X+M) + \Delta(Y+M)]. \tag{4.47}$$

On applying the sum uncertainty relation by Sazim, Pati *et al*, we have:

$$\Delta(X+M) + \Delta(Y+M) \leq \Delta X + \Delta M + \Delta Y + \Delta M, \tag{4.48}$$

which will enable us to write:

$$\Delta(X|M)^2 + \Delta(Y|M)^2 \geq \Delta X^2 + \Delta Y^2 + 4\Delta M^2 + 2CC_{)X\otimes M,Y\otimes M)} - 2\Delta M[\Delta X + \Delta M + \Delta Y + \Delta M], \quad (4.49)$$

which simplifies to:

$$\Delta(X|M)^2 + \Delta(Y|M)^2 \geq \Delta X^2 + \Delta Y^2 + 2CC_{(X\otimes M,Y\otimes M)} - 2\Delta M[\Delta X + \Delta Y]. \quad (4.50)$$

If we examine the first two terms on the RHS of the relation, it seems possible that we could replace them with either one of the bounds specified by Maccone-Pati in their uncertainty relation (2.19), which, if we recall is of the form:

$$\Delta X^2 + \Delta Y^2 \geq \max(\mathcal{L}_1, \mathcal{L}_2), \quad (4.51)$$

with

$$\mathcal{L}_1 = \frac{1}{2}|\langle\psi^{\perp}_{X+Y}|\, X + Y\, |\psi\rangle|^2$$
$$\mathcal{L}_2 = \pm i[\hat{X}, \hat{Y}] + |\langle\psi|\, \hat{X} \pm i\hat{Y}\, |\psi^{\perp}\rangle|^2.$$

Unfortunately, try as we may, there could not be a resolution regarding the final term $2\Delta M[\Delta X + \Delta Y]$ on the RHS, despite attempts at condensing it into a more useful form. This led us to turn our attention, albeit briefly, to alternative approaches towards achieving our goal.

4.6 The Conditional Amplitude Operator: A Potential Approach Towards a New Relation

Another prospect that was explored in the duration of the work was the conditional amplitude operator proposed by Cerf and Adami[3].

The fact that conditional uncertainty could not be expressed in terms of an operator was a major reason that led us to adopt the path detailed in the above section. But the conditional amplitude operator actually provides an explicit density matrix representation for the conditional states which would arise due to the presence of quantum memory. This leads us to believe that we may have a prospective candidate towards formulating the desired relation in this operator.

Cerf and Adami devised the conditional amplitude operator as a quantum analogue to the expression for classical conditional probability:

$$p(a|b) = \frac{p(a,b)}{p(b)}. \tag{4.52}$$

Modelled after the above expression. the operator is formulated thus:

$$\rho_{A|B} = \exp_2[\log_2 \rho_{AB} - \log_2(\mathbb{1} \otimes \rho_B)], \tag{4.53}$$

which, on applying the Trotter symmetrization reduces to:

$$\rho_{A|B} = \lim_{n\to\infty} [\rho_{AB}^{1/n} (\mathbb{1} \otimes \rho_B)^{-1/n}]^n. \tag{4.54}$$

A point to be noted is that this operator is defined only in the support of ρ_{AB} (i,e., only those eigenvectors which have *non-zero* eigenvalues),due to the fact that $sup(\rho_{AB}) \cap ker(\mathbb{1} \otimes \rho_B) = \varnothing$.
This restriction is important, because computing $\rho_{A|B}$ for states in the kernel of $(\mathbb{1} \otimes B)$ can lead to singularities.

The idea behind using this approach was that once we had a conditional amplitude operator defined, we could use the trace to calculate variance in the form $\Delta(X|M)^2 = Tr[X^2 \rho_{X|M}] - Tr[X \rho_{X|M}]$, as is done for the case of usual density matrices. Further research in this direction stands to be done.

Chapter 5

Future Scope of the Project

As we saw, a conclusion could not be reached using the conditional variance (eq (4.18))attempt we detailed in section (4.5). We did however, consider how the conditional amplitude operator in section (4.6) could be put to use towards finding a relation, but the operator was proving to be quite tricky to manipulate, and it has been understood that more time would be needed to gain some familiarity with the operator.

The future course of the work, naturally, would be to try to gain a deeper understanding of the conditional amplitude operator and hopefully try and use it to construct our uncertainty relation.

Appendices

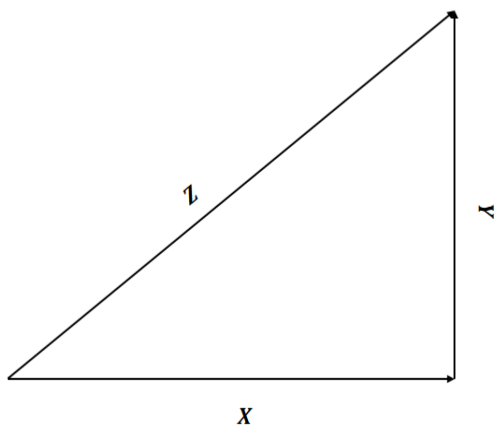
Appendix A

The Cauchy-Schwarz Inequality

The Triangle and Cauchy-Schwarz Inequalities

The Triangle Inequality

This expression is basically an encapsulation of the fact that the sum of the lengths of any two sides of a triangle cannot be less than the length of its third side. Considering the image on the right to be an illustration of a triangle in a linear vector space, the above condition translates to:



$$\left\|\vec{X}\right\| + \left\|\vec{Y}\right\| \geq \left\|\vec{Z}\right\|.$$

But $\left\|\vec{Z}\right\|$ is basically $\left\|\vec{X}\right\| + \left\|\vec{Y}\right\|$.
Hence the inequality becomes:

$$\left\|\vec{X}\right\| + \left\|\vec{Y}\right\| \geq \left\|\vec{X} + \vec{Y}\right\|$$

The Cauchy-Schwarz Inequality

Proof. We start off from the triangle inequality:

$$\left\|\vec{X}\right\| + \left\|\vec{Y}\right\| \geq \left\|\vec{X} + \vec{Y}\right\|$$

and square the equation on both sides, giving us:

$$\left(\left\|\vec{X}\right\| + \left\|\vec{Y}\right\|\right)^2 \geq \left\|\vec{X} + \vec{Y}\right\|^2$$

$$\Rightarrow \quad \left\|\vec{X}\right\|^2 + \left\|\vec{Y}\right\|^2 + 2\left\|\vec{X}\right\|\left\|\vec{Y}\right\| \geq \left(\vec{X} + \vec{Y}\right)\left(\vec{X} + \vec{Y}\right)$$

$$\Rightarrow \quad \left\|\vec{X}\right\|^2 + \left\|\vec{Y}\right\|^2 + 2\left\|\vec{X}\right\|\left\|\vec{Y}\right\| \geq \left\|\vec{X}\right\|^2 + \left\|\vec{Y}\right\|^2 + 2\left|\vec{X} \cdot \vec{Y}\right|$$

$$\Rightarrow \quad \left\|\vec{X}\right\| + \left\|\vec{Y}\right\| \geq \left|\vec{X} \cdot \vec{Y}\right|$$

What we have just arrived at, is an alternative form of the triangle inequality.

Upon squaring this, we get the Cauchy-Schwarz inequality:

$$\left\|\vec{X}\right\|^2 \left\|\vec{Y}\right\|^2 \geq \left|\vec{X} \cdot \vec{Y}\right|^2$$

■

Appendix B

The Parallelogram Law

The Parallelogram Law of Vector Addition

The parallelogram law of vector addition plays a pivotal role in the derivation of the Maccone - Pati uncertainty relations, and as such it is important that it be treated in more detail.

The Parallelogram Law

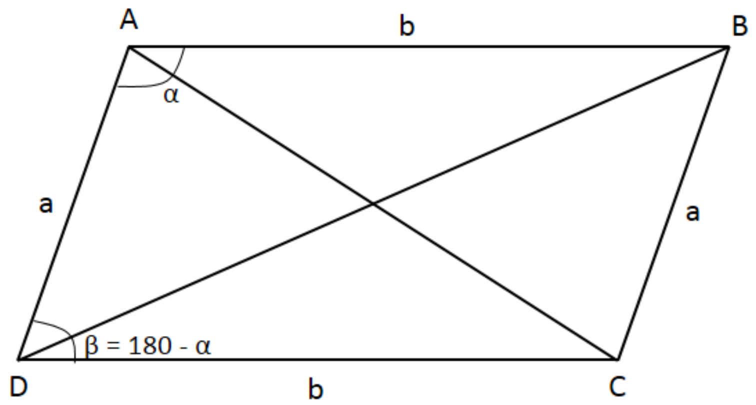


Figure B.1: A Euclidean Parallelogram

The parallelogram law in Euclidean space goes as follows:

$$AC^2 + BD^2 = 2(a^2 + b^2)$$

Proof. The opposite sides of a parallelogram being equal, we name our sides as follows:

$$AD = BC = a$$
$$AB = CD = b$$

Turning our attention now to ΔBAD, we apply the law of cosines to get:

$$BD^2 \; = \; a^2 + b^2 - 2ab\cos\alpha$$

We follow the same procedure with ΔADC, and obtain:

$$AC^2 \; = \; a^2 + b^2 - 2ab\cos\beta$$

The adjacent angles of a parallelogram being supplementary i.e., $\alpha + \beta = \pi$, we have:

$$\beta = \pi - \alpha$$

So that $AC^2 \; = \; a^2 + b^2 - 2ab\cos\beta$ becomes:

$$\begin{aligned} AC^2 &= a^2 + b^2 - 2ab\cos\pi - \alpha \\ &= a^2 + b^2 + 2ab\cos\alpha \end{aligned}$$

Now, on adding the expressions for AC^2 and BD^2, we obtain:

$$\begin{aligned} AC^2 + BD^2 \; &= a^2 + b^2 + 2ab\cos\alpha + \\ & a^2 + b^2 - 2ab\cos\alpha \end{aligned}$$

Thus leaving us with the parallelogram law of vector addition in Euclidean space:

$$AC^2 + BD^2 \; = \; 2(a^2 + b^2)$$

■

Now that we have an idea about what the parallelogram law is trying to state, we move on to the extension of the above expression to Hilbert Spaces:

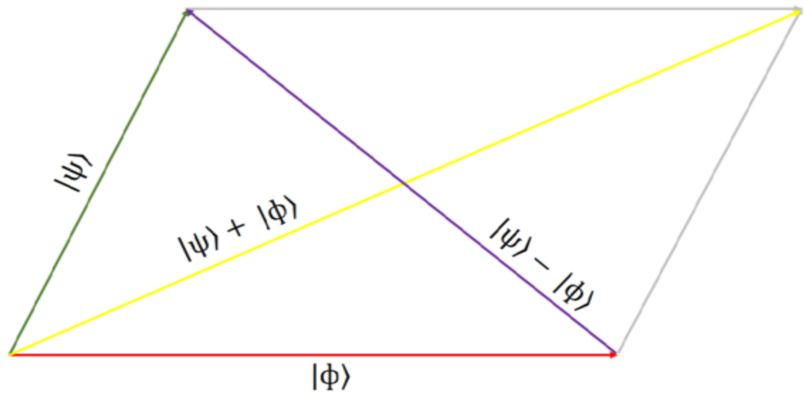


Figure B.2: Parallelogram in an Inner Product Space

The Parallelogram Law in Inner Product Spaces

Here we have two vectors $|\psi\rangle$ and $|\phi\rangle$ as shown above, and the parallelogram law arises as the result of a very straightforward application of the norm in the inner product space[10]:

Proof. We proceed by taking the square of $\|\psi + \phi\|$:

$$\begin{aligned}\|\psi + \phi\|^2 &= (\langle\psi| + \langle\phi|)(|\psi\rangle + |\phi\rangle) \\ &= \langle\psi|\psi\rangle + \langle\phi|\phi\rangle + \langle\phi|\psi\rangle + \langle\psi|\phi\rangle \\ &= \|\phi\|^2 + \|\psi\|^2 + \langle\psi|\phi\rangle + \langle\phi|\psi\rangle\end{aligned}$$

The next step involves taking the square of the expression $\||\psi\rangle - |\phi\rangle\|$

$$\begin{aligned}\|\psi - \phi\|^2 &= (\langle\psi| - \langle\phi|)(|\psi\rangle - |\phi\rangle) \\ &= \langle\psi|\psi\rangle + \langle\phi|\phi\rangle - \langle\phi|\psi\rangle - \langle\psi|\phi\rangle \\ &= \|\phi\|^2 + \|\psi\|^2 - \langle\psi|\phi\rangle - \langle\phi|\psi\rangle\end{aligned}$$

Adding the above two equations together, we end up with

$$\|\psi + \phi\|^2 + \|\psi - \phi\|^2 = 2(\|\phi\|^2 + \|\psi\|^2)$$

Which is what we set out to prove. ∎

Appendix C

The Fano Inequality

The Fano Inequality

We here detail the proof for the Fano inequality, as outlined in Nielsen and Chuang's Quantum Computation and Quantum Information.[11]

Proof. We first declare an *Error Random Variable* **E**.

$$E \equiv \begin{cases} 1 & \text{if } X \neq \widetilde{X} \\ 0 & \text{if } X = \widetilde{X} \end{cases} \tag{C.1}$$

We here note that $H(E) = H(p_e)$

We use the chaining rule for conditional entropies and define

$$\begin{aligned} H(E, X|Y) &= H(E, X, Y) - H(Y) \\ &= H(E, X, Y) + H(X, Y) - H(X, Y) - H(Y) \\ &= H(E|X, Y) + H(X|Y) \end{aligned}$$

We do this step because we will only have access to Y, from which we will have to acquire both E and X.

In the above equation, the value of $H(E|X, Y)$ will be 0, because if we know both X, *and* Y, there is no more uncertainty regarding the value of E.

This leaves us with:

$$H(E, X|Y) = H(X|Y) \tag{C.2}$$

Next, we use the chain rule, but this time, with a different permutation of the random variables:

$$\begin{aligned} H(E,X|Y) &= H(E,X,Y) - H(Y) \\ &= H(E,X,Y) + H(E,Y) - H(E,Y) - H(Y) \\ &= H(X|E,Y) + H(E|Y) \end{aligned}$$

We know conditioning reduces entropy. Hence,

$$\begin{aligned} H(E,X|Y) &\leq H(X|E,Y) + H(E) \\ &= H(X|E,Y) + H(p_e) \end{aligned}$$

But, since $H(E,X|Y) = H(X|Y)$, we have:

$$H(X|Y) \leq H(X|E,Y) + H(p_e) \tag{C.3}$$

Now, we know from the definition of conditional entropy, that

$$H(A|B) = \sum_{b \in \mathcal{B}} p(B=b) H(A|B=b)$$

$\mathcal{B}$ here being the support of B. i.e., the subset of B which has non-zero probabilities. So, in eq(),

$$\begin{aligned} H(X|E,Y) &= \sum_{e=0,1} p(E=e) H(X|E=e,Y) \\ &= p(E=0)H(X|E=0,Y) + p(E=1)H(X|E=1,Y) \end{aligned}$$

But, if $E = 0$, it means that the value of X we get from Y has no error, and hence, no uncertainty associated with it.

This leaves us with:

$$\begin{aligned} H(X|E,Y) &= p(E=1)H(X|E=1,Y) \\ &= p_e H(X|E=1,Y) \end{aligned}$$

We now put an upper bound on $H(X|E=1,Y)$.

We know, if we have $E \equiv 1$, it means we have made a wrong guess, leaving us with at most $(|X|-1)$ possibilities for the value of X i.e., all values in the set X *except* the correct one.

Since the maximum entropy of a system is given by its cardinality i.e., $H(A) \leq \log{(dimA)}$, we may bound the above relation as:

$$H(X|E,Y) \leq p_e \log{(|X|-1)} \tag{C.4}$$

Putting this in eq(A.3) gives us the Fano Bound:

$$H(p_e) + p_e \log{(|X|-1)} \geq H(X|Y) \tag{C.5}$$

■

Bibliography

[1] Shrobona Bagchi, Chandan Datta, and Pankaj Agrawal. “Entanglement dependent bounds on conditional-variance uncertainty relations”. In: *arXiv preprint arXiv:1909.11486* (2019).

[2] Mario Berta et al. “The uncertainty principle in the presence of quantum memory”. In: *Nature Physics* 6.9 (2010), pp. 659–662.

[3] Nicolas J Cerf and Christoph Adami. “Quantum extension of conditional probability”. In: *Physical Review A* 60.2 (1999), p. 893.

[4] Patrick J. Coles et al. “Entropic uncertainty relations and their applications”. In: *Rev. Mod. Phys.* 89 (1 Feb. 2017), p. 015002. DOI: 10.1103/RevModPhys.89.015002. URL: https://link.aps.org/doi/10.1103/RevModPhys.89.015002.

[5] Patrick J. Coles et al. “Uncertainty Relations from Simple Entropic Properties”. In: *Phys. Rev. Lett.* 108 (21 May 2012), p. 210405. DOI: 10.1103/PhysRevLett.108.210405. URL: https://link.aps.org/doi/10.1103/PhysRevLett.108.210405.

[6] Werner Heisenberg. “Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik”. In: *Original Scientific Papers Wissenschaftliche Originalarbeiten*. Springer, 1985, pp. 478–504.

[7] Earle H Kennard. “Zur quantenmechanik einfacher bewegungstypen”. In: *Zeitschrift für Physik* 44.4-5 (1927), pp. 326–352.

[8] Hans Maassen and J. B. M. Uffink. “Generalized entropic uncertainty relations”. In: *Phys. Rev. Lett.* 60 (12 Mar. 1988), pp. 1103–1106. DOI: 10.1103/PhysRevLett.60.1103. URL: https://link.aps.org/doi/10.1103/PhysRevLett.60.1103.

[9] Lorenzo Maccone and Arun K. Pati. "Stronger Uncertainty Relations for All Incompatible Observables". In: *Phys. Rev. Lett.* 113 (26 Dec. 2014), p. 260401. DOI: 10.1103/PhysRevLett.113.260401. URL: https://link.aps.org/doi/10.1103/PhysRevLett.113.260401.

[10] Jonas Maziero. "The Maccone-Pati uncertainty relation". In: *arXiv preprint arXiv:1705.09139* (2017).

[11] Michael A Nielsen and Isaac Chuang. *Quantum computation and quantum information*. American Association of Physics Teachers, 2002.

[12] AK Pati and PK Sahu. "Sum uncertainty relation in quantum theory". In: *Physics Letters A* 367.3 (2007), pp. 177–181.

[13] H. P. Robertson. "The Uncertainty Principle". In: *Phys. Rev.* 34 (1 July 1929), pp. 163–164. DOI: 10.1103/PhysRev.34.163. URL: https://link.aps.org/doi/10.1103/PhysRev.34.163.

[14] Sk Sazim et al. "Mutual uncertainty, conditional uncertainty, and strong subadditivity". In: *Physical Review A* 98.3 (2018), p. 032123.

[15] E Schrödinger. "About Heisenberg uncertainty relation". In: *Proc. Prussian Acad. Sci. Phys. Math. XIX* 293 (1930).

[16] Claude E Shannon. "A mathematical theory of communication". In: *Bell system technical journal* 27.3 (1948), pp. 379–423.

[17] Willi-Hans Steeb and Yorick Hardy. *Problems and solutions in quantum computing and quantum information*. World Scientific Publishing Company, 2018.

[18] Lev Vaidman. "Minimum time for the evolution to an orthogonal quantum state". In: *American journal of physics* 60.2 (1992), pp. 182–183.

[19] Michael M. Wolf et al. "Area Laws in Quantum Systems: Mutual Information and Correlations". In: *Phys. Rev. Lett.* 100 (7 Feb. 2008), p. 070502. DOI: 10.1103/PhysRevLett.100.070502. URL: https://link.aps.org/doi/10.1103/PhysRevLett.100.070502.

[20] Nouredine Zettili. *Quantum mechanics: concepts and applications*. 2003.