

Violating the CHSH Inequality: Empirical Evidence of Quantum Nonlocality (PHY 3035 Quantum Mechanics Honors)

Damien Koon

April 2025

Abstract

The Clauser–Horne–Shimony–Holt (CHSH) inequality provides an experimentally testable framework for Bell’s theorem, and led to John Clauser being awarded the 2022 Nobel Prize in Physics for his work involving this pioneering quantum information experiment. Experimentally proving Bell’s theorem using the CHSH experiment shows that quantum mechanics is incompatible with local hidden variable theories. This paper implements two CHSH experiments: One using Qiskit’s Aer simulator and another one IBM quantum hardware. In each, a Bell state is prepared and the expectation values for the CHSH parameter are measured, demonstrating violations of the classical bound, $|S| \leq 2$. Our results empirically validate Bell’s theorem while bridging the theoretical foundations of quantum computation.

Contents

Introduction	3
Gates, Phase, and the Bloch Sphere	4
Gates	4
Phase	7
Bloch Sphere	8
Quantum Circuits, Measurement, and Entanglement	9
Quantum Circuits	9
Measurement	10
Bell States and Entanglement	10
CHSH Experiment	12
Bell's Theorem	12
Expectation Values	12
The CHSH value, S	13
Tsirelson's bound	14
Aer Simulator Code	16
Results	18
Execution on IBM Quantum Hardware	19
Results	20
Conclusion	20
Bibliography	22

Introduction

Proving Bell's theorem with the CHSH experiment is a foundational result in Quantum Information theory. To begin, paper will first begin by introducing the required background knowledge adopted from the formalism of quantum information as described by Neilsen and Chuang [1], and then it will go through the quantum code to experimentally reproduce the results of the CHSH experiment.

Qubits, or "quantum bits," are quantum representations of information. Physically, a qubit can be made from any quantum particle that has two distinct states. For example a photon of light being polarized either horizontally or vertically or an electron being spin up or spin down. Physically, qubits are usually either spin up/spin down instead of being in a "zero/one" state.

$$|0\rangle = |\uparrow\rangle \quad |1\rangle = |\downarrow\rangle$$

A singular qubit can be either zero or one (or both!), and represented mathematically by

$$|\psi\rangle = |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |\psi\rangle = |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Similarly to classical bits, qubits can be either 0 or 1, however, they can also leverage superposition and be in both states simultaneously

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

Where the square of $\frac{1}{\sqrt{2}}$ is the probability of each state, so it's 50% $|0\rangle$ and 50% $|1\rangle$ The qubit will stay in this superposition until interaction with the environment or measurement, then it will collapse into either $|\psi\rangle = |0\rangle$ or $|\psi\rangle = |1\rangle$

A more general way to represent qubits is like this, where $|\alpha|^2$ and $|\beta|^2$ are the probabilities of being in the respective states.

$$|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \alpha |0\rangle + \beta |1\rangle$$

So this is how a single qubit can be represented, how do we consider a multi-qubit system? To add new qubits to a system, they need to be tensored together.

A 5-qubit system can be represented by

$$|\psi\rangle = |01011\rangle$$

This 5-qubit system is quite messy to represent with matrices and can be done with the tensor product.

$$|01011\rangle = |0\rangle \otimes |1\rangle \otimes |0\rangle \otimes |1\rangle \otimes |1\rangle$$

What does the tensor product actually do?

Mathematically:

$$(\alpha|0\rangle + \beta|1\rangle) \otimes (\gamma|0\rangle + \delta|1\rangle) = \alpha\gamma|00\rangle + \beta\gamma|10\rangle + \alpha\delta|01\rangle + \beta\delta|11\rangle$$

By tensoring these qubits together, we have collected all possible outcomes of the qubits and their probabilities, where now $|\alpha\gamma|^2$ is the probability for measuring $|00\rangle$

Gates, Phase, and the Bloch Sphere

Gates

In classical computing, there are logic gates, which operate on binary inputs. For example, the AND gate and the OR gate for two binary inputs.

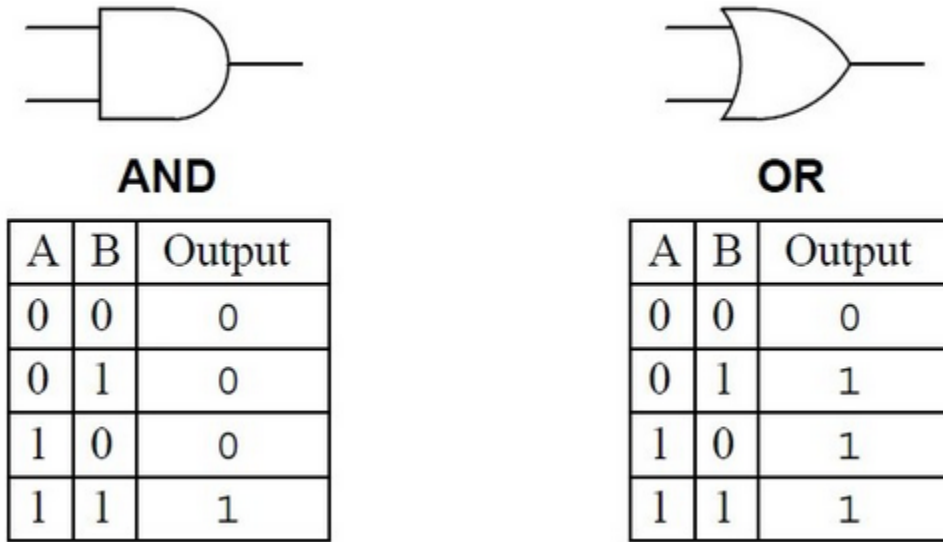


Figure 1: AND and OR Logic Gates
[4]

This leads into the idea of quantum logic gates, or simply qubit gates.[1]

Before going into qubit gates, it is of note that it is also of note that each of these gates is Hermitian.

To define Hermitian matrices, we must first define the unitary matrix, U^\dagger . To create a unitary matrix, you must take the transpose of the matrix and then apply the complex conjugate, or vice versa.

$$(U^T)^* = (U^*)^T = U^\dagger$$

if $A = A^\dagger$, then in addition to the matrix being unitary, it is also Hermitian. I.E. All Hermitian matrices are unitary, but not all unitary matrices are Hermitian.

It is of note that unitary and Hermitian matrices are their own inverse For a unitary matrix:

$$A^\dagger A = I \quad A \neq A^\dagger$$

For a Hermitian matrix:

$$A^\dagger A = I \quad A = A^\dagger$$

Now to explain qubit gates, I will show various examples of how they change qubits. However, for the CHSH experiment, we will focus on the H, CNOT, and RY gates.

The Pauli-X Gate:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

The X gate is one of the simplest gates, and it can be used to change a $|0\rangle$ to a $|1\rangle$ or a $|1\rangle$ to a $|0\rangle$. Geometrically, we will later see that the X gate flips the qubit π radians around x-axis on the Bloch Sphere

$$X|0\rangle = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \cdot 1 + 1 \cdot 0 \\ 1 \cdot 1 + 0 \cdot 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = |1\rangle.$$

$$X|1\rangle = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \cdot 0 + 1 \cdot 1 \\ 1 \cdot 0 + 0 \cdot 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle.$$

The Pauli-Y Gate:

$$Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

The Y gate does the same as the X gate geometrically, but along the y-axis instead of the x-axis.

$$Y|0\rangle = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \cdot 1 + (-i) \cdot 0 \\ i \cdot 1 + 0 \cdot 0 \end{bmatrix} = \begin{bmatrix} 0 \\ i \end{bmatrix} = i|1\rangle.$$

$$Y|1\rangle = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \cdot 0 + (-i) \cdot 1 \\ i \cdot 0 + 0 \cdot 1 \end{bmatrix} = \begin{bmatrix} -i \\ 0 \end{bmatrix} = -i|0\rangle.$$

The Pauli-Z Gate:

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Finally, the Z gate rotates the qubit π radians about the Z-axis of the Bloch Sphere

$$Z|0\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \cdot 1 + 0 \cdot 0 \\ 0 \cdot 1 + (-1) \cdot 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle.$$

$$Z|1\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \cdot 0 + 0 \cdot 1 \\ 0 \cdot 0 + (-1) \cdot 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} = -|1\rangle.$$

The three Pauli matrices are the simplest of the Quantum Gates, and while they are not explicitly used for the CHSH experiment, they aid in understanding how a qubit can be transforms.

Before we introduce the idea of the Bloch sphere, let us finish introducing the gates we will need for the CHSH experiment.

The Hadamard Gate:

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

The Hadamard Gate is central to entanglement, because it creates an equal superposition state from a basis state

$$H|0\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \cdot 1 + 1 \cdot 0 \\ 1 \cdot 1 + (-1) \cdot 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle).$$

$$H|1\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \cdot 0 + 1 \cdot 1 \\ 1 \cdot 0 + (-1) \cdot 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle).$$

We originally start with $|0\rangle$, but after applying the Hadamard gate, we have an equal superposition where the qubit is 50% in $|0\rangle$ and 50% in $|1\rangle$. For the CHSH experiment, we will not be using $H|1\rangle$, but it is instructive.

The Controlled X Gate:

$$CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

This gate and the Hadamard gate are central to the entanglement between two qubits, as we will see later. Consider the CNOT gate to be an X gate on a target qubit, but it only activates if the second qubit is a 1. Looking forward to entanglement, what happens if that second (control) qubit is in an equal superposition of $|0\rangle$ and $|1\rangle$?

Rotation operator gate about y-axis

$$RY(\theta) = \begin{bmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{bmatrix}$$

You might recall that we have the Y gate, which is a rotation of π radians about the y-axis.

$$Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

So what happens if we plug in π to the RY gate?

$$RY(\pi) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

This looks similar to the Y gate, but it's missing the i term! Why is that? This is due to the idea of global and relative phase.

Phase

First let us take a base qubit, $|0\rangle$, and then let us apply a Hadamard gate to split it into a superposition:

$$|\psi\rangle = H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

Now that we have an equal superposition, let's apply a Z gate and see what happens.

$$Z\left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$Z|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

This is very similar to just applying a Hadamard gate to $|0\rangle$, but it has changed $|1\rangle$ to $-|1\rangle$. Recall Euler's identity and note that we can rewrite this like this.

$$Z|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\pi}|1\rangle)$$

This is the idea of relative phase. Recall that the Z gate rotates the qubit around the z-axis by π , so what if we want to rotate it by some angle ϕ ?

A general form:

$$|\psi\rangle = \alpha|0\rangle + e^{i\phi}\beta|1\rangle$$

But why is the $e^{i\phi}$ only on $|1\rangle$? Let's try putting it on both the 0 and 1 kets.

$$|\psi\rangle = e^{i\phi}\alpha|0\rangle + e^{i\phi}\beta|1\rangle = e^{i\phi}(\alpha|0\rangle + \beta|1\rangle)$$

The global phase, $e^{i\phi}$, has no observable consequences and doesn't change the probabilities of getting a measurement result and is considered to be physically irrelevant, so it can be left off.[1]

$$e^{i\phi}(\alpha |0\rangle + \beta |1\rangle) = \alpha |0\rangle + \beta |1\rangle$$

Note that if both phases are different, it just reduces to the relative phase.

$$|\psi\rangle = e^{i\theta}\alpha |0\rangle + e^{i\phi}\beta |1\rangle = \alpha |0\rangle + e^{i(\phi-\theta)}\beta |1\rangle$$

Also note that the probabilities do not change since $|e^{i\phi}|^2 = 1$

The phase factor on the 1 ket instead of the 0 ket by standard convention. It can be rewritten to be on the 0 ket instead if needed

$$|\psi\rangle = \alpha |0\rangle + e^{i\phi}\beta |1\rangle = e^{i2\pi}\alpha |0\rangle + e^{i\phi}\beta |1\rangle = e^{i(2\pi-\phi)}\alpha |0\rangle + e^{i\phi}\beta |1\rangle$$

Bloch Sphere

We need more ways to represent qubits other than Dirac notation. It would be beneficial to have a geometric representation, so we can better visualize changes to a qubit. Thus, the physicist, Felix Bloch, introduced a sphere on which qubits can be represented as a point on the surface.

Higher vertically means that the qubit is more likely to be $|0\rangle$ and vice versa for $|1\rangle$

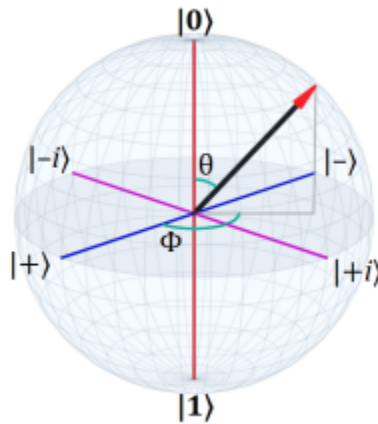


Figure 2: The Bloch Sphere
[5]

But what are these different states along the equator?

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \quad |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

$$|+i\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \quad |-i\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)$$

There are in between $|0\rangle$ and $|1\rangle$, so that would imply that the states are equally probable. The only difference between them is their phase.

$$\begin{aligned} |+\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + e^{i2\pi}|1\rangle) & |-\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + e^{i\pi}|1\rangle) \\ |+i\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + e^{i\frac{\pi}{2}}|1\rangle) & |-i\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + e^{i\frac{3\pi}{2}}|1\rangle) \end{aligned}$$

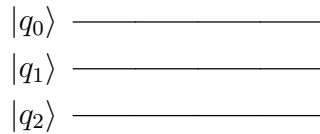
Quantum Circuits, Measurement, and Entanglement

Quantum Circuits

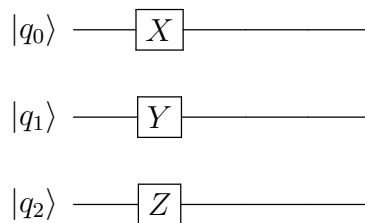
So far, we have discussed gates, phase, and the Bloch sphere, but each of these discussions has only involved a single qubit. How can we apply a gate to a single qubit in a multi-qubit system, and how can we represent this system graphically?

This leads into the idea of the quantum circuit — a graph with lines as "wires" which read left to right as different points in time with each qubit in a system having a separate wire.

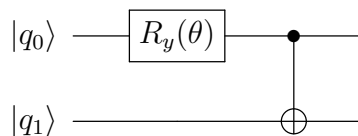
To draw the base diagram, first consider a 3 qubit system without any gates applied



Now what if we want to apply an X gate to qubit 0, a Y gate to qubit 1, and a Z gate to qubit 2?



Quantum circuits are intuitive and very useful for visualizing a qubit system. Let's explore a few more examples.

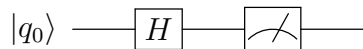


This system has an R_Y gate applied on qubit 0, and a CNOT gate controlled by qubit 0 and targeted to qubit 1.

Measurement

We briefly mentioned measurement previously when discussing the probabilities, but now we can see how it is implemented in a quantum circuit.

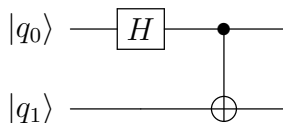
We have created a quantum circuit with one qubit with a Hadamard gate applied, and then after it is applied, the qubit is measured.



Recall that the Hadamard gate takes a basis qubit, and splits it into equal probabilities. Thus, if it is measured, the qubit will have a 50/50 chance to collapse into either state.

Bell States and Entanglement

As we saw from the Hadamard gate, it takes a basis state, and splits them into a superposition. So what if we take two qubits, apply a Hadamard gate, and then apply a CNOT gate?



This two qubit circuit is the classical example of entanglement and is referred to as a Bell state. More specifically, this is a "maximally entangled two-qubit Bell state."

This Bell state can be mathematically represented by

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

There are 4 different maximally entangled Bell states, but $|\Phi^+\rangle$ is the most common and will be used in the CHSH experiment.

$$\begin{aligned} |\Phi^+\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) & |\Phi^-\rangle &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \\ |\Psi^+\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) & |\Psi^-\rangle &= \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \end{aligned}$$

But what does it mean for two qubits to be entangled?

First, consider a two qubit system

$$|\psi_0\rangle = \alpha |0\rangle + \beta |1\rangle$$

$$|\psi_1\rangle = \gamma |0\rangle + \delta |1\rangle$$

And as we saw from the previous example earlier in the introduction,

$$|\psi_0\rangle \otimes |\psi_1\rangle = \alpha\gamma |00\rangle + \beta\gamma |10\rangle + \alpha\delta |01\rangle + \beta\delta |11\rangle$$

now compare this to the entangled Bell state, $|\Phi^+\rangle$

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

Let's try to match the coefficients of the original two qubit system to the Bell state. I.E. we need to find coefficients such that

$$\alpha\gamma = 0 \quad \alpha\delta = \frac{1}{\sqrt{2}} \quad \beta\gamma = \frac{1}{\sqrt{2}} \quad \beta\delta = 0$$

It is impossible to match these coefficients from the sub system to the system of a Bell state. This implies non-separability of the qubits in the Bell state. This is the core idea of entanglement.

CHSH Experiment

Bell's Theorem

Bell's theorem states that no local hidden variable theory can reproduce all of the predictions of quantum mechanics. [2]

A local hidden variable theory assumes that the universe operates with "hidden variables" which cannot be measured and there is no way for us to obtain information about them, with the additional principle of locality stating that objects must be influence only by its immediate surroundings.

The principle of locality evolved from the theories of classical physics, and is central to special and general relativity. One of Einstein's primary postulates is one of locality — No information or causal influence can travel faster than the speed of light.

Einstein personally advocated for local realism, which states that there is no instantaneous action at a distance and that physical properties exist independently of observation. However, the CHSH experiment is evidence against this idea of local realism.

The CHSH inequality provides a practical way to test Bell's theorem by bounding correlations between two spatially separated systems under the assumptions of a local hidden variable theory. [3]

Expectation Values

We will soon see that the CHSH parameter is composed of expectation values and is written as

$$S = E(A_0B_0) - E(A_0B_1) + E(A_1B_0) + E(A_1B_1)$$

How can we calculate these individual expectation values experimentally?

$$E(A_iB_j) = \langle \psi | A_i \otimes B_j | \psi \rangle$$

Recall that the tensor product extrapolates each possible permutation for A_i and B_j , where the solution is the product of Alice and Bob's outcomes for settings A_i and B_j . We write this in terms of correlations as done by Xiao Feng. [6]

$$E(A_iB_j) = \frac{1}{N} \sum_{k=1}^N a_k b_k$$

Such that $a_k, b_k \in \{\pm 1\}$ are the outcomes for the total number of trials.

If both of the outcomes are $1\pm$, then the outcomes agree, and are positively correlated. However, if one correlation is $+1$ and the other is -1 , then the outcomes disagree, and are negatively correlated. Thus we can write the expectation values as such

$$E(A_i, B_j) = \frac{N_{++} - N_{+-} - N_{-+} + N_{--}}{N_{total}}$$

The CHSH value, S

The CHSH parameter, S, is given by

$$S = E(A_0B_0) - E(A_0B_1) + E(A_1B_0) + E(A_1B_1)$$

Where $E(A_iB_j)$ represents the expectation values of the product of the measurement outcomes for two qubits, of which we will name Alice and Bob.

This expression arise by strategic choice. It is designed to maximize violation of the CHSH inequality and to simplify experimentation from Bell's original inequality, which was ineffective against experimental noise. [3]

To proceed with the derivation of the inequality, assume that outcomes are predetermined by a hidden variable, λ . Locality ensures that outcomes only depend on local settings, and realism assumes that outcomes exist independently outside of measurement.

for a fixed value of λ , with binary outcomes

$$A_0(\lambda), A_1(\lambda), B_0(\lambda), B_1(\lambda) \in \{\pm 1\}$$

Then the CHSH value becomes

$$S(\lambda) = A_0(\lambda)B_0(\lambda) - A_0(\lambda)B_1(\lambda) + A_1(\lambda)B_0(\lambda) + A_1(\lambda)B_1(\lambda)$$

$$S(\lambda) = A_0(\lambda)[B_0(\lambda) - B_1(\lambda)] + A_1(\lambda)[B_0(\lambda) + B_1(\lambda)]$$

All possible permutations				
B_0	B_1	$B_0 - B_1$	$B_0 + B_1$	$S(\lambda)$
+1	+1	0	+2	$+2A_1$
+1	-1	+2	0	$+2A_0$
-1	+1	-2	0	$-2A_0$
-1	-1	0	-2	$-2A_1$

Thus, in all cases and for all hidden variables, $|S(\lambda)| \leq 2$ holds. Due to the absolute value, the bound $S \leq -2$ is also valid, but $S \geq 2$ is more commonly used

This inequality is experimentally testable and can be violated with a Bell state, but how can we maximize this violation to avoid experimental noise?

Tsirelson's bound

If we find that the CHSH value is greater than 2, we have proven the contrapositive of the assumptions used in deriving the inequality, but what is the maximum possible violation?

This maximum value is known as the Tsirelson bound, named after the mathematician Boris S. Tsirelson. In his 1980 paper, Tsirelson derived the the upper bound for the CHSH and highlighted the differences in Bell's 1964 framework of local hidden variable theories. [7] [2]

Using a slightly different notation than before, we will use the operator norm of the CHSH operator, T . This is not the same as S , but it is directly related.

$$T = A_0 \otimes B_0 - A_0 \otimes B_1 + A_1 \otimes B_0 + A_1 \otimes B_1$$

As an intermediate step

$$T = A_0 \otimes (B_0 - B_1) + A_1 \otimes (B_0 + B_1)$$

$$T^2 = (A_0 \otimes (B_0 - B_1) + A_1 \otimes (B_0 + B_1))^2$$

$$T^2 = (A_0 \otimes (B_0 - B_1))^2 + (A_1 \otimes (B_0 + B_1))^2 + A_0 A_1 \otimes ((B_0 - B_1)(B_0 + B_1)) + A_1 A_0 \otimes ((B_0 + B_1)(B_0 - B_1))$$

First consider the square terms and recall that $A_i^2 = B_j^2 = I$

$$(B_0 - B_1)^2 = 2I - (B_0 B_1 + B_1 B_0)$$

$$(B_0 + B_1)^2 = 2I + (B_0 B_1 + B_1 B_0)$$

Thus,

$$(A_0 \otimes (B_0 - B_1))^2 = A_0^2 \otimes (B_0 - B_1)^2 = I \otimes (2I - (B_0 B_1 + B_1 B_0))$$

$$(A_1 \otimes (B_0 + B_1))^2 = A_1^2 \otimes (B_0 + B_1)^2 = I \otimes (2I + (B_0 B_1 + B_1 B_0))$$

Adding them together gives,

$$I \otimes [2I - (B_0 B_1 + B_1 B_0) + 2I + (B_0 B_1 + B_1 B_0)] = I \otimes 4I = 4I$$

Now consider the cross terms, and remember that terms are not necessarily commutative.

$$(B_0 - B_1)(B_0 + B_1) = B_0B_1 - B_1B_0 = [B_0, B_1]$$

and similarly,

$$(B_0 + B_1)(B_0 - B_1) = B_1B_0 - B_0B_1 = [B_1, B_0] = -[B_0, B_1]$$

Thus the cross terms become

$$\begin{aligned} A_0A_1 \otimes [B_0, B_1] + A_1A_0 \otimes (-[B_0, B_1]) &= A_0A_1 \otimes [B_0, B_1] - A_1A_0 \otimes [B_0, B_1] \\ &= [A_0A_1 - A_1A_0] \otimes [B_0, B_1] = [A_0, A_1] \otimes [B_0, B_1] \end{aligned}$$

Combining all of the terms,

$$T^2 = 4I + [A_0, A_1] \otimes [B_0, B_1]$$

Notice that if the commutators are zero, then we see the classical bound, $T = 2I$

To evaluate these commutators requires the Simultaneous diagonalization theorem [1], which leads into the commutation relations for the Pauli matrices as stated by Nielsen and Chuang. This results will not be fully shown here.

If we assume the commutation relations for Z and X, e.g. $[Z, X] = 2iY$, then we get

$$[A_0, A_1] = 2iY \quad [B_0, B_1] = 2iY$$

$$T^2 = 4I + [A_0, A_1] \otimes [B_0, B_1] = 4I + 2iY \otimes 2iY = 4I - 4(Y \otimes Y)$$

Since $Y^2 = I$, the eigenvalues are ± 1 , thus it follows that

$$T^2 = 4I \pm 4I$$

ignoring $T^2 = 0$,

$$||T^2|| \leq 8 \implies |S| \leq 2\sqrt{2}$$

With quantum mechanics, we find the maximum value for the CHSH value is $2\sqrt{2}$. This is the theoretical limit, even with higher-dimensional entanglement, no quantum state or measurement strategy will exceed this. Physically it demonstrates that quantum correlations are stronger than classical, however still bounded by $2\sqrt{2}$

Aer Simulator Code

Qiskit was created by IBM researchers to execute programs on quantum computers. To recreate the CHSH experiment, I first used the Aer simulator module with Qiskit. [8]

We will skip over all of the Qiskit and Aer simulator setup and only include pseudo-code relating to the CHSH experiment. The complete code can be found on <https://github.com/acroscopic/CHSH/>

The first step was to create a bell state between two qubits, Alice, $q[0]$ and Bob, $q[1]$. 2 classical bits are also created for storing the results of measurement.

```
1 q = QuantumRegister(2, 'q')
2 c = ClassicalRegister(2, 'c')
3 def bell():
4     qc = QuantumCircuit(q, c)
5     qc.h(q[0])
6     qc.cx(q[0], q[1])
7     return qc
```

The next step is to measure the qubits and store them in the classical bits. However, before measurement, the RY gate is applied to the measured qubit. The angle at which the RY gate is applied will be discussed shortly.

```
1 def measure(qc, angle, qubit, cbit):
2     qc.ry(-angle, qubit)
3     qc.measure(qubit, cbit)
4     return qc
```

Now we need to figure out a way to calculate the expectation values. Recall that the individual expectation values are calculated by

$$E(A_i, B_j) = \frac{N_{++} - N_{+-} - N_{-+} + N_{--}}{N_{total}}$$

```
1 def expectation(counts):
2     total = sum(counts.values())
3     count_00 = counts.get('00', 0)
4     count_11 = counts.get('11', 0)
5     count_01 = counts.get('10', 0)
6     count_10 = counts.get('01', 0)
7     E = (count_00 + count_11 - count_01 - count_10) / total
8     return E
```

You might notice that $count_{01} = counts.get('10')$ and $count_{10} = counts.get('01')$. This is due to Qiskit inherently ordering bitstrings from right to left instead of left to right.

Now that the setup is done, it is time to implement the CHSH experiment directly

```
1 def CHSH():
2     S = 0 # initializes the CHSH value
3
4     # These angles provide the maximum violation of the CHSH
5     # inequality, known as Tsirelson's bound
6     angles = {
7         'A0': np.pi / 2,
8         'A1': 0,
9         'B0': np.pi / 4,
10        'B1': -np.pi / 4
11    }
12    measurement_settings = [
13        ('A0', 'B0'),
14        ('A0', 'B1'),
15        ('A1', 'B0'),
16        ('A1', 'B1'),
17    ]
18    for alice, bob in measurement_settings:
19        qc = bell()
20
21        # Apply Alice's A0 and A1 measurements on the 0th qubit and
22        # store it in the 0th bit for each measurement setting
23        qc = measure(qc, angles[alice], 0, 0)
24
25        # Apply Bob's B0 and B1 measurements on the 1st qubit and
26        # store it in the 1st bit for each measurement setting
27        qc = measure(qc, angles[bob], 1, 1)
28
29        # 10 million iterations for the expectation values
30        transpiled_qc = transpile(qc, backend)
31        job = backend.run(transpiled_qc, shots=10000000)
32        result = job.result()
33        counts = result.get_counts()
34
35        # Calculate the expectation value for each measurement setting
36        E = expectation(counts)
37
38        #  $S = E(A_0 B_0) - E(A_0 B_1) + E(A_1 B_0) + E(A_1 B_1)$ 
39        if alice == 'A0' and bob == 'B1':
40            S -= E
41        else:
42            S += E
43
44    return S
```

Results

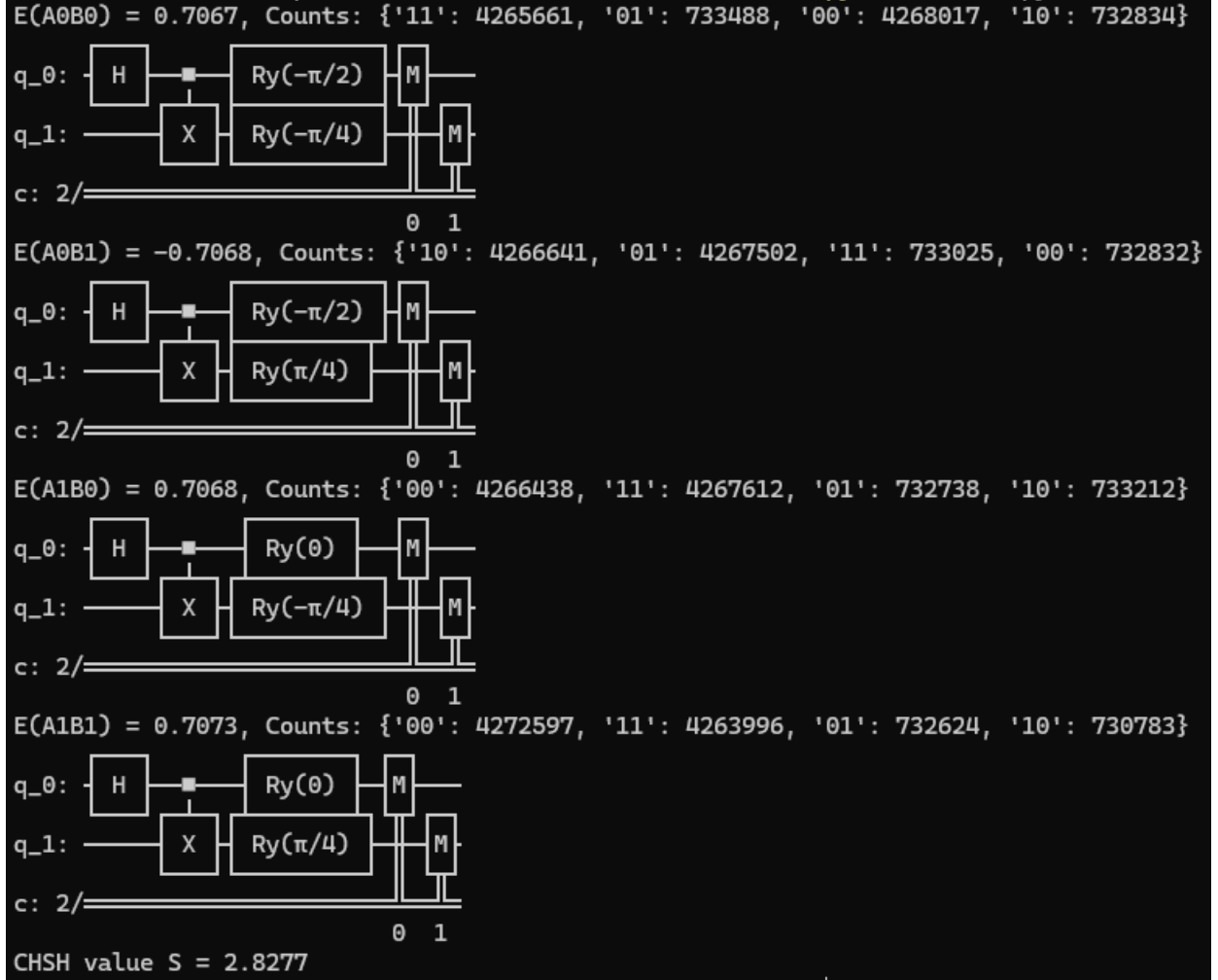


Figure 3: Results from running Aer.py

After running the code, with a total of 10 million iterations per expectation value, we get a final CHSH value of 2.8277, which is just under the maximum bound of $2\sqrt{2} \approx 2.828...$

The Aer simulator has provided a near-noiseless environment to run the CHSH experiment, however running the experiment on real hardware does not have such a luxury. To run on IBM hardware, the code must be modified so that the violation can still be clearly observed even with experimental noise

Execution on IBM Quantum Hardware

Previously, the section of code was ran using the Aer simulator and did not interface with any quantum hardware. It is highly relevant to empirically replicate the results of the CHSH experiment using a real quantum computer [9]. The setup is different, but the theory behind it is not.

Since there is a lot of setup for IBM hardware, more Qiskit specific aspects of the code will be shown. However, this is still pseudo-code, and the full version can be found on <https://github.com/acroscopic/CHSH/>

First, we need to set up the type of backend we want to use. We will filter for hardware with 127 qubits, which is the modern "Eagle" processor-based systems via IBM Quantum. [9]

```
1 try:
2     backend = service.least_busy(simulator=False, operational=True,
3                                   min_num_qubits=127)
4 except:
5     print("No backend available. Using simulator instead.")
6     backend = service.get_backend("ibmq_qasm_simulator")
```

For this version of the code instead of running millions of calculations to find the expectation values, the angle for the RY gate will be parametrized between equally spaced values over a full 0 to 2π rotation. It can do it this way because of the difference in calculating the expectation values.

```
1 theta = Parameter("$\\theta$")
2
3 qc = QuantumCircuit(2)
4
5 qc.h(0)
6 qc.cx(0, 1)
7
8 qc.ry(theta, 0)
9
10 number_of_phases = 20
11 phases = np.linspace(0, 2 * np.pi, number_of_phases)
```

One of the biggest difference between the previous Aer simulator code and this version, is that expectation values are written in terms of the observables as Pauli matrices, the X and Z gates. Where X and Z are tensored together. This is just an alternative method to calculate the same expectation values as before.

```
1 # S = <ZZ> - <ZX> + <XZ> + <XX>
2 observable = SparsePauliOp.from_list([
3     ("ZZ", 1), # E(A_0 B_0)
4     ("ZX", -1), # - E(A_0 B_1)
5     ("XZ", 1), # E(A_1 B_0)
6     ("XX", 1), # E(A_1 B_1)
```

7])

Since we are running this code on hardware, we need to modify our data to fit the specific hardware

```
1 target = backend.target
2 pm = generate_preset_pass_manager(target=target, optimization_level
  =3)
3 qc_isa = pm.run(qc)
4 isa_observable = observable.apply_layout(layout=qc_isa.layout)
5 estimator = Estimator(mode=backend)
6 pub = (
7     qc_isa, # remapped ISA circuit
8     [[isa_observable]], # remapped observables
9     individual_phases # Parameter values to test
10 )
```

And then finally, we submit the job to queue and wait to receive the results.

```
1 job = estimator.run(pubs=[pub])
2 result = job.result()
3
4 S_values = result[0].data.evs[0] # Expectation values for the
  observable
5 print(S_values)
6
7 violation = np.any(np.abs(S_values) > 2)
8 print(f"CHSH violation detected: {violation}")
9
10 print(f"Job ID: {job.job_id()}") # Unique identifier for job
11 print(f"Job status: {job.status()}") # Should be 'DONE' if
  successful
```

Results

After running the code on IBM's quantum computer in Ukraine, Kyiv, we get multiple CHSH values corresponding to the different angles of the RY gates, many of which violate the classical bound $|S| \leq 2$.

Conclusion

By executing the CHSH experiment on both the Aer simulator and IBM hardware, we observed the CHSH parameter, $S > 2$, violating the classical bound of $|S| \leq 2$. This result is evidence of the violation of the CHSH inequality.

By violating this inequality, we have empirically validated Bell's theorem, demonstrating that non-local realism theories can fully describe quantum correlations.

```
Using backend: ibm_kyiv
[ 1.92082573  1.17968406  0.26785827 -0.67216313 -1.52004531 -2.20630625
 -2.66448486 -2.78884763 -2.70778526 -2.23802631 -1.56787715 -0.75121153
  0.20743911  1.0905658  1.85939958  2.43086412  2.75158915  2.78280571
  2.47215054  1.87903581]
CHSH violation detected: True
Job ID: czrx8paqnmvg008v0jvg
Job status: DONE
```

Figure 4: Results from running CHSH.py

This violation implies that quantum mechanics exhibits non-locality. While this does not imply instantaneous communication, it confirms that quantum systems cannot be described by local deterministic methods

These findings underscore the necessity of embracing quantum theory and its inherently non-classical framework to be able to describe the fundamental aspects of nature.

Bibliography

References

- [1] Nielsen, M. A., & Chuang, I. L. *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.
- [2] Bell, J. S. "On the Einstein Podolsky Rosen Paradox." *Physics Physique Fizika*, vol. 1, pp. 195–200, 1964. https://cds.cern.ch/record/111654/files/vol1p195-200_001.pdf
- [3] Clauser, J. F., Horne, M. A., Shimony, A., & Holt, R. A. "Proposed Experiment to Test Local Hidden-Variable Theories." *Physical Review Letters*, vol. 23, no. 15, pp. 880–884, 1969. <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.23.880>
- [4] Toheeb Olaide/Class Notes, "logic-gates-symbol-csc-classnotesng.jpg," classnotes.ng, Date, <https://classnotes.ng/lesson/logic-gate>
- [5] Wyrđ Smythe/Logos Con Carne, "bloch-hdr.png," logosconcarne.com, March 15, 2021, <https://logosconcarne.com/2021/03/15/qm-101-bloch-sphere/>
- [6] Xiao Feng, "Proof of CHSH Inequality and Tsirelson's Bound Using Same Method," 2024, <https://math.ucsd.edu/sites/math.ucsd.edu/files/XiaoFeng.pdf>
- [7] Tsirelson, B. S. "Quantum Generalizations of Bell's Inequality." *Letters in Mathematical Physics*, vol. 4, pp. 93–100, 1980.
- [8] IBM, *Qiskit: An Open-Source Framework for Quantum Computing*, 2025, 2.0, <https://qiskit.org>
- [9] IBM, *IBM Quantum Computing*, 2025, <https://quantum-computing.ibm.com>