

Violating the CHSH Inequality: Empirical Evidence of Quantum Nonlocality

Damien Koon

September 2025

Abstract

The Clauser–Horne–Shimony–Holt (CHSH) inequality provides an experimentally testable framework for Bell’s theorem, leading to John Clauser’s pioneering Bell tests which were recognized by the 2022 Nobel Prize in Physics. This paper acts as a pedagogical exposition aimed at undergraduate students and simultaneously replicates and extends CHSH experiments to explore quantum nonlocality across three computational settings: (1) an ideal (noiseless) simulation, (2) a noisy simulation (to emulate decoherence), and (3) execution on IBM quantum hardware. In each case, entangled Bell states are prepared and measured with optimized settings, leading to the demonstration of violations of the classical bound, $|S| \leq 2$. Our results empirically validate Bell’s theorem and demonstrate that quantum correlations remain robust even under realistic noise. This work serves as a practical guide for educators, students, and researchers who wish to reproduce this foundational experiment in quantum information in their own environments.

Contents

1	Introduction	3
2	Qubits	3
2.1	Qubit Manipulation	5
2.1.1	Matrix Operations	5
2.1.2	Qubit Gates	5
2.1.3	The Pauli Matrices	6
2.1.4	Essential Quantum Gates	7
2.1.5	Phase	9
2.1.6	Bloch Sphere	10
2.2	Quantum Circuits and Entanglement	11
2.2.1	Quantum Circuits	11
2.2.2	Measurement	12
2.2.3	Bell States and Entanglement	13
3	CHSH Inequality	15
3.1	Bell's Theorem	15
3.2	The CHSH parameter, S	15
3.3	Expectation Values	16
3.4	The Classical Bound, $S \leq 2$	18
3.5	Tsirelson's bound	19
4	CHSH Experiment	24
4.1	Aer Simulator and IBM Hardware	24
4.2	Statevector Simulator	24
4.2.1	Results	26
4.3	Qasm Simulator	26
4.3.1	Results	29
4.4	Execution on IBM Quantum Hardware	30
4.4.1	Results	31
5	Conclusion	32
6	Acknowledgements	33
7	Bibliography	34

1 Introduction

We begin by introducing the required background knowledge adopted from the formalism of quantum information as described by Nielsen and Chuang [1]. We then explain and derive the CHSH inequality, including Tsirelson's bound. Finally, we will then go through the code to experimentally reproduce the results of the CHSH experiment on IBM quantum hardware[6]. By experimentally violating the CHSH inequality, we provide empirical support for Bell's theorem from observing violations inconsistent with local hidden-variable models. The CHSH experiment delivers unambiguous evidence of nonlocal quantum entanglement, ruling out local realism, and laying foundational groundwork for quantum information theory.

2 Qubits

Quantum bits, or "qubits," are quantum representations of information. Physically, a qubit can be made from any quantum system with 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 either spin up or spin down with respect to an external magnetic field instead of being in a "zero" or "one" state.

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

A singular qubit can be either zero or one (or both!) and they are 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.

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

In this case, $\frac{1}{\sqrt{2}}$ is the "probability amplitude" of each state. To find the individual probability, we square the modulus of the values, $|\frac{1}{\sqrt{2}}|^2 = 0.5$.¹ Therefore, the qubit is 50% in the $|0\rangle$ state and 50% in the $|1\rangle$ state. We can write the probability of measuring these states as follows

$$P(|0\rangle) = |\frac{1}{\sqrt{2}}|^2 = 0.5 \quad P(|1\rangle) = |\frac{1}{\sqrt{2}}|^2 = 0.5$$

¹The modulus might seem superfluous in this case, but is standard practice when working with complex numbers, which we will see later.

The qubit will stay in this superposition until it interacts with the environment or is measured, then it will collapse into one of the states, depending on the probability. Either $|\psi\rangle = |0\rangle$ or $|\psi\rangle = |1\rangle$ could be observed.

A more general way to represent qubits is as follows, where $|\alpha|^2$ and $|\beta|^2$ are the probabilities of being in the $|0\rangle$ and $|1\rangle$ states, respectively.

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

If this is how a single qubit can be represented, how do we represent a multi-qubit system?

For now, let's continue using the same matrix notation, but with a 2-qubit system. We previously represented $|0\rangle$ and $|1\rangle$ by

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

We will now represent a 2-qubit system by

$$\begin{aligned} |00\rangle &= \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} & |01\rangle &= \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \\ |10\rangle &= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} & |11\rangle &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \end{aligned}$$

Writing out a five-qubit state as a 32×1 column would be cumbersome, so instead we build a 32-dimensional space by "tensoring" five two-dimensional spaces.

This tensor notation is typically reduced by putting it all inside of a ket. Note that we read this notation from right to left. Thus, the first qubit in this system is 1, and not a 0 as you might expect.

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

The tensor product takes the possible states and expands it to show every possible permutation of the states, it combines separate quantum systems into one joint system, keeping track of all possible combinations of their states.

For the rest of the paper, we will be focusing on two-qubit systems. This is how we can generally represent two qubits tensored together:

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

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 $|00\rangle$ and so on.

2.1 Qubit Manipulation

2.1.1 Matrix Operations

Before we proceed with manipulating qubits, we must first define a matrix operation, the Hermitian conjugate, U^\dagger , pronounced "U dagger." To perform this operation, you must take the transpose of the matrix and then apply the complex conjugate, or vice versa. Basic examples will be shown in the next section.

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

For a unitary matrix, a matrix times its Hermitian conjugate, will equal the identity.

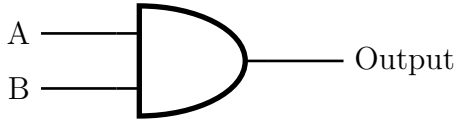
$$U^\dagger U = I$$

For a Hermitian operator, we have $U = U^\dagger$ and for a unitary operator, we have $U^\dagger U = I$.

Additionally, an operator is both unitary and Hermitian if and only if $U^2 = I$.

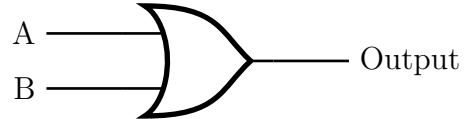
2.1.2 Qubit Gates

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. They are frequently used in computer engineering to manipulate classical information.



AND

A	B	Output
0	0	0
0	1	0
1	0	0
1	1	1



OR

A	B	Output
0	0	0
0	1	1
1	0	1
1	1	1

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

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.

Recall that all quantum gates are unitary operators, but only those satisfying $U^2 = I$ happen to be Hermitian. For example, the Pauli gates and the Hadamard gate are also Hermitian.

2.1.3 The Pauli Matrices

We will start with the Pauli matrices, X, Y, Z (Also referred to as $\sigma_x, \sigma_y, \sigma_z$). Each of these gates apply to a single qubit.

The Pauli-X Gate:

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

The X gate is both unitary and Hermitian, but what does it do?

$$XX^\dagger = I \quad X = X^\dagger$$

The X gate can be used to change a $|0\rangle$ to a $|1\rangle$ or a $|1\rangle$ to a $|0\rangle$.

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

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

The Pauli-Y Gate:

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

$$YY^\dagger = I \quad Y = Y^\dagger$$

The Y gate is noticeably different with the addition of the imaginary number, i. We will later see how this can be visualized with the Bloch sphere. For now, we will just see what happens when we apply it to $|0\rangle$ and $|1\rangle$

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

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

The Pauli-Z Gate:

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

$$ZZ^\dagger = I \quad Z = Z^\dagger$$

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

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

The three Pauli matrices are the simplest of the quantum gates and they aid in understanding how a qubit can be transformed. We will later see how these qubits appear geometrically on the Bloch sphere.

2.1.4 Essential Quantum Gates

We will next introduce the rest of the gates that will be needed 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 $|0\rangle$ and $|1\rangle$ 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:

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

So far, we have only considered single qubit gates, but the CNOT is a two qubit gate. Consider the CNOT gate to be an X gate on a target qubit, but it only activates if the second, "control" qubit is a 1.

$$CNOT \cdot |00\rangle = CNOT \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = |00\rangle$$

$$CNOT \cdot |01\rangle = CNOT \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = |01\rangle$$

$$CNOT \cdot |10\rangle = CNOT \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = |11\rangle$$

$$CNOT \cdot |11\rangle = CNOT \cdot \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = |10\rangle$$

If this seems confusing, the Visual examples of the CNOT gate will be shown in the later section, Quantum Circuits.

Looking forward to entanglement, what happens if the control qubit is in an equal superposition of $|0\rangle$ and $|1\rangle$? As we will soon see, the CNOT gate and the Hadamard gate are central to the entanglement between two qubits.

The RY gate is a rotation operator gate about y-axis. This gate is mentioned in particular because it relates to the CHSH experiment. RX and RZ gates exist as well.

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

It is of note that the explicit matrix form is not always needed. The RY gate can also be represented by

$$e^{-i\frac{\theta}{2}Y}$$

Which can be expanded out by using the Taylor series

$$\begin{aligned} e^{-i\frac{\theta}{2}Y} &= \cos\left(\frac{\theta}{2}\right)I - i\sin\left(\frac{\theta}{2}\right)Y \\ &= \cos\left(\frac{\theta}{2}\right) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - i\sin\left(\frac{\theta}{2}\right) \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{bmatrix} \end{aligned}$$

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. The Y gate represents a rotation by π radians about the y-axis, but with an additional global phase of -i

$$RY(\pi) = -iY$$

The Pauli-Y gate is defined with a conventional global phase to satisfy anti-commutativity along with the Pauli-X and Pauli-Z operators. This is not highly relevant, and the key takeaway should be that $RY(\pi)$ and Y perform the same physical rotation, but differ by a global phase of -i.

2.1.5 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.^[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 $|e^{i\phi}|^2 = 1$

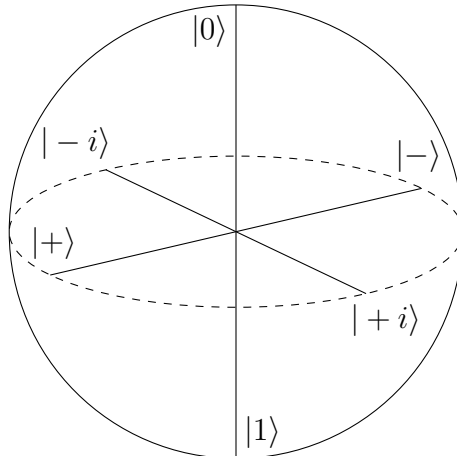
The phase factor on the 1 ket instead of the 0 ket by standard convention. It can be rewritten to be on the 0 ket instead if needed

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

2.1.6 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.

Where the location of the point simultaneously represents the phase and the probability for both $|0\rangle$ and $|1\rangle$.



But what are these different states along the equator?

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

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.

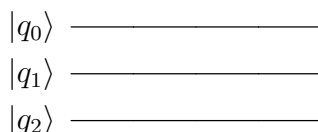
$$\begin{aligned} |+\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + e^{i2\pi} |1\rangle) & |-\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + e^{i\pi} |1\rangle) \\ |+i\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + e^{i\frac{\pi}{2}} |1\rangle) & |-i\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + e^{i\frac{3\pi}{2}} |1\rangle)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

2.2 Quantum Circuits and Entanglement

2.2.1 Quantum Circuits

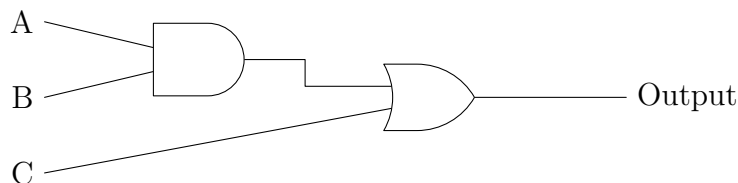
So far, we have discussed gates, phase, and the Bloch sphere, but each of these discussions has only involved a single qubit.² How can we apply a gate to a single qubit in a multi-qubit system, and how can we represent this system graphically?

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



To understand why this diagram is chosen, let us recall the AND and OR gates from the previous section, and let's create a classical circuit.

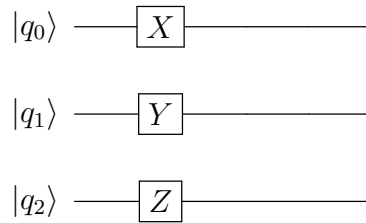
The Classical Circuit in diagram takes 3 inputs, A, B, and C and provides 1 output, after passing A, B, and C through an AND gate and 1 OR gate.



This idea from classical computing 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.

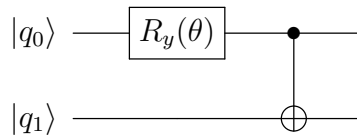
²Except for the CNOT gate

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?

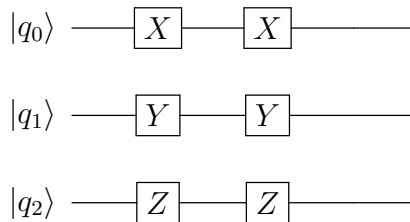


Quantum circuits are intuitive and very useful for visualizing a qubit system. Let's explore a few more examples.

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



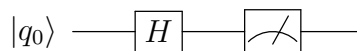
This system has two of each Pauli gate applied to its respective qubit. Recall that the X, Y, and Z gates are unitary, so this circuit is equivalent to a circuit without any gates applied, like the one at the start of this section.



2.2.2 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 $|0\rangle$ or a $|1\rangle$, and splits it into equal probabilities. Thus, if it is measured, the qubit will have a 50/50 chance to collapse into either state.

Consider the X-basis states, $|+\rangle$ and $|-\rangle$, and a qubit in the state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, such that $|\alpha|^2 + |\beta|^2 = 1$

By the Born rule[1], the probability of the qubit being in the $|+\rangle$ is

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

Where the probability is

$$p(+)=\left|\frac{\alpha+\beta}{\sqrt{2}}\right|^2=\frac{|\alpha+\beta|^2}{2}$$

And similarly for $|-\rangle$

$$\langle\psi|-\rangle\langle-|\psi\rangle = |\langle-|\psi\rangle|^2 = \frac{\langle 0| - \langle 1|}{\sqrt{2}}(\alpha|0\rangle + \beta|1\rangle) = \frac{\alpha - \beta}{\sqrt{2}}$$

Where the probability is

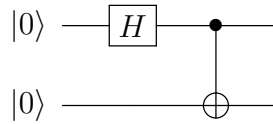
$$p(-)=\left|\frac{\alpha-\beta}{\sqrt{2}}\right|^2=\frac{|\alpha-\beta|^2}{2}$$

So if $\alpha = 1, \beta = 0$, an equal probability emerges.

$$p(+)=\frac{|1+0|^2}{2}=1/2 \qquad p(-)=\frac{|1-0|^2}{2}=1/2$$

2.2.3 Bell States and Entanglement

As we know from the Hadamard gate, we can take a basis state, and split it into a superposition. So what if we take two $|0\rangle$ qubits, apply a Hadamard gate, and then apply a CNOT gate?



This two qubit circuit is a very important circuit, and we will be seeing it frequently. This 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 this 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? What is interesting about Bell states? First, let's consider a two qubit system, $|\psi_0\rangle$ and $|\psi_1\rangle$

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

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

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 the state is not separable into a product state. This is the core idea of entanglement.

3 CHSH Inequality

The CHSH inequality is a more robust version of Bell’s inequality. Both are used to prove Bell’s theorem, however Bell’s inequality requires a noiseless environment and is practically difficult to implement.

The section will outline the theory behind the CHSH experiment, including derivations of the CHSH inequality, how the expectation values are calculated, and the true upper bound for the CHSH parameter, known as Tsirelson’s bound.

3.1 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 influenced only by their 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 independent of observation. However, the CHSH experiment is empirical 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]

3.2 The CHSH parameter, S

The CHSH parameter³, 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 arises by strategic choice rather than by any derivation. Bell’s original inequality was ineffective against experimental noise, and the CHSH improved upon Bell’s ideas and designed the CHSH parameter to maximize violation of the CHSH inequality and to simplify experimentation. [3] The idea of a maximum violation will be expanded upon in future sections.

³Throughout this paper, S is referred to using the terms ”parameter” and ”value” interchangeably. However, ”parameter” is the more accurate term.

3.3 Expectation Values

The CHSH parameter, S , is composed of four expectation values

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

Where A_i and B_j represent the "measurement settings" for two qubits, which we will name Alice and Bob respectively. To compute S , we must first calculate these individual expectation values.

These measurement settings correlate to what basis the qubits are being measured in. Specifically, the qubits will either be measured with respect to the X or Z basis.

We choose the standard CHSH settings: Alice uses A_0 in the X basis and A_1 in the Z basis, Bob uses B_0 along $(X + Z)/\sqrt{2}$ (a $+45^\circ$ rotation about y) and B_1 along $(X - Z)/\sqrt{2}$ (a -45° rotation). In our code this corresponds to angles $A_0 : \pi/2$, $A_1 : 0$, $B_0 : \pi/4$, $B_1 : -\pi/4$ applied via $R_Y(-\theta)$.

To ground the next calculations, let us recall that expectation values quantify how measurement outcomes on entangled qubits correlate when projected onto different bases. In particular, computing the eigenvalues and eigenvectors of the Pauli operators, Z and X reveals the possible outcomes as ± 1 and their associated states. By explicitly finding these determinants, we confirm each measurement, setting the stage for constructing the CHSH parameter.

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

Where the eigenstates are $|0\rangle$ and $|1\rangle$, with eigenvalues of ± 1

$$\det(Z - \lambda I) = 0 = (1 - \lambda)(-1 - \lambda) \implies \lambda = \pm 1$$

For $\lambda = +1$:

$$(Z - I)\mathbf{v} = 0$$

$$\begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \rightarrow v_2 = 0$$

$$\mathbf{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle$$

For $\lambda = -1$:

$$(Z + I)\mathbf{v} = 0$$

$$\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \rightarrow v_1 = 0$$

$$\mathbf{v} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = |1\rangle$$

Having determined that the eigenvalues of Z are ± 1 with eigenstates $|0\rangle$ and $|1\rangle$, we now turn our attention to the Pauli- X operator. We will find that we get the same eigenvalues ± 1 , however now with eigenvectors $|+\rangle$ and $|-\rangle$ respectively.

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

Where the eigenstates are $|+\rangle$ and $|-\rangle$, with eigenvalues of ± 1

$$\det(X - \lambda I) = 0 = \lambda^2 - 1 \implies \lambda = \pm 1$$

For $\lambda = +1$:

$$(X - I)\mathbf{v} = 0$$

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \implies v_1 = v_2$$

$$\mathbf{v} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = |+\rangle$$

For $\lambda = -1$:

$$(X + I)\mathbf{v} = 0$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \implies v_1 = -v_2$$

$$\mathbf{v} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = |-\rangle$$

With the eigenvalue and vectors for Z and X determined, we can now formalize the expectation value for a two-qubit measurement.

An individual expectation value $E(A_i B_j)$ is computed by

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

Where $|\psi\rangle$ is the shared entangled Bell state, $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$

3.4 The Classical Bound, $S \leq 2$

To continue with the derivation of the classical bound, we must now assume realism and locality. This means that before Alice and Bob pick how they will measure their qubits, the outcomes are already predetermined by some underlying hidden variable, λ .

We can then write the expectation values from repeated trials where Alice and Bob record their outcomes, a_k, b_k for the measurement settings A_i, B_j .⁴

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

Where $a_k, b_k \in \{\pm 1\}$ are the outcomes of a single observation, summed for the total number of trials.

If both of the outcomes are ± 1 , 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

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

To proceed with the derivation of the classical inequality, we want to make the assumption that outcomes are predetermined by some hidden variable, which we will call λ . Locality ensures that outcomes only depend on local settings, and realism assumes that outcomes exist independent 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

$$S(\lambda) = a_0(\lambda)b_0(\lambda) - a_0(\lambda)b_1(\lambda) + a_1(\lambda)b_0(\lambda) + a_1(\lambda)b_1(\lambda)$$

Since we know the possible values of the correlations and the value for S , we can then determine the bound of S with the assumptions of realism and locality. But first, let's rewrite S to be a bit easier to work with.

$$S(\lambda) = a_0(\lambda)[b_0(\lambda) - b_1(\lambda)] + a_1(\lambda)[b_0(\lambda) + b_1(\lambda)]$$

Recall that each value can only take on the values of ± 1 , so now we can tabulate all of the possible combinations of values and evaluate from there.

⁴In practice, one replaces the ensemble average by a finite sum over N trials, incurring only negligible statistical error for large N .

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	$-2a_1$

In all cases and for any single hidden variable, λ , we can deduce the classical bound $|S(\lambda)| \leq 2$.

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

To optimize the violation of the CHSH inequality, it has been found that we want Alice and Bob's measurement bases to be misaligned by 45° on the Bloch sphere. The Z and X bases are orthogonal, so to get these optimal angles we use the RY gate to rotate by 45° . The RY gate will be applied to both qubits after they have been entangled.

Looking ahead and knowing that the value $|S(\lambda)| \leq 2$ will be violated, then what is the maximum possible violation?

3.5 Tsirelson's bound

Looking ahead and knowing that the classical bound will be violated, what is the maximum possible violation within the realm of quantum mechanics?

This maximum value is known as the Tsirelson bound, named after the mathematician Boris S. Tsirelson. In his 1980 paper, Tsirelson derived the upper bound for the CHSH parameter and highlighted the differences in Bell's 1964 framework of local hidden variable theories. [4] [2]

We will now use a slightly different notation than before, where expectation values are now placed inside angle brackets. Additionally, we will use the operator norm of the CHSH operator, T for our derivation. This is not the same as S , but it is directly related by the expectation value.

$$S = \langle A_0 \otimes B_0 \rangle - \langle A_0 \otimes B_1 \rangle + \langle A_1 \otimes B_0 \rangle + \langle A_1 \otimes B_1 \rangle$$

$$S = \langle \psi | T | \psi \rangle$$

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

For any normalized value, $|\psi\rangle$, we can use the operator norm to find the inequality

$$|\langle \psi | T | \psi \rangle| \leq \|T\| = \sqrt{\|T^2\|}$$

Therefore,

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

Before we continue with tensor algebra manipulation of T , it is vitally important to note that operations are not commutative. i.e. $A_0A_1 \neq A_1A_0$

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

Now as an intermediate step, let's square both sides.

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

$$T^2 = (A_0 \otimes (B_0 - B_1))^2 + (A_1 \otimes (B_0 + B_1))^2 + A_0A_1 \otimes ((B_0 - B_1)(B_0 + B_1)) + A_1A_0 \otimes ((B_0 + B_1)(B_0 - B_1))$$

Let's break this into parts and first consider the square terms and recall that $A_i^2 = B_j^2 = I$

$$(B_0 - B_1)^2 = 2I - (B_0B_1 + B_1B_0)$$

$$(B_0 + B_1)^2 = 2I + (B_0B_1 + B_1B_0)$$

Thus,

$$(A_0 \otimes (B_0 - B_1))^2 = A_0^2 \otimes (B_0 - B_1)^2 = I \otimes (2I - (B_0B_1 + B_1B_0))$$

$$(A_1 \otimes (B_0 + B_1))^2 = A_1^2 \otimes (B_0 + B_1)^2 = I \otimes (2I + (B_0B_1 + B_1B_0))$$

Adding them together gives,

$$I \otimes [2I - (B_0B_1 + B_1B_0) + 2I + (B_0B_1 + B_1B_0)] = I \otimes 4I = 4I$$

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

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

Combining all of the terms,

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

Notice that if the commutators are zero,

$$T^2 = 4I \implies ||T|| = 2$$

Thus the maximum eigenvalue is 2, and we naturally see the classical bound,

$$|S| = |\langle \Phi^+ | T | \Phi^+ \rangle| \leq 2$$

To continue with the derivation, we must evaluate these commutators with the commutation relations for the Pauli matrices as stated by Nielsen and Chuang[1].

Recall that the values A_0, B_0 correlate to Z, and the values A_1, B_1 correlate the X. We can then derive the commutation relations for Z and X.

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

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

$$[Z, X] = ZX - XZ = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}.$$

Compare with $2iY$:

$$2iY = 2i \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}.$$

$$[Z, X] = 2iY$$

We now distinguish Y_A and Y_B as acting on Alice's and Bob's qubit, respectively.

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

$$[A_0, A_1] = [Z, X] = 2iY_A \quad [B_0, B_1] = [Z, X] = 2iY_B$$

$$T^2 = 4I + 2iY_A \otimes 2iY_B = 4I - 4(Y_A \otimes Y_B)$$

While Y_A and Y_B appear mathematically indistinguishable, we must deal with them separately. Let's do this by evaluating the tensor product with respect to the bell state.

$$\begin{aligned} & \langle \Phi^+ | Y_A \otimes Y_B | \Phi^+ \rangle \\ &= \frac{1}{(\sqrt{2})^2} (\langle 00 | + \langle 11 | | Y_A \otimes Y_B | | 00 \rangle + | 11 \rangle) \end{aligned}$$

It is vitally important that we keep track of which qubit we're working with.

$$= \frac{1}{2} (\langle 0_A | \langle 0_B | + \langle 1_A | \langle 1_B | | Y_A \otimes Y_B | | 0_A \rangle | 0_B \rangle + | 1_A \rangle | 1_B \rangle)$$

After expanding it out, this becomes the rather messy form

$$\begin{aligned} &= \frac{1}{2} [(\langle 0_A | \langle 0_B | Y_A \otimes Y_B | 0_A \rangle | 0_B \rangle) \\ &\quad + (\langle 0_A | \langle 0_B | Y_A \otimes Y_B | 1_A \rangle | 1_B \rangle) \\ &\quad + (\langle 1_A | \langle 1_B | Y_A \otimes Y_B | 0_A \rangle | 0_B \rangle) \\ &\quad + (\langle 1_A | \langle 1_B | Y_A \otimes Y_B | 1_A \rangle | 1_B \rangle)] \end{aligned}$$

With some matrix multiplication, we find that $\langle 0 | Y | 0 \rangle = \langle 1 | Y | 1 \rangle = 0$, which greatly simplifies our equation to just the cross terms

$$= \frac{1}{2} [\langle 0_A | Y_A | 1_A \rangle \langle 0_B | Y_B | 1_B \rangle + \langle 1_A | Y_A | 0_A \rangle \langle 1_B | Y_B | 0_B \rangle]$$

Now all that remains is finish the matrix multiplication.

$$\langle 0 | Y | 1 \rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -i$$

Similarly we find,

$$\langle 1 | Y | 0 \rangle = \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = i$$

Which then brings our original form to the rather simple equation,

$$\frac{1}{2}((-i)^2 + (i)^2) = -1$$

Recall that we had to use the bell state to get this value, which simultaneously transformed our T^2 value

$$\langle T^2 \rangle = 4 - 4\langle Y_A \otimes Y_B \rangle = 8$$

We then get the result

$$|S| \leq ||T|| = 2\sqrt{2}$$

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 $2\sqrt{2}$

4 CHSH Experiment

Qiskit, or the Quantum Information Software Kit, is an open source software development kit (SDK) created by IBM researchers to execute programs on quantum computers. Qiskit has various tools for creating and manipulating quantum information.

Qiskit allows usage of their prototype quantum devices on the IBM Quantum Platform or by using locally run simulators. In this section there are 3 different programs. Two of which were written using the Aer simulator, and one which is compatible with IBM quantum hardware.

Only snippets of the full code will be shown. Only code that directly relates to the CHSH experiment will be mentioned and little to no code relating to the Qiskit backend will be shown. The complete code can be found on <https://github.com/acroscopic/CHSH/> and it is recommended to reference while reading this paper.

4.1 Aer Simulator and IBM Hardware

We will begin by using the Aer simulator module for Qiskit to recreate the CHSH experiment, and then we will use real IBM Quantum Hardware.[5]. Three total backends were used. Two Aer simulator backends and a backend to connect with an IBM Quantum Computer. The two Aer simulators used are the Qasm Simulator which simulates a noisy quantum circuit, and the Statevector Simulator, which is an ideal or noiseless quantum circuit.

4.2 Statevector Simulator

Let's begin by using the Aer Statevector simulator. This is a noiseless backend, which means that it will produce the same value each time the program is run.

We begin by defining angles. We will later reference these angles when we use the RY gate.

```
1 # These are the angles at which we're measuring our qubits
2 angles = {
3     # Pure X-basis measurement
4     'A0': np.pi / 2,
5
6     # Pure Z-basis measurement
7     'A1': 0,
8
9     # X-basis (Rotated 45 degrees)
10    'B0': np.pi / 4,
11
12    # Z-basis (Rotated 45 degrees)
13    'B1': -np.pi / 4
14 }
```

Next, let's define the measurement settings we'll be using,

```
1 # Recall  $S = E(A_0B_0) - E(A_0B_1) + E(A_1B_0) + E(A_1B_1)$ 
2 # This is just setting up the settings so we can loop over them
3 # We will subtract the  $E(A_0B_1)$  later
4 measurement_settings = [
5     ('A0', 'B0'), # Add
6     ('A0', 'B1'), # Subtract (later)
7     ('A1', 'B0'), # Add
8     ('A1', 'B1'), # Add
9 ]
```

We need to loop through the measurement settings, create a Bell state, and apply the measurement rotations.

```
1     # Start looping through the measurement settings
2     # We need to use both qubits, which we've named Alice and Bob
3     for alice, bob in measurement_settings:
4
5         # Create a quantum circuit with 2 qubits
6         qc = QuantumCircuit(2)
7
8         # First, we need to create Bell state between Alice and Bob
9         # Put a Hadamard gate on Alice (Qubit 0)
10        qc.h(0)
11        # Then, put a CNOT gate on Alice and Bob (Controlled by
12        # Alice)
13        qc.cx(0, 1)
14
15        # Apply RY gate rotations to the qubits to align the
16        # measurement basis
17        qc.ry(-angles[alice], 0) # Alice's basis
18        qc.ry(-angles[bob], 1)   # Bob's basis
```

Now that we have prepared the system, we will make a measurement and find the expectation values. This is done by simulating the statevectors for each measurement setting, and then squaring the probability amplitude. All that remains is compute the CHSH value.

```
1 if alice == 'A0' and bob == 'B1':
2     S -= E # Subtract only  $E(A_0B_1)$ 
3 else:
4     S += E # Add everything else
5 print(f"E({alice}{bob}) = {E}")
6
7 print(f"S = {S}")
8 print(f"2*sqrt(2) = {2*np.sqrt(2)}")
```

Since this is a noiseless environment, we expect the analytic values which lead to $S = 2\sqrt{2}$

$$E(A_0B_0) = E(A_1B_0) = E(A_1B_1) = \frac{\sqrt{2}}{2}, E(A_0B_1) = -\frac{\sqrt{2}}{2}$$

4.2.1 Results

```
Calculating individual expectation values:
E(A0B0) = 0.70710678
E(A0B1) = -0.70710678
E(A1B0) = 0.70710678
E(A1B1) = 0.70710678

Calculating CHSH Value:
S = 2.8284271247...
2√2 = 2.8284271247...
```

Figure 1: Terminal output from running `Aer_StateVector.py`

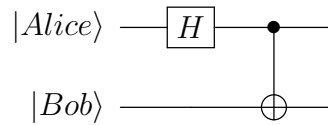
The Aer statevector simulator has provided a noiseless environment to run the CHSH experiment. We can see that the value obtained is identical to $2\sqrt{2}$ and will always produce this result when run. However running the CHSH experiment on real hardware does not have the luxury of a noiseless environment. To see what we might expect, let's continue with the Aer simulator, but with a new noisy backend.

4.3 Qasm Simulator

The Qasm Simulator works similarly to the Statevector Simulator. We will begin again by creating 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
4 def bell():
5     # Create a quantum circuit
6     qc = QuantumCircuit(q, c)
7     alice = q[0]           # The first qubit
8     bob = q[1]             # The second qubit
9     qc.h(alice)            # Hadamard Gate
10    qc.cx(alice, bob)       # CNOT Gate
11
12    return qc
```

Recall the graphical visualization of this bell state. The only difference between this circuit and the code, is that two classical bits are created for storing the results after measurement.

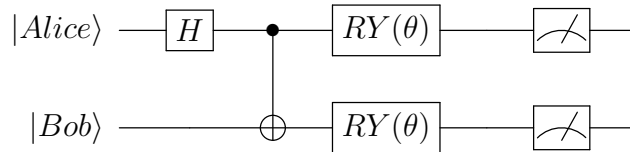


The next step is to measure the qubits and store them in the classical bits. However, before measurement, the RY gate is applied.

```

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

```

4.3.1 Results

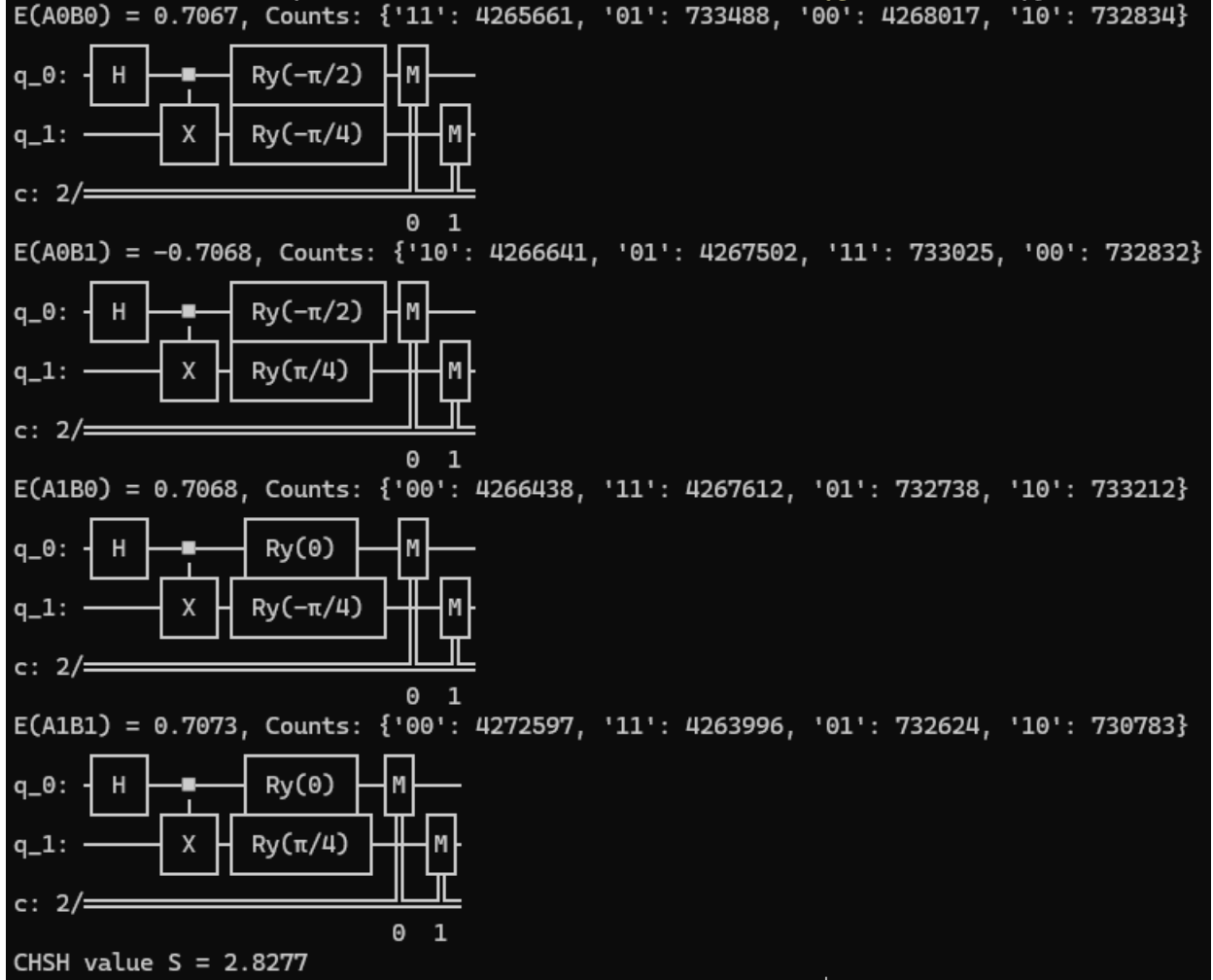


Figure 2: Terminal output from running Aer_Qasm.py

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 $2\sqrt{2} \approx 2.82842...$

Since the Qasm simulator is a noisy quantum circuit simulator, if you do not have a high enough shot count, values above Tsirelson's bound can be achieved. This would normally be a major alarm, but this is an issue inherent in finite sampling within the Qasm simulator. Using a higher shot count, the values get closer to the bound, $2\sqrt{2}$, and are more likely to stay under the bound.

CHSH value S = 2.8288

Figure 3: Finite sampling artifact above Tsirelson's bound.

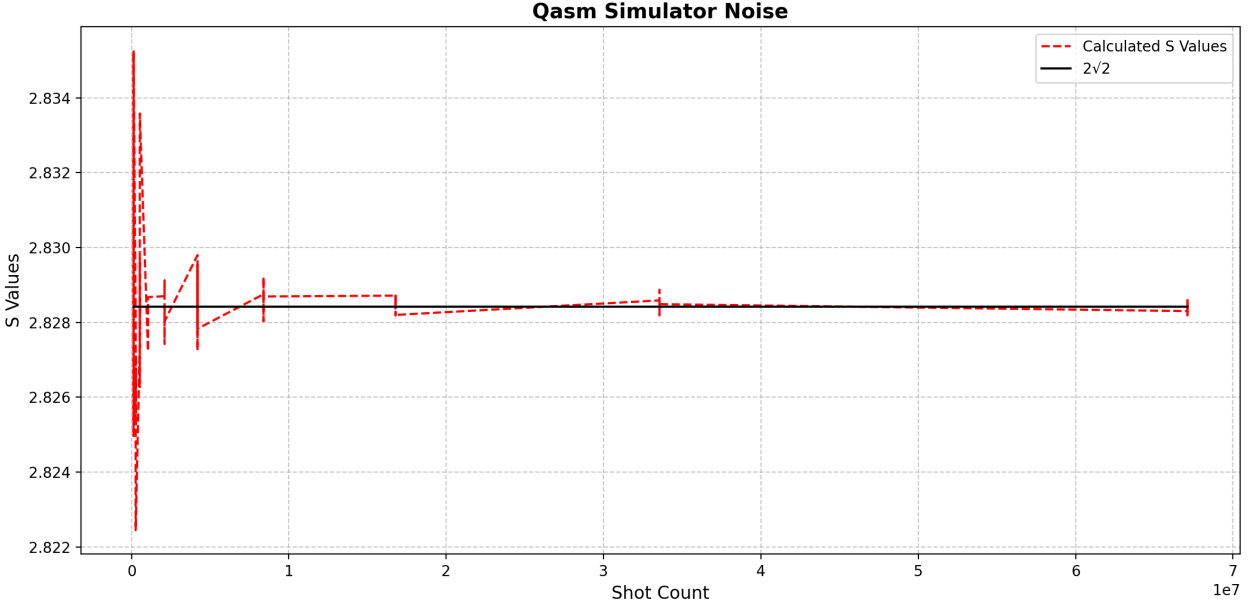


Figure 4: Tabulated S parameter values alongside the shots used, highlighting an issue inherent in finite sampling. We see that as higher amounts of calculations as done, we start to converge to $2\sqrt{2}$, however it will not remain strictly below or above the line.

4.4 Execution on IBM Quantum Hardware

Previously, all code has been run using the Aer simulator and did not interface with any real quantum hardware. It is highly relevant to empirically replicate the results of the CHSH experiment using a real quantum computer [6]. 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. For this section in particular, it is recommended to view the full code while reading: https://github.com/acroscopic/CHSH/blob/main/CHSH_Experiment.py. However, only code snippets will be included in this paper for brevity.

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[6].

```
1 backend = service.least_busy(simulator=False, operational=True,
    min_num_qubits=127)
```

Instead of running lots of calculations to find the expectation values as before, hardware requires a more analytic approach. In addition, the angle for the RY gate will be parametrized between equally spaced values over a full 0 to 2π rotation. This will allow us to see values that do and don't violate the inequality, and how we can optimize it.

```
1 # Parametrize different angles for the RY gate
2 theta = Parameter("$\\theta$")
3 number_of_phases = 20
```

```

4 phases = np.linspace(0, 2 * np.pi, number_of_phases)
5
6 # Creating the quantum circuit in the same manner as before
7 qc = QuantumCircuit(2)
8 qc.h(0)
9 qc.cx(0, 1)
10 qc.ry(theta, 0)

```

One of the biggest differences 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 that is available before we submit the job to queue and wait to receive the results. After doing that, we can run the job.

```

1 # Running the CHSH Experiment after aligning with the specifics of
  the backend.
2 job = estimator.run(pubs=[pub])
3 result = job.result()
4
5 # Expectation values for the observable
6 S_values = result[0].data.evs[0]
7 print(S_values)
8
9 violation = np.any(np.abs(S_values) > 2)
10 print(f"CHSH violation detected: {violation}")
11
12 # Unique identifier for job
13 print(f"Job ID: {job.job_id()}")
14
15 # Should be 'DONE' if successful
16 print(f"Job status: {job.status()}")

```

4.4.1 Results

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

```

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 5: Terminal output from running CHSH_experiment.py

5 Conclusion

After executing the CHSH experiment on both a noiseless and a noisy Aer Simulators and on IBM hardware, we observed the CHSH parameter, $S > 2$, violating the classical bound of $|S| \leq 2$. This empirical violation of the CHSH inequality provides strong evidence against local hidden variable theories, which assume that outcomes are determined by hidden properties independent of distant measurements.

By violating this inequality, we have empirically validated Bell’s theorem, demonstrating that quantum mechanics exhibits non-classical correlations, incompatible with local realism. More specifically, the results reject local realism, which is the joint assumptions of locality and realism. Quantum mechanics inherently incorporates non-local correlations through entangled states, as seen by the Bell state used in this experiment.

This violation implies that quantum mechanics exhibits non-locality. While this does not imply instantaneous communication due to the no-signaling theorem, it confirms that quantum systems cannot be described by the local deterministic methods from classical physics without violating this empirical evidence.

The noisy simulator and hardware results highlight the fragility of quantum correlations in practical settings. Despite this, the measured violations consistently surpass $S = 2$.

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. The results align with the Copenhagen interpretation, and the violation challenges many classically held intuitions about causality and locality.

In summary, this experiment shows that quantum mechanics is not solely a statistical tool, but a fundamental framework to describe reality.

6 Acknowledgements

This paper was written as an honors project for the undergraduate Quantum Mechanics course, PHY 3035, at the Florida Institute of Technology.

I owe my deepest gratitude to Dr. Souvik Das, whose support made this work possible and for doing an outstanding job structuring my introduction to the world of quantum. He generously stayed late to guide my work, furthered my knowledge, meticulously proofread, and found more typos and grammar slip-ups than I care to admit. Dr. Das also provided me with invaluable presentation opportunities: twice with his Quantum Computing Research Group and once at a lecture attended by students from mathematics, physics, and space science, and my fellow peers in PHY 3035.

My thanks go to Krupa Pothiwala and Grant Lindell for their enthusiastic collaboration. From the earliest brainstorming sessions in the problem sets for PHY 3035, their constructive feedback in proofreading, and their dedication that pushed this project to a higher standard.

I am also grateful to Scott Demarest for his clear explanation of the fundamental limitations on separability of entangled particles, and for kindly allowing me to use his derivation demonstrating that entangled particles cannot be decomposed into a product state of two subsystems.

Finally, I would like to extend my thanks to Florida Tech for fostering an environment where undergraduate research can flourish, and I am thankful to all of the faculty, staff, and fellow students who contributed their time and expertise, directly or indirectly, to my education.

7 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] Tsirelson, B. S. "Quantum Generalizations of Bell's Inequality." *Letters in Mathematical Physics*, vol. 4, pp. 93–100, 1980.
- [5] IBM, *Qiskit: An Open-Source Framework for Quantum Computing*, 2025, 2.0, <https://qiskit.org>
- [6] IBM, *IBM Quantum Computing*, 2025, <https://quantum-computing.ibm.com>