

Lectures on Quantum Information Science

Artur Ekert

2020-11-15

Contents

| | |
|--|-----------|
| Overview | 5 |
| 1 Quantum interference | 7 |
| 1.1 Two basic rules | 7 |
| 1.2 Quantum interference (the failure of probability theory) | 8 |
| 1.3 Superpositions | 11 |
| 1.4 Interferometers | 12 |
| 1.5 Qubits, gates, and circuits | 12 |
| 1.6 Quantum decoherence | 12 |
| 1.7 Computation: deterministic, probabilistic, and quantum | 12 |
| 1.8 Computational complexity | 12 |
| 1.9 Outlook | 13 |
| 1.10 Notes and Exercises | 14 |
| 1.11 Supplement: Physics against logic, via beamsplitters | 14 |
| 1.12 Supplement: Quantum interference revisited (still about beam- splitters) | 14 |
| 2 Qubits | 15 |
| 3 Measurements | 17 |
| 4 Quantum entanglement | 19 |
| 5 Quantum algorithms | 21 |

| | | |
|-----------|--|-----------|
| 6 | Bell's theorem | 23 |
| 6.1 | Quantum correlations | 23 |
| 6.2 | Hidden variables | 24 |
| 6.3 | CHSH inequality | 25 |
| 6.4 | Quantum correlations revisited | 26 |
| 6.5 | Tsirelson's inequality | 26 |
| 7 | Decoherence, and elements of quantum error correction | 29 |
| 8 | Density matrices | 31 |
| 9 | Quantum channels (or CP maps) | 33 |
| 10 | Quantum error correction and fault tolerance | 35 |

Overview

Chapter 1

Quantum interference

About complex numbers, called probability amplitudes, that, unlike probabilities, can cancel each other out, leading to quantum interference and qualitatively new ways of processing information.

The classical theory of computation does not usually refer to physics. Pioneers such as Alan Turing, Alonzo Church, Emil Post and Kurt Gödel managed to capture the correct classical theory by intuition alone and, as a result, it is often falsely assumed that its foundations are self-evident and purely abstract. They are not!¹

The concepts of information and computation can be properly formulated only in the context of a physical theory — information is stored, transmitted and processed always by *physical* means. Computers are physical objects and computation is a physical process. Indeed, any computation, classical or quantum, can be viewed in terms of physical experiments, which produce **outputs** that depend on initial preparations called **inputs**. Once we abandon the classical view of computation as a purely logical notion independent of the laws of physics it becomes clear that whenever we improve our knowledge about physical reality, we may also gain new means of computation. Thus, from this perspective, it is not very surprising that the discovery of quantum mechanics in particular has changed our understanding of the nature of computation. In order to explain what makes quantum computers so different from their classical counterparts, we begin with the rudiments of quantum theory.

1.1 Two basic rules

Quantum theory, at least at some instrumental level, can be viewed as a modification of probability theory. We replace positive numbers (probabilities) with

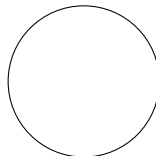
¹Computation is a physical process. Computation is a physical process. Computation is ...

complex numbers z (called **probability amplitudes**) such that the squares of their absolute values, $|z|^2$, are interpreted as probabilities.

Definition 1.1. The correspondence between probability amplitude z and probability $p = |z|^2$ is known as **Born's Rule**.

The rules for combining amplitudes are very reminiscent of the rules for combining probabilities:

1. Whenever something can happen in a sequence of independent steps, we multiply the amplitudes of each step.



2. Whenever something can happen in several alternative ways, we add the amplitudes for each separate way.

That's it! These two rules are basically all you need to manipulate amplitudes in any physical process, no matter how complicated.² They are universal and apply to any physical system, from elementary particles through atoms and molecules to white dwarfs stars. They also apply to information, since, as we have already emphasised, information is physical. The two rules look deceptively simple but, as you will see in a moment, their consequences are anything but trivial.

1.2 Quantum interference (the failure of probability theory)

Modern mathematical probability theory is based on three axioms, proposed by Andrey Nikolaevich Kolmogorov (1903–1987) in his monograph with the impressive German title *Grundbegriffe der Wahrscheinlichkeitsrechnung* (“Foundations of Probability Theory”). The **Kolmogorov axioms** are simple and intuitive:³

1. Once you identify all elementary outcomes, or events, you may then assign probabilities to them.

²We will, however, amend the two rules later on when we touch upon particle statistics.

³I always found it an interesting coincidence that the two basic ingredients of modern quantum theory, namely probability and complex numbers, were discovered by the same person, an extraordinary man of many talents: a gambling scholar by the name of Girolamo Cardano (1501–1576).

1.2. QUANTUM INTERFERENCE (THE FAILURE OF PROBABILITY THEORY)9

2. Probability is a number between 0 and 1, and an event which is certain has probability 1.
3. Last but not least, the probability of any event can be calculated using a deceptively simple rule — the *additivity axiom*:

Whenever an event can occur in several mutually exclusive ways, the probability for the event is the sum of the probabilities for each way considered separately.

Obvious, isn't it? So obvious, in fact, that probability theory was accepted as a mathematical framework theory, a language that can be used to describe actual physical phenomena. Physics should be able to identify elementary events and assign numerical probabilities to them. Once this is done we may revert to mathematical formalism of probability theory. The Kolmogorov axioms will take care of the mathematical consistency and will guide us whenever there is a need to calculate probabilities of more complex events. This is a very sensible approach, apart from the fact that it does not always work! Today, we know that probability theory, as ubiquitous as it is, fails to describe many common quantum phenomena. In order to see the need for quantum theory let us consider a simple experiment in which probability theory fails to give the right predictions.

1.2.1 The double slit experiment

In a double slit experiment, a particle emitted from a source S can reach the detector D by taking two different paths, e.g. through an upper or a lower slit in a barrier between the source and the detector. After sufficiently many repetitions of this experiment we can evaluate the frequency of clicks in the detector D and show that it is inconsistent with the predictions based on probability theory. Let us use the quantum approach to show how the discrepancy arises.

The particle emitted from a source S can reach detector D by taking two different paths, with amplitudes z_1 and z_2 respectively. We may say that the upper slit is taken with probability $p_1 = |z_1|^2$ and the lower slit with probability $p_2 = |z_2|^2$. These are two mutually exclusive events. With the two slits open, probability theory declares (by the additivity axiom) that the particle should reach the detector with probability $p_1 + p_2 = |z_1|^2 + |z_2|^2$. But this is not what happens experimentally!

Following the “quantum rules’’, first we add the amplitudes and then we square the absolute value of the sum to get the probability. Thus, the particle will

reach the detector with probability

$$\begin{aligned}
 p = |z|^2 &= |z_1 + z_2|^2 = |z_1|^2 + |z_2|^2 + z_1^* z_2 + z_1 z_2^* \\
 &= p_1 + p_2 + |z_1||z_2| (e^{i(\varphi_2 - \varphi_1)} + e^{-i(\varphi_2 - \varphi_1)}) \\
 &= p_1 + p_2 + 2\sqrt{p_1 p_2} \cos(\varphi_2 - \varphi_1) \\
 &= p_1 + p_2 + \text{interference terms}
 \end{aligned}$$

where we have expressed the amplitudes in their polar forms

$$\begin{aligned}
 z_1 &= |z_1| e^{i\varphi_1} \\
 z_2 &= |z_2| e^{i\varphi_2}.
 \end{aligned}$$

The appearance of the interference terms marks the departure from the classical theory of probability. The probability of any two seemingly mutually exclusive events is the sum of the probabilities of the individual events, $p_1 + p_2$, *modified* by the **interference term** $2\sqrt{p_1 p_2} \cos(\varphi_2 - \varphi_1)$. Depending on the **relative phase** $\varphi_2 - \varphi_1$, the interference term can be either negative (which we call **destructive** interference) or positive (**constructive** interference), leading to either suppression or enhancement of the total probability p .

The algebra is simple; our focus is on the physical interpretation. Firstly, note that the important quantity here is the *relative* phase $\varphi_2 - \varphi_1$ rather than the individual values φ_1 and φ_2 . This observation is not trivial at all. If a particle reacts only to the difference of the two phases, each pertaining to a separate path, then it must have, somehow, experienced the two paths, right? Thus we cannot say that the particle has travelled either through the upper or the lower slit, because it has travelled through *both*. In the same way, quantum computers follow, in some tangible way, all computational paths simultaneously, producing answers that depend on all these alternative calculations. Weird, but this is how it is! Secondly, what has happened to the additivity axiom in probability theory — what is wrong with it? One thing that is wrong is the assumption that the processes of taking the upper or the lower slit are mutually exclusive. In reality, as we have just mentioned, the two transitions *both occur*, simultaneously. However, we cannot learn this from probability theory, or any other a priori mathematical construct. There is no fundamental reason why Nature should conform to the additivity axiom.⁴ We find out how nature works by making intelligent guesses, running experiments, checking what happens and formulating physical theories. If our guess disagrees with experiments then it is wrong, so we try another intelligent guess, and another, etc. Right now, quantum theory is the best guess we have: it offers good explanations and predictions that have not been falsified by any of the existing experiments. This said, be assured that one day quantum theory will be falsified and we will have to start guessing again.

⁴According to the philosopher Karl Popper (1902–1994) a theory is genuinely scientific only if it is possible, in principle, to establish that it is false. Genuinely scientific theories are never finally confirmed because no matter how many confirming observations have been made observations that are inconsistent with the empirical predictions of the theory are always possible.

1.3 Superpositions

Amplitudes are more than just tools for calculating probabilities: they tell us something about physical reality. When we deal with probabilities, we may think about them as numbers that quantify our lack of knowledge. Indeed, when we say that a particle goes through the upper or the lower slit with some respective probabilities it does go through one of the two slits, we just do not know which one. In contrast, according to quantum theory, a particle that goes through the upper and the lower slit with certain amplitudes does explore both of the two paths, not just one of them. This is a statement about a real physical situation, about something that is out there and something we can experiment with. The assumption that the particle goes through one of the two slits, but we do not know which one, is inconsistent with the experimental observations. We have to accept that apart from some easy to visualise states, also known as the basis states, such as the particle at the upper slit or the particle at the lower slit, there are infinitely many other states, all of them equally real, in which the particle is in a *superposition* of the two basis states. This rather bizarre picture of reality is the best we have at the moment, and it works, at least for now. Physicists write such states as ⁵

$$|\psi\rangle = \alpha|\text{at the upper slit}\rangle + \beta|\text{at the lower slit}\rangle,$$

meaning the particle at the upper slit with amplitude α and at the lower slit with amplitude β . Mathematically, you can think about this expression as a vector $|\psi\rangle$ in a two-dimensional complex vector space written in terms of the two basis vectors $|\text{at the upper slit}\rangle$ and $|\text{at the lower slit}\rangle$. You can also write this vector as a column vector with two complex entries α and β , but then you have to explain the physical meaning of the basis states. Here, we use the $|\cdot\rangle$ notation, introduced by Paul Dirac in the early days of the quantum theory as a useful way to write and manipulate vectors. In Dirac notation you can put into the box $|\cdot\rangle$ anything that serves to specify what the vector is. It could be $|\uparrow\rangle$ for spin up and $|\downarrow\rangle$ for spin down, or $|0\rangle$ for a quantum bit holding logical 0 and $|1\rangle$ for a quantum bit holding logical 1, etc. As we shall see soon, there is much more to this notation.

⁵Dirac notation will likely be familiar to physicists, but may look odd to mathematicians or computer scientists. Love it or hate it (and I suggest the former), the notation is so common that you simply have no choice but to learn it, especially if you want to study anything related to quantum theory.

1.4 Interferometers

1.5 Qubits, gates, and circuits

1.6 Quantum decoherence

1.7 Computation: deterministic, probabilistic, and quantum

1.8 Computational complexity

Is there a compelling reason why we should care about quantum computation? It may sound like an extravagant way to compute something that can be computed anyway. Indeed, your standard laptop, given enough time and memory, can simulate pretty much any physical process. In principle, it can also simulate any quantum interference and compute everything that quantum computers can compute. The snag is, this simulation, in general, is very inefficient. And efficiency does matter, especially if you have to wait more than the age of the Universe for your laptop to stop and deliver an answer!⁶

In order to solve a particular problem, computers (classical or quantum) follow a precise set of instructions — an **algorithm**. Computer scientists quantify the efficiency of an algorithm according to how rapidly its running time, or the use of memory, increases when it is given ever larger inputs to work on. An algorithm is said to be *efficient* if the number of elementary operations taken to execute it increases no faster than a polynomial function of the size of the input.⁷ We take the input size to be the total number of binary digits (bits) needed to specify the input. For example, using the algorithm taught in elementary school, one can multiply two n digit numbers in a time that grows like the number of digits squared, n^2 . In contrast, the fastest-known method for the reverse operation — factoring an n -digit integer into prime numbers — takes a time that grows exponentially, roughly as 2^n . That is considered inefficient.

The class of problems that can be solved by a deterministic computer in polynomial time is represented by the capital letter P, for *polynomial* time. The class of problems that can be solved in polynomial time by a probabilistic computer is called BPP, for *bounded-error probabilistic polynomial* time. It is clear that BPP contains P, since a deterministic computation is a special case of a probabilistic computation in which we never consult the source of randomness. When we run

⁶The age of the Universe is currently estimated at 13.772 billion years.

⁷Note that the technological progress alone, such as increasing the speed of classical computers, will never turn an inefficient algorithm (exponential scaling) into an efficient one (polynomial scaling). Why?

a probabilistic (a.k.a. randomised) computation many times on the same input, we will not get the same answer every time, but the computation is useful if the probability of getting the right answer is high enough. Finally, the complexity class BQP, for *bounded-error quantum polynomial*, is the class of problems that can be solved in polynomial time by a quantum computer. Since a quantum computer can easily generate random bits and simulate a probabilistic classical computer, BQP certainly contains the class BPP. Here we are interested in problems that are in BQP but not known to be in BPP. The most popular example of such a problem is factoring. A quantum algorithm, discovered by Peter Shor in 1994, can factor n -digit numbers in a number of steps that grows only as n^2 , as opposed to the 2^n that we have classically.⁸ Since the intractability of factorisation underpins the security of many methods of encryption, Shor's algorithm was soon hailed as the first 'killer application' for quantum computation: something very useful that only a quantum computer could do. Since then, the hunt has been on for interesting things for quantum computers to do, and at the same time, for the scientific and technological advances that could allow us to build quantum computers.

1.9 Outlook

When the physics of computation was first investigated, starting in the 1960s, one of the main motivations was a fear that quantum-mechanical effects might place fundamental bounds on the accuracy with which physical objects could render the properties of the abstract entities, such as logical variables and operations, that appear in the theory of computation. It turned out, however, that quantum mechanics itself imposes no significant limits, but does break through some of those that classical physics imposed. The quantum world has a richness and intricacy that allows new practical technologies, and new kinds of knowledge. In this course we will merely scratch the surface of the rapidly developing field of quantum computation. We will concentrate mostly on the fundamental issues and skip many experimental details. However, it should be mentioned that quantum computing is a serious possibility for future generations of computing devices. At present it is not clear how and when fully-fledged quantum computers will eventually be built; but this notwithstanding, the quantum theory of computation already plays a much more fundamental role in the scheme of things than its classical predecessor did. I believe that anyone who seeks a fundamental understanding of either physics, computation or logic must incorporate its new insights into their world view.

⁸It must be stressed that not all quantum algorithms are so efficient, in fact many are no faster than their classical counterparts. Which particular problems will lend themselves to quantum speed-ups is an open question.

1.10 Notes and Exercises**1.11 Supplement: Physics against logic, via beamsplitters****1.12 Supplement: Quantum interference revisited (still about beamsplitters)**

Chapter 2

Qubits

Chapter 3

Measurements

Chapter 4

Quantum entanglement

Chapter 5

Quantum algorithms

Chapter 6

Bell's theorem

About quantum correlations, which are stronger than any correlations allowed by classical physics, and about the CHSH inequality which demonstrates this fact.

6.1 Quantum correlations

Consider two entangled qubits in the singlet state

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle)$$

and note that the projector $|\psi\rangle\langle\psi|$ can be written as¹

$$|\psi\rangle\langle\psi| = \frac{1}{4} (\text{id} \otimes \text{id} - \sigma_x \otimes \sigma_x - \sigma_y \otimes \sigma_y - \sigma_z \otimes \sigma_z).$$

Any single qubit observable with values ± 1 can be represented by the operator

$$\vec{a} \cdot \vec{\sigma} = a_x \sigma_x + a_y \sigma_y + a_z \sigma_z,$$

where \vec{a} is a unit vector in the three-dimensional Euclidean space. Suppose Alice and Bob choose measurements defined by vectors \vec{a} and \vec{b} , respectively. For example, if the two qubits are spin-half particles, they may measure the spin components along the directions \vec{a} and \vec{b} . We write the corresponding observable as the tensor product

$$A \otimes B = (\vec{a} \cdot \vec{\sigma}) \otimes (\vec{b} \cdot \vec{\sigma}).$$

¹There are other, more elementary, ways of deriving this result but here I want you to hone your skills. Now that you've learned about projectors, traces, and Pauli operators, why not put them to good use.

The eigenvalues of $A \otimes B$ are the products of eigenvalues of A and B . Thus $A \otimes B$ has two eigenvalues: $+1$, corresponding to the instances when Alice and Bob registered identical outcomes, i.e. $(+1, +1)$ or $(-1, -1)$; and -1 , corresponding to the instances when Alice and Bob registered different outcomes, i.e. $(+1, -1)$ or $(-1, +1)$. This means that the expected value of $A \otimes B$, in any state, has a simple interpretation:

$$\langle A \otimes B \rangle = \Pr(\text{outcomes are the same}) - \Pr(\text{outcomes are different}).$$

This expression can take any numerical value from -1 (perfect anti-correlations) through 0 (no correlations) to $+1$ (perfect correlations). We now evaluate the expectation value in the singlet state:

$$\begin{aligned} \langle \psi | A \otimes B | \psi \rangle &= \text{tr} [(\vec{a} \cdot \vec{\sigma}) \otimes (\vec{b} \cdot \vec{\sigma}) |\psi\rangle\langle\psi|] \\ &= -\frac{1}{4} \text{tr} [(\vec{a} \cdot \vec{\sigma}) \sigma_x \otimes (\vec{b} \cdot \vec{\sigma}) \sigma_x + (\vec{a} \cdot \vec{\sigma}) \sigma_y \otimes (\vec{b} \cdot \vec{\sigma}) \sigma_y + (\vec{a} \cdot \vec{\sigma}) \sigma_z \otimes (\vec{b} \cdot \vec{\sigma}) \sigma_z] \\ &= -\frac{1}{4} \text{tr} [(a_x b_x + a_y b_y + a_z b_z) \text{id} \otimes \text{id}] \\ &= -\vec{a} \cdot \vec{b} \end{aligned}$$

where we have used the fact that $\text{tr}(\vec{a} \cdot \vec{\sigma}) \sigma_k = a_k$ ($k = x, y, z$). So if Alice and Bob choose the same observable, $\vec{a} = \vec{b}$, then their outcomes will be always opposite: whenever Alice registers $+1$ (resp. -1) Bob is bound to register -1 (resp. $+1$).

6.2 Hidden variables

The story of “hidden variables” dates back to 1935 and grew out of Einstein’s worries about the completeness of quantum theory. Consider, for example, a qubit. No quantum state of a qubit can be an eigenstate of two non-commuting operators, say σ_x and σ_z . If the qubit has a definite value of σ_x its value of σ_z must be indeterminate, and vice versa. If we take quantum theory to be a complete description of the world, then we must accept that it is impossible for both σ_x and σ_z to have definite values for the same qubit at the same time. Einstein felt very uncomfortable about all this. He argued that quantum theory is incomplete, and that observables σ_x and σ_z may both have simultaneous definite values, although we only have knowledge of one of them at a time. This is the hypothesis of **hidden variables**. In this view, the indeterminacy found in quantum theory is merely due to our ignorance of these “hidden variables” that are present in nature but not accounted for in the theory. Einstein came up with a number of pretty good arguments for the existence of “hidden variables”. Probably the most compelling one was described in his 1935 paper (known as the EPR paper), co-authored with his younger colleagues, Boris Podolsky and Nathan Rosen. It stood for almost three decades as the most significant challenge to the completeness of quantum theory. Then, in 1964, John Bell showed that the hidden variable hypothesis can be tested and refuted.

6.3 CHSH inequality

Upper bound on *classical* correlations.

I will describe the most popular version of Bell's argument, introduced in 1969 by John Clauser, Michael Horne, Abner Shimony, and Richard Holt (CHSH). Let us assume that the results of any measurement on any individual system are predetermined. Any probabilities we may use to describe the system merely reflect our ignorance of these hidden variables.

Now, imagine the following scenario. Alice and Bob, two characters with a predilection for wacky experiments, are equipped with appropriate measuring devices and sent to two distant locations. Somewhere in between them there is a source that emits pairs of qubits that fly apart, one towards Alice and one towards Bob. Let us label the two qubits in each pair as \mathcal{A} and \mathcal{B} respectively, and let us assume that both Alice and Bob have well defined values of their observables. We ask Alice and Bob to measure one of the two pre-agreed observables. For each incoming qubit, Alice and Bob choose randomly, and independently from each other, which particular observable will be measured. Alice chooses between A_1 and A_2 , and Bob between B_1 and B_2 . Each observable has value $+1$ or -1 , and so we are allowed to think about them as random variables A_k and B_k , for $k = 1, 2$, which take values ± 1 . Let us define a new random variable, the **CHSH quantity** S , as

$$S = A_1(B_1 - B_2) + A_2(B_1 + B_2).$$

It is easy to see that one of the terms $B_1 \pm B_2$ must be equal to zero and the other to ± 2 , hence $S = \pm 2$. The average value of S must lie somewhere in-between, i.e.

$$-2 < \langle S \rangle < 2.$$

That's it! Such a simple and yet profound mathematical statement about correlations, which we refer simply to as the **CHSH inequality**. No quantum theory is involved because the CHSH inequality is not specific to quantum theory: it does not really matter what kind of physical process is behind the appearance of binary values of A_1 , A_2 , B_1 , and B_2 ; it is a statement about correlations, and for all classical correlations we must have

$$|\langle A_1 B_1 \rangle - \langle A_1 B_2 \rangle + \langle A_2 B_1 \rangle + \langle A_2 B_2 \rangle| < 2.$$

There are essentially two two assumptions here:

1. **Hidden variables.** Observables have definite values
2. **Locality.** Alice's choice of measurements (A_1 or A_2) does not affect the outcomes of Bob's measurement, and vice versa.

I will not discuss the locality assumption here in detail but let me comment on it briefly. In the hidden variable world \mathbf{a} , statement such as “if Bob were to measure B_1 then he would register $+1$ ” must be either true or false *prior to Bob's measurement*. Without the locality hypothesis, such a statement is ambiguous, since the value of B_1 could depend on whether A_1 or A_2 will be chosen by Alice. We do not want this for it implies the instantaneous communication. It means that, say, Alice by making a choice between A_1 and A_2 , affects Bob's results. Bob can immediately “see” what Alice “does”.

6.4 Quantum correlations revisited

In quantum theory the observables A_1, A_2, B_1, B_2 become 2×2 Hermitian matrices with two eigenvalues ± 1 , and $\langle S \rangle$ becomes the expected value of the 4×4 **CHSH matrix**

$$S = A_1 \otimes (B_1 - B_2) + A_2 \otimes (B_1 + B_2).$$

We can now evaluate $\langle S \rangle$ using quantum theory. For example, if the two qubits are in the singlet state, then we know that

$$\langle A \otimes B \rangle = -\vec{a} \cdot \vec{b}.$$

We choose vectors $\vec{a}_1, \vec{a}_2, \vec{b}_1$, and \vec{b}_2 as shown in **!!!TODO!!!**, and then, with these choices,

$$\begin{aligned} \langle A_1 \otimes B_1 \rangle &= \langle A_2 \otimes B_1 \rangle = \langle A_2 \otimes B_2 \rangle = \frac{1}{\sqrt{2}} \\ \langle A_1 \otimes B_2 \rangle &= -\frac{1}{\sqrt{2}}. \end{aligned}$$

Thus

$$\langle A_1 B_1 \rangle - \langle A_1 B_2 \rangle + \langle A_2 B_1 \rangle + \langle A_2 B_2 \rangle = -2\sqrt{2},$$

which obviously violates CHSH inequality. And this, to be sure, has been observed in a number of painstakingly careful experiments. The early efforts were truly heroic, and the experiments had many layers of complexity. Today, however, such experiments are routine.

Motto. The behaviour of entangled quantum systems cannot be explained by any local hidden variables.

6.5 Tsirelson's inequality

Upper bound on *quantum* correlations.

One may ask (and indeed one of you did ask) if $|\langle S \rangle| = 2\sqrt{2}$ is the maximal violation of the CHSH inequality, and yes, it is: quantum correlations cannot achieve any value of $|\langle S \rangle|$ larger than $2\sqrt{2}$. This is because, for any state $|\psi\rangle$, the expected value $\langle S \rangle = \langle \psi | S | \psi \rangle$ cannot exceed the largest eigenvalue of S , and we can put an upper bound on the largest eigenvalues of S . To start with, the largest eigenvalue (in absolute value) of a Hermitian matrix M , denoted by $\|M\|$, is a matrix norm, and it has the following properties:

$$\begin{aligned}\|M \otimes N\| &= \|M\| \|N\| \\ \|MN\| &\leq \|M\| \|N\| \\ \|M + N\| &\leq \|M\| + \|N\|\end{aligned}$$

Given that $\|A_i\| = 1$ and $\|B_j\| = 1$ ($i, j = 1, 2$), it is easy to show that $\|S\| < 4$. One can, however, derive a tighter bound. We can show (do it) that

$$S^2 = 4(\text{id} \otimes \text{id}) + [A_1, A_2] \otimes [B_1, B_2].$$

The norm of each of the commutators ($\|[A_1, A_2]\|$ and $\|[B_1, B_2]\|$) cannot exceed 2, and $\|S^2\| = \|S\|^2$, which all together gives

$$\|S\| < 2\sqrt{2} \implies |\langle S \rangle| < 2\sqrt{2}.$$

This result is known as the **Tsirelson inequality**.

Chapter 7

Decoherence, and elements of quantum error correction

Chapter 8

Density matrices

Chapter 9

Quantum channels (or CP maps)

Chapter 10

Quantum error correction and fault tolerance