On the complexity of multi-qubit quantum states

Johannes Kepler University Linz / Spring 2025

Univ. Prof. Dr. Richard Kueng, MSc ETH

Contents

1. Quantum information processing units (QPUs)

Date: April 2, 2025

This chapter is a succinct summary of my lecture notes an introduction to quantum computing. We refer to this source for a more detailed discussion of all the concepts introduced here.

Agenda:

1. high-level overview
2. single-qubit processors
3. multi-qubit processors
4. implications

1.1 High-level overview

A QPU operates on two fundamentally different levels. Input and output do correspond to conventional bit strings. However, the logic in-between is executed on extremely small scales – the realm of individual atoms and photons (particles of light). There, genuine quantum effects become available and can be used to perform completely new types of (quantum) logic. Fig. 1.1 illustrates such a setting. Throughout the course of today's lecture, we will explore the workings of such a hybrid quantum-classical architecture. We will discover that the quantum logic part is captured by a nice deterministic and even reversible

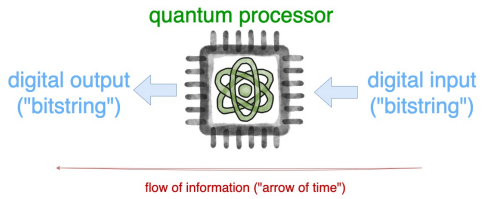


Figure 1.1 *Schematic illustration of a quantum processing unit (QPU):* on a high level, a QPU maps bitstrings to bitstrings. Also, in this class we read circuit diagrams from right to left. The red arrow underscores this convention.

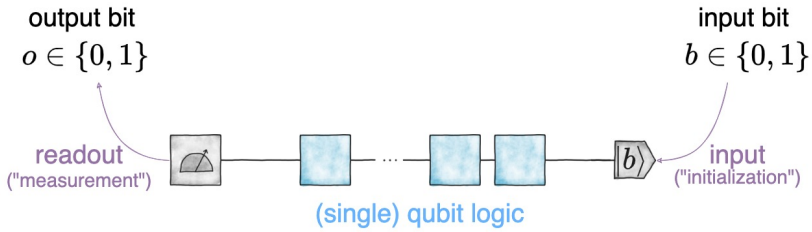


Figure 1.2 *Schematic illustration of a single-qubit processor (QPU):* input (very right) and output (very left) of a single-qubit QPU are conventional bits. Inbetween, single qubit logic (blue) is used to process the input bit directly at the quantum level. Disruptive effects happen at the quantum-classical interface (purple arrows), in particular the readout stage.

formalism. The interfaces between quantum and classical realm are more disruptive by comparison. Readout, in particular, can produce true randomness – something that is impossible for conventional (deterministic) hardware. Let us now start to discover the workings and interplay of these different constituents in a step-by-step fashion.

1.2 Single-qubit quantum processors

1.2.1 Gaining intuition

The easiest QPU only involves a single input bit, a single output bit and a single quantum bit, or *qubit*, inbetween, see Fig. 1.2 for an illustration. There are two Boolean functions that depend nontrivially on the input in question:

identity ($\mathbb{I}$) and bit flip ($\boldsymbol{X}$) gates

$$\begin{aligned} \mathbb{I}(b) &= b \quad \text{for } b \in \{0,1\} && \text{(identity)} \\ \boldsymbol{X}(b) &= \neg b \quad \text{for } b \in \{0,1\} && \text{(negation).} \end{aligned}$$

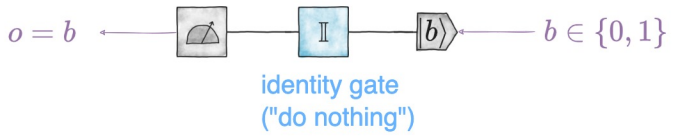
Equivalently, we can capture the action of these operation by the following 2×2 truth tables:

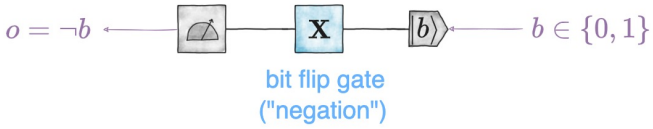
$\mathbb{I}$	0	1
0	1	0
1	0	1

$\boldsymbol{X}$	0	1
0	0	1
1	1	0

(1.1)

A single-qubit QPU can implement these basic functionalities. A full pipeline with qubit initialization, classical gate and readout looks as follows:

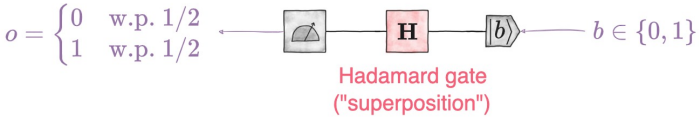




Note that these two operations are special: in contrast to other elementary logical gates (like AND and OR) they are *reversible*. They do not erase any information about the input bit. This is an essential feature of quantum logical operations. In fact, identity and bit-flip are their own inverses.

Let us now introduce our first genuine quantum operation that doesn't have a classical counterpart. The *Hadamard* or superposition gate $\boldsymbol{H}$ can apparently create true randomness: First and foremost, there is the *Hadamard* or *superposition gate*:

Hadamard/superposition ($\boldsymbol{H}$)



, (1.2)

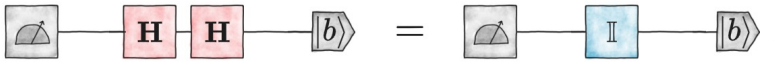
where 'w.p. $1/2$' is short for 'with probability $1/2$'. This classical-quantum-classical pipeline takes an arbitrary single-bit input b and produces a uniformly random output bit. We write

$$o \overset{unif}{\sim} \{0,1\}$$

to denote that $o = 0$ and $o = 1$ happen with equal probability $1/2$ each. This feature is a striking deviation from conventional logic which is fundamentally deterministic. The execution of a Hadamard gate uses an interesting quantum effect, called *superposition*: a binary quantum system (qubit) can assume both bit values at the same time.

If we readout (measure) such a superposition of binary values, the outcome bit we obtain is truly random: both $o = 0$ and $o = 1$ occur with equal probability. This is the same situation as a fair coin flip. The Hadamard gate provides the means to observe true randomness by bringing a qubit into equal superposition. And, equally strikingly, we can use another Hadamard gate to exit superposition again. Much like identity and bit-flip, the Hadamard gate is

also reversible. In fact, it is its own inverse as well:



(1.3)

Together, Eq. (1.2) and Eq. (1.3) reveal a striking quantum phenomenon. The first equation showcases that the Hadamard gate can be used to generate uniformly random bits. In standard binary logic, randomization requires an external seed and cannot be undone without erasing the bit in question. Or, put differently: the only way to map a random bit r into a deterministic bit o is to erase and reset. This breaks any correlations with the original input bit b. The Hadamard gate, however, is not like this at all! We can apply it twice to completely undo its effect and recover a perfect correlation between input bit b and output bit o. This is impossible in classical logic (even when we allow for true randomness).

The 'truth table' of the Hadamard gate reflects this, because it doesn't adhere to the rules of conventional logic:

	0	1
0	$1/\sqrt{2}$	$1/\sqrt{2}$
1	$1/\sqrt{2}$	$-1/\sqrt{2}$

(Hadamard 'truth table').

The detailed numbers in this table should become clear later on. For now, we emphasize two things:

- **(i)** the magnitude of each entry is the same, that is 0 and 1 feature in equal measure within the superposition;
- **(ii)** the two rows (columns) are distinct. This means that information about the input qubit is actually preserved.

1.2.2 Mathematical formalism (matrix-vector multiplication)

We now present these rules for completely describing a single-qubit quantum processor.

state of a qubit is a normalized 2D vector

Definition 1.1 (single-qubit state vector). The *state* of a single qubit keeps track of its quantum logical value. At each point, it is given by a 2-dimensional vector, normalized to unit length:

$$|\psi\rangle := \boldsymbol{\psi} = \begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix} \in \mathbb{C}^2 \text{ s.t. } \langle \boldsymbol{\psi}, \boldsymbol{\psi} \rangle = \boldsymbol{\psi}^* \boldsymbol{\psi} = |\psi_0|^2 + |\psi_1|^2 = 1. \tag{1.4}$$

The somewhat strange notation $|\psi\rangle$ to denote state vectors $\boldsymbol{\psi}$ is called a *ket* and features prominently in the quantum computing literature.

Example 1.2 (Qubit initialization). The following two state vectors

$$|0\rangle := \boldsymbol{e}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |1\rangle := \boldsymbol{e}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

are valid state vectors that tell us how to imprint a classical bit $b \in \{0,1\}$ into the state vector of a qubit. ∎

As Definition 1.1 suggests, state vectors can be used to keep track of qubits throughout a sequence of quantum logical gates. This, however, necessitates a formalism to unambiguously characterize the action of quantum gates. And a way to imprint their action onto the current state vector of a qubit.

Definition 1.3 (single-qubit gate action). A single-qubit gate is fully described by a unitary 2×2 matrix

single-qubit gates are unitary 2×2 matrices ('truth tables')

$$\boldsymbol{U} = \begin{pmatrix} U_{0,0} & U_{0,1} \\ U_{1,0} & U_{1,1} \end{pmatrix} \in \mathbb{C}^{2\times 2} \text{ s.t. } \boldsymbol{U}^*\boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^* = \mathbb{I}.$$

The action of gate $\boldsymbol{U}$ on quantum state $|\psi\rangle = \boldsymbol{\psi} \in \mathbb{C}^2$ is captured by matrix-vector multiplication :

gate action on qubit state = matrix-vector multiplication

$$|\psi_{\text{out}}\rangle = \boldsymbol{\psi}_{\text{out}} = \boldsymbol{U}\boldsymbol{\psi}_{\text{in}} = \boldsymbol{U}|\psi_{\text{in}}\rangle.$$

The following instructive example showcases how this matrix-vector multiplication formalism allows us to recover classical logical operations.

Example 1.4 (matrix representation of classical gates). The matrix representations of the classical operations *identity* ($\mathbb{I}$) and *bit-flip* ($\boldsymbol{X}$) are in one-to-one correspondence with their truth tables:

$$\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \boldsymbol{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

It is easy to check that both matrices are unitary matrices. What is more,

$$\begin{aligned} \mathbb{I}|0\rangle &= \mathbb{I}\boldsymbol{e}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \boldsymbol{e}_0 = |0\rangle, \\ \mathbb{I}|1\rangle &= \mathbb{I}\boldsymbol{e}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \boldsymbol{e}_1 = |1\rangle, \end{aligned}$$

which puts 'do nothing' into concrete formulas. Likewise

$$\begin{aligned} \boldsymbol{X}|0\rangle &= \boldsymbol{X}\boldsymbol{e}_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \boldsymbol{e}_1 = |1\rangle, \\ \boldsymbol{X}|1\rangle &= \boldsymbol{X}\boldsymbol{e}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \boldsymbol{e}_0 = |0\rangle, \end{aligned}$$

puts formulas to the action of a bit-flip. It is not a coincidence that these matrix representations are in one-to-one correspondence to the truth tables of the corresponding logical functionalities. Matrix-vector multiplication is just another way of reading logical truth tables. ∎

Exercise 1.5 (matrix representation of the Hadamard gate). The matrix representation of the Hadmard gate is given as

$$\boldsymbol{H} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \in \mathbb{C}^{2\times 2}.$$

1 Show that this matrix is unitary by verifying $\boldsymbol{U}^*\boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^* = \mathbb{I}$.
2 The action of a Hadamard gate maps deterministic bit states $|0\rangle$ and $|1\rangle$ into uniform superposition states $|+\rangle$ and $|-\rangle$. Use matrix-vector multiplication to verify the following state vector representations:

$$|+\rangle = \boldsymbol{H}|0\rangle = \boldsymbol{H}\boldsymbol{e}_0 = \frac{1}{\sqrt{2}}(\boldsymbol{e}_0 + \boldsymbol{e}_1) = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \tag{1.5}$$

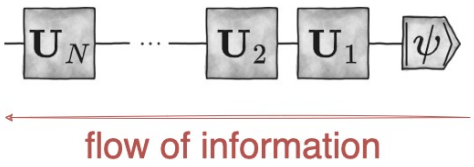
$$|-\rangle = \boldsymbol{H}|1\rangle = \boldsymbol{H}\boldsymbol{e}_1 = \frac{1}{\sqrt{2}}(\boldsymbol{e}_0 - \boldsymbol{e}_1) = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \tag{1.6}$$

Both expressions on the very right should be interpreted as 'both 0 and 1 in equal measure'. Note, however that one combination has a '+' inbetween, while the other one has a '−'. This sign difference ensures that both superpositions remain distinct and can be undone again.
3 Verify reversibility by showing $\boldsymbol{H}^2 = \mathbb{I}$ via matrix-vector multiplication.

Suppose that we sequentially apply N single-qubit gates $\boldsymbol{U}_1, \ldots, \boldsymbol{U}_N$ to an arbitrary starting state (vector) $|\psi\rangle = \boldsymbol{\psi} \in \mathbb{C}^2$. Then, we can compute the final state vector as

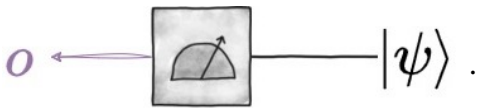
$$|\psi_{\text{final}}\rangle = \boldsymbol{\psi}_{\text{final}} = \boldsymbol{U}_N \times \cdots \times \boldsymbol{U}_2 \times \boldsymbol{U}_1 \boldsymbol{\psi} \in \mathbb{C}^2.$$

In words: we start our matrix-vector multiplication on the very right and keep going. The convention to read circuit diagrams from left to right as well exactly resembles this ordering:



. (1.7)

We now have all the pieces in place to initialize a qubit and keep track of its state throughout a sequence of arbitrary many single qubit gates. All that is missing now is a formula for executing the readout at the very end. That is, we need to assign meaning to the following operation:



.

Definition 1.6 (single-qubit readout). A single-qubit readout (measurement) operation always produces a valid bit value $o \in \{0, 1\}$. But it does so probabilistically. The probability of obtaining outcome $o \in \{0, 1\}$ depends on the underlying state vector $|\psi\rangle = \boldsymbol{\psi} \in \mathbb{C}^2$:

outcome probabilities = squared magnitudes of state vector entries

$$p_0 = \Pr_{|\psi\rangle}[o = 0] = |\psi_0|^2 \geq 0 \text{ and } p_1 = \Pr_{|\psi\rangle}[o = 1] = |\psi_1|^2 \geq 0. \tag{1.8}$$

This should be read as: 'the probability of obtaining outcome $o = 0$ ($o = 1$) when reading out a qubit in state $|\psi\rangle$ is $|\psi_0|^2$ ($|\psi_1|^2$).

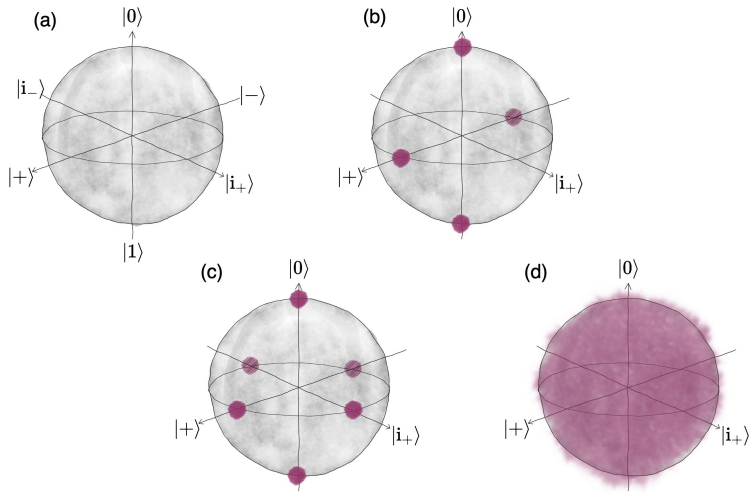


Figure 1.3 *A visualization of single-qubit state vectors reached by elementary gate combinations:* (a) the Bloch sphere representation of all single-qubit state vectors (b) The single qubit quantum states reached with only combinations of $\boldsymbol{X}$ and $\boldsymbol{H}$; (c) The single qubit quantum states reached with only Clifford gates; (d) The single qubit quantum states reached with arbitrary long combinations of Clifford gates and the $\boldsymbol{T}$ gate as stated in Theorem 1.7.

Note that normalization of the state vector (Eq. (1.4) in Defininition 1.1) ensures that p_0 and p_1 define a valid binary probability distribution (think: coin toss): $p_0, p_1 \geq 0$ and

$$p_0 + p_1 = |\psi_0|^2 + |\psi_1|^2 = \|\boldsymbol{\psi}\|^2 = 1.$$

1.2.3 Discrete gate sets and universality

Note that we have phrased the mathematical formalism in an exceedingly general fashion: a state vector can by *any* complex-valued 2D vector and a circuit can be any 2×2 unitary matrix. In stark contrast, the concrete examples of quantum gates we have seen so far – bit flip ($\boldsymbol{X}$) and the Hadamard gate ($\boldsymbol{H}$) – are still very discrete in nature. Here, we analyze this apparent gap between continuous degrees of freedom and discrete quantum gates.

In fact, combining these two gates only yields 4 functionally distinct gates[1]:

$$\mathbb{I}, \boldsymbol{H}, \boldsymbol{X}, \boldsymbol{Z} = \boldsymbol{HXH} \quad \text{and} \quad \boldsymbol{XZ}.$$

In order to get more functionalities, we need additional gates. The *phase gate*

$$\boldsymbol{S} = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{i} \end{pmatrix}$$

[1]Note that global sign differences cannot affect the functionality of a gate, because signs/phases are absorbed by the readout rule

is one such gate that introduces complex numbers into the formalism. Together, $\boldsymbol{H}$ and $\boldsymbol{S}$ can produce a total of 24 functionally distinct gates known as Clifford gates. These include all Pauli gates, e.g. $\boldsymbol{X} = \boldsymbol{H}\boldsymbol{S}^2\boldsymbol{H}$. A more substantial increase in expressiveness occurs if we replace the phase gate ($\boldsymbol{S}$) with the *T-gate*:

$$\boldsymbol{T} = \begin{pmatrix} 1 & 0 \\ 0 & \exp(\mathrm{i}\pi/4) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & (1+\mathrm{i})/\sqrt{2} \end{pmatrix}.$$

Note that the T-gate can be thought of as the 'square root' of the phase gate: $\boldsymbol{T}^2 = \boldsymbol{S}$. Hence, replacing $\boldsymbol{S}$ by $\boldsymbol{T}$ can only increase the number of quantum gates that can be reached by Hadamard and T. The actual gain is astonishing.

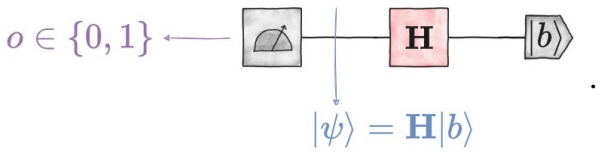
Theorem 1.7 (universal gate set). Together, the elementary Clifford gates $\boldsymbol{H}$, $\boldsymbol{S}$ and the T-gate $\boldsymbol{T}$ form a *universal gate set*: Every 2×2 unitary matrix can be approximated to an arbitrary degree with sequences comprised of *only* $\boldsymbol{H}$ and $\boldsymbol{T}$ (we actually don't need $\boldsymbol{S}$, because $\boldsymbol{S} = \boldsymbol{T}^2$).

Hadamard+T generate every 2×2 unitary matrix

This powerful statement stems from group theory and its implications are visualized in Figure 1.3. Remarkably, the actual cost of approximating a target unitary $\boldsymbol{U}$ up to accuracy ϵ only scales poly-logarithmically in the desired accuracy. This result, known as Solovay-Kitaev Theorem, ensures that the approximation error of a quantum circuit approximation diminishes exponentially with the number of elementary quantum gates one is willing to invest.

1.2.4 Example: random number generator

Consider the following quantum circuit which we read from left to right:



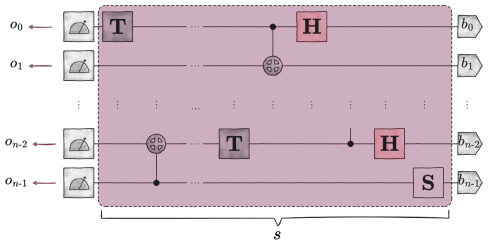
Let us do the computation for $b = 1$ (the case for $b = 0$ is similar). We start by using matrix-vector multiplication to compute the final state vector (blue):

$$|\psi\rangle = \begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix} = \boldsymbol{H}|1\rangle = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}.$$

Hence, $\psi_0 = 1/\sqrt{2}$, $\psi_1 = -1/\sqrt{2}$ and perfect randomness generation follows from invoking the rule for single-qubit readout (Definition 1.6):

$$\Pr_{|\psi\rangle}[o = 0] = |\langle 0|\psi\rangle|^2 = |\psi_0|^2 = \left|1/\sqrt{2}\right|^2 = 1/2,$$
$$\Pr_{|\psi\rangle}[o = 1] = |\langle 1|\psi\rangle|^2 = |\psi_1|^2 = \left|-1/\sqrt{2}\right|^2 = 1/2.$$

This is just the definition of a perfect random bit $o \overset{unif}{\sim} \{0,1\}$.



n-qubit architecture: n input qubits, n readout bits and a combination of elementary quantum gates inbetween

Figure 1.4 *A general n-qubit architecture* has n input bits $b_0, \ldots, b_{n-1}$, a central block of quantum logic and a final readout stage that recovers n bits $o_0, \ldots, o_{n-1}$. the central block is solely comprised of elementary quantum gates, e.g. $\boldsymbol{H}, \boldsymbol{T}, \boldsymbol{S}$ and $\boldsymbol{CNOT}$. The number s of elementary quantum gates is called the *size* of the quantum circuit.

1.3 Multi-qubit quantum processors

The single-qubit formalism introduced above readily extends to n-qubit QPUs. Visualized in Fig. 1.4, such an architecture features a n-bit initialization stage (left), a central block of reversible quantum logic and a n-bit readout stage (right). Readout and initialization can give rise to one out of $D = 2^n$ different bitstring configurations:

$$|\boldsymbol{b}\rangle = |b_0 \cdots b_{n-1}\rangle \quad \text{with} \quad b_0, \ldots, b_{n-1} \in \{0, 1\}\,.$$

A general quantum state vector assigns *amplitudes* to each of these bit configurations:

n-qubit state vector has 2^n complex-valued amplitudes

$$|\psi\rangle = \sum\nolimits_{b_0,\cdots,b_{n-1}=0}^{1} \psi_{b_0\cdots b_{n-1}} |b_0 \cdots b_{n-1}\rangle = \boldsymbol{\psi} = \begin{pmatrix} \psi_{0\cdots 00} \\ \psi_{0\cdots 01} \\ \vdots \\ \psi_{1\cdots 11} \end{pmatrix} \in \mathbb{C}^{2^n}.$$

The normalization condition still requires that the sum of squared absolute values adds up to one:

$$\|\boldsymbol{\psi}\|_2^2 = \boldsymbol{\psi}^* \boldsymbol{\psi} = \sum\nolimits_{b_0,\cdots,b_{n-1}=0}^{1} \left|\psi_{b_0\cdots b_{n-1}}\right|^2 = 1.$$

These squared absolute values of the 2^n possible bit string configurations also tell us the probability of obtaining one of these bitstrings when performing a readout operation on all n qubits:

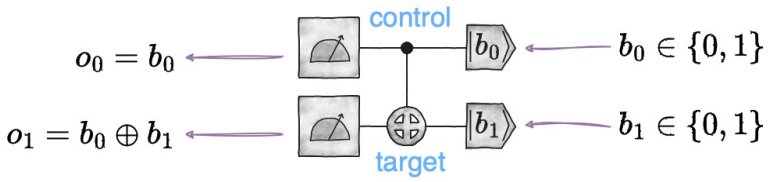
Born's rule: n-qubit readout produces a single n-bit string, possibly in random fashion

$$\boldsymbol{o} = \boldsymbol{v} = v_0 \cdots v_{n-1} \text{ occurs with } \Pr_{|\psi\rangle}[\boldsymbol{o} = \boldsymbol{v}] = \left|\psi_{v_0\cdots v_{n-1}}\right|^2.$$

This is known as *Born's rule*. In words: the probability of observing a certain outcome bitstring $\boldsymbol{v}$ is given by the squared magnitude of the amplitude $\psi_{\boldsymbol{v}}$. The absolute values ensure nonnegativity, while the state normalization condition ensures that the sum of all these probabilities is indeed one. So, this is a valid probability distribution over 2^n different outcomes.

Likewise, a general n-qubit circuit is described by a $2^n \times 2^n$ unitary matrix $\boldsymbol{U}$. Similar to classical circuits, we obtain such functionalities by combining elementary gates. We already know that Hadamard ($\boldsymbol{H}$) and T-gate ($\boldsymbol{T}$) comprise a universal gate set for single qubit circuits. For multi-qubit functionalities, we also need at least one conditional two-qubit gate. A popular choice is the *CNOT-gate* which corresponds to a reversible realization of an XOR operation:

CNOT is a controlled-not gate on two qubits



. (1.9)

Note that the choice of control qubit (solid circle) and target qubit (cross) matter and we actually obtain two functionally different CNOT gates with the following 4× matrix (truth table) representations:

$$\boldsymbol{CNOT}_{0\to1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \boldsymbol{CNOT}_{1\to0} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{1.10}$$

Together, Hadamard ($\boldsymbol{H}$), T-gate ($\boldsymbol{T}$) and CNOT gates between any qubit pair can be used to construct more expressive n-qubit circuits. This has been visualized in the central block of Fig. 1.4. The resulting $2^n \times 2^n$ matrix $\boldsymbol{U}$ can be built up from the matrix descriptions of these elementary gates by using different types of matrix products:

1. *Parallel gate applications use the Kronecker product* '$\otimes$' of the individual gate matrices involved (including $\mathbb{I} \in \mathbb{C}^{2\times2}$ for qubit wires where nothing happens). For n qubits, this always produces a single $2^n \times 2^n$ matrix for each gate layer.
2. Combining *sequential gate layers uses the matrix product* '$\times$' of $2^n \times 2^n$ gate layer matrices.

Note that both Kronecker and matrix produce preserve unitarity. Hence, any n-qubit circuit comprised of only elementary quantum gates produces a big $2^n \times 2^n$ circuit matrix $\boldsymbol{U}$ that is unitary. Hence, every quantum circuit is a valid n-qubit time evolution. Remarkably, the converse is also true.

Theorem 1.8 (Solovay-Kitaev (n-qubit case)). Every unitary $2^n \times 2^n$ matrix $\boldsymbol{U}$ can be approximated to arbitrary precision by a n-qubit quantum circuit that is solely comprised of $\boldsymbol{H}, \boldsymbol{T}, \boldsymbol{CNOT}$ (Clifford) and $\boldsymbol{T}$-gates.

every $2^n \times 2^n$ unitary matrix admits a n-qubit circuit approx. (Solovay-Kitaev)

In words: every $2^n \times 2^n$ unitary matrix can be approximated to arbitrary precision by a n-qubit quantum circuit comprised of only elementary gates.

This is conceptually very interesting, but far from trivial. For instance, given a unitary matrix $\boldsymbol{U}$ it is a challenging problem to find such an approximation in the first place (transpilation). Worse, most $2^n \times 2^n$ unitaries require exponentially deep circuits to approximate. This, and related questions, will be the content of the next lectures.

1.4 Implications

Let us conclude this brief introductory lecture with a couple of general comments and remarks.

- Every classical circuit can be executed on a QPU with (at most) linear overhead in the number of qubits and the number of quantum gates. This is achieved by using additional qubits to implement AND and OR in a reversible fashion.
- Every quantum circuit can be simulated on a classical computer using matrix-vector multiplication. This incurs (at most) an exponential overhead. In turn, quantum computers have no impact on computability, only computational complexity.
- The best example of this form is Shor's algorithm for factoring a n-bit product of two distinct primes $N = p \times q$. This is actually a hybrid quantum-classical algorithm that first reduces this factoring problem to the problem of finding the order of a number in modular arithmetic. A quantum circuit comprised of $O(n^3)$ can find this order with probability at least $2/3$, while no efficient (randomized) classical algorithm is known.
- The fact that any $2^n \times 2^n$ unitary matrix $\boldsymbol{U}$ can be approximated by a digital quantum circuit is also relevant for simulating quantum physics. The time-dependent Schroedinger equation of a closed, D-dimensional system gives rise to a family of $D \times D$ unitary matrices $\boldsymbol{U}(t) = \exp(-\mathrm{i}t\boldsymbol{H})$ parametrized by the evolution time t. Theorem 1.8 assures that a QPU comprised of $n = \lceil \log_2(D) \rceil$ qubits can, in principle, approximate any such unitary to arbitrary precision.

2. Complexity of generic states/circuits

Date: 3 April 2025

2.1 Motivation

Agenda:

1. motivation
2. main result
3. proof:
   - expected behavior
   - concentration
   - union bound

Last lecture, we introduced the formalism of QPUs with an emphasiz on n-qubit QPUs. We found some parallels to classical n-bit circuits, like elementary gate sets and universality, but there are also some notable differences: quantum circuits are reversible and can explore continuous degrees of freedom. Also, the readout stage is very different from classical hardware.

Today we start asking some fundamental questions about the *complexity* of quantum circuits and also states. Roughly speaking, the complexity of a quantum circuit is the minimal cost required to accurately approximate it with elementary gates. For Boolean circuits, a famous counting argument by Shannon shows that almost all Boolean functions require circuits of exponential size to accurately represent. Or, in slightly different words: a generic Boolean function has exponentially large complexity in the sense that it cannot be realized by a circuit of sub-exponential size.

Today, we showcase how to generalize this type of results to the quantum case. To keep things simple, we won't focus on the complexity of a n-qubit quantum circuit, but on the complexity of a n-qubit quantum state. Roughly speaking, the complexity of a quantum state $\boldsymbol{\psi} = |\psi\rangle$ is the minimal cost of a circuit that takes a designated starting state, e.g. $\boldsymbol{\psi}_0 = |\psi_0\rangle = |0\cdots0\rangle$ and (approximately) maps it to the n-qubit state vector $\boldsymbol{\psi}$. Note that the classical counterpart of this state complexity is not particularly interesting. The complexity of any n-bit string is at most n. Today, we will prove that the quantum case is very different. In fact, almost all n-qubit state vectors have a state complexity that scales exponentially in n.

2.2 Setting and main result

Recall that every n-qubit quantum state is fully described by a $D = 2^n$-dimensional state vector $\boldsymbol{\psi} = |\psi\rangle \in \mathbb{C}^D$ with complex-valued entries and Euclidean unit length: $\langle \boldsymbol{\psi}, \boldsymbol{\psi} \rangle = \boldsymbol{\psi}^* \boldsymbol{\psi} = 1$. So, geometrically speaking, a state vector is an element of an exponentially large complex unit sphere $\mathbb{S}^{D-1}$ with $D = 2^n$. Every state vector is a point on that sphere and, conversely, any point on that sphere corresponds to a state vector. A natural distance measure between quantum state vectors is the *fidelity*:

fidelity between two quantum states

$$F(\boldsymbol{\psi}, \boldsymbol{\varphi}) = |\langle \boldsymbol{\psi}, \boldsymbol{\varphi} \rangle|^2 = |\boldsymbol{\psi}^* \boldsymbol{\varphi}|^2 \in [0, 1].$$

The fidelity is 1 if and only if the two state vectors are the same (up to a complex phase) and 0 if the state vectors are orthogonal to each other. So, the fidelity is large if and only if two state vectors are (very) similar and the fidelity is small else if the two state vectors are very distinct from each other.

To define state complexity, we first fix a universal gate set, e.g. $\boldsymbol{H}, \boldsymbol{T}$ and $\boldsymbol{CNOT}$ between any pair of qubits (and in any direction). Next, we let V_R denote the set of all n-qubit circuits that can be generated with R such gates or fewer. In other words: V_R is the set of all quantum circuits of size (at most) R. Informally speaking, we say that a n-qubit state vector has state complexity at most R if there exists a circuit $\boldsymbol{V} \in \mathsf{V}_R$ that approximately generates it.

formal definition of state complexity

Definition 2.1 (state complexity). Let V_R be the set of all size-R circuits on n qubits with gates from a universal gate set (e.g. Hadamard $\boldsymbol{H}$, T-gate $\boldsymbol{T}$ and CNOT gates), let $|\psi_0\rangle = \boldsymbol{\psi}_0$ be a fixed 'simple' state. For $\delta \in (0, 1)$, we say that $|\psi\rangle = \boldsymbol{\psi} \in \mathbb{C}^D$ has δ-complexity at most R if

$$\max_{\boldsymbol{V} \in \mathsf{V}_R} F\left(\boldsymbol{\psi}, \boldsymbol{V}\boldsymbol{\psi}_0\right) \geq 1 - \delta^2.$$

If this is the case, we write $C_\delta(|\psi\rangle) \leq R$.

The main result of today addresses the state complexity of a 'generic' n-qubit state vector. the geometry of the $D = 2^n$-dimensional unit sphere provides us with a uniform measure from which we can sample these state vectors. It's the unique unitarily invariant measure on $\mathbb{S}^{D-1}$ that assigns the same infinitesimal probability weight to every possible unit vector. This measure is sometimes also called the *Haar measure*. Today we will prove the following result.

Theorem 2.2 (generic states have exponentially high state complexity). A Haar-random state $\boldsymbol{h}|h\rangle \overset{unif}{\sim} \mathbb{S}^{D-1}$ obeys

$$\Pr_{\boldsymbol{h} \overset{unif}{\sim} \mathbb{S}^{D-1}} \left[C_\delta(|h\rangle) \leq R\right] \leq (2n)^{2R} \mathrm{e}^{-D(1-\delta^2)/2} \quad \text{for any } R \in \mathbb{N}_+.$$

This probability remains tiny until

$$R \approx \frac{D(1-\delta^2)}{4\log(2n)} = \frac{2^n(1-\delta)}{4\log(2n)} = \Omega\left(2^n/\log(n)\right).$$

Note that the Haar measure is fair in the sense that it assigns the same infinitesimal weight to each D-dimensional quantum state. Laplace's definition of probability therefore allows us to interpret Eq. (2.2) as a bound on the relative volume of complexity-R states. This volume remains tiny until R approaches the overall Hilbert space dimension $D = 2^n$. In other words: *almost all states have almost maximal state complexity*.

almost all states have almost maximal state complexity

2.3 Proof of the main result

We use a general randomized proof technique to establish Theorem 2.2 that works in three steps: (i) analyzing the expected behavior, (ii) establish concentration around this expectation and (iii) apply a union bound over all possible instances. This proof techniques is very versatile and may be of independent interest.

2.3.1 Step 1: Expected behavior

Let us start by computing the expected fidelity between a Haar-random state vector and a fixed target state vector.

Lemma 2.3 Let $|v\rangle = \boldsymbol{v} \in \mathbb{C}^D$ be a fixed unit vector (state). Then, a random vector $|h\rangle = \boldsymbol{h} \overset{unif}{\sim} \mathbb{S}^{D-1}$ obeys

$$\mathbb{E}_{\boldsymbol{h} \overset{unif}{\sim} \mathbb{S}^{D-1}} \left[F\left(|v\rangle, |h\rangle\right)\right] = \mathbb{E}_{\boldsymbol{h} \overset{unif}{\sim} \mathbb{S}^{D-1}} \left[|\langle \boldsymbol{v}, \boldsymbol{h}\rangle|^2\right] = \frac{1}{D}.$$

In words: a uniformly random state vector is almost maximally far away of any fixed unit vector in expectation.

Note that it is possible to compute this expectation value exactly using Gaussian integration. This strategy works in two parts: (i) use unitary invariance of the distribution of $\boldsymbol{h}$ to repace $\boldsymbol{v}$ with a simpler vector, e.g. $\boldsymbol{e}_0 = (10\cdots)^*$; (ii) replace $\boldsymbol{h}$ by a standard complex Gaussian vector and use the fact that the length (nonnegative number) and direction (unit vector) of a standard Gaussian vector are statistically independent to reduce the remaining task to two independent Gaussian integrals that are easy to solve (one for the real part and one for the complex part). While Gaussian integration is powerful, it does require quite a bit of dedication to carry out successfully. Let us now present an alternative formalism that has become very popular in quantum information over the past years. Additional details and exposition can be found in my Caltech lecture notes on Quantum and Classical Information Processing with Tensors.

Fact 2.4 (Haar integration; folklore). Let $|h\rangle = \boldsymbol{h} \overset{Haar}{\sim} \mathbb{C}^D$ ($D = d^n$) be a Haar random state vector. Then, for all $k \in \mathbb{N}_+$ Haar integration formula

$$\mathbb{E}_{\boldsymbol{h} \overset{unif}{\sim} \mathbb{S}^{D-1}} \left[(\boldsymbol{h}\boldsymbol{h}^*)^{\otimes k} \right] = \int_{\mathbb{S}^{D-1}} (\boldsymbol{h}\boldsymbol{h}^*)^{\otimes k} \, \mathrm{d}\mu(h) = \binom{D+k-1}{k}^{-1} \boldsymbol{P}_{\vee^k}, \quad (2.1)$$

where $P_{\vee k}$ is the projector onto the totally symmetric subspace of $\left(\mathbb{C}^D\right)^{\otimes k}$. ∎

This formula readily allows us to compute the expected fidelity of a Haar random state vector.

Proof of Lemma 2.3. The key observation is that the squared modulus becomes linear in the outer products of the vectors: $|\langle \boldsymbol{v}, \boldsymbol{h}\rangle|^2 = |\boldsymbol{v}^*\boldsymbol{h}|^2 = \mathrm{trace}\,(\boldsymbol{v}\boldsymbol{v}^*\boldsymbol{h}\boldsymbol{h}^*)$. This allows us to employ Fact 2.4 for $k = 1$ and $\boldsymbol{P}_{\vee^1} = \mathbb{I}$:

$$\begin{aligned} \mathbb{E}_{\boldsymbol{h}} \left[|\boldsymbol{v}, \boldsymbol{h}\rangle|^2 \right] =& \mathbb{E}_{\boldsymbol{h}} \left[\mathrm{tr}\,(\boldsymbol{v}\boldsymbol{v}^* \ \boldsymbol{h}\boldsymbol{h}^*) \right] = \mathrm{trace}\,(\boldsymbol{v}\boldsymbol{v}^* \mathbb{E}_{\boldsymbol{h}} \left[\boldsymbol{h}\boldsymbol{h}^* \right]) \\ =& \mathrm{trace} \left(\boldsymbol{v}\boldsymbol{v}^* \frac{1}{D} \mathbb{I} \right) = \frac{1}{D} \boldsymbol{v}^* \mathbb{I} \boldsymbol{v} = \frac{1}{D}. \end{aligned}$$

■

2.3.2 Step 2: Concentration around expected behavior

High-dimensional probability theory tells us that concrete realizations of the uniform state vector $\boldsymbol{h} \in \mathbb{S}^{D-1}$ will concentrate sharply around the expected behavior we just computed. Concentration of measure, exemplified by Levy's Lemma, tells us that

$$\Pr_{\boldsymbol{h} \overset{unif}{\sim} \mathbb{S}^{D-1}} \left[\left| |\boldsymbol{v}^*\boldsymbol{h}|^2 - 1/D \right| \geq \tau \right] \leq 2\mathrm{e}^{-D\tau^2/(9\pi^3)} \quad \text{for any } \tau > 0.$$

In words: the probability that a randomly selected $\boldsymbol{h}$ is not maximally far away from a fixed reference state vector $\boldsymbol{h}$ diminishes exponentially in D. And $D = 2^n$ already scales exponentially in the number of qubits. This, in turn, ensures that the probability of a random state vector being close in fidelity to an arbitrary fixed reference state is doubly-exponentially suppressed in the number of qubits. Let us prove a weaker variant of such a statement directly using the Haar integration result from above.

Lemma 2.5 Let $\boldsymbol{v} \in \mathbb{C}^D$ ($D = 2^n$) be a fixed unit vector (n-qubit state state). Then, a random vector $\boldsymbol{h} \overset{unif}{\sim} \mathbb{S}^{D-1}$ obeys concentration of measure for one fixed fidelity

$$\Pr \left[|\boldsymbol{v}^*\boldsymbol{h}|^2 \geq (1-\delta) \right] \leq 2\mathrm{e}^{-D(1-\delta)/2} \quad \text{for any } \delta \in (0,1).$$

Proof. Let us start by computing the moments of the random variable $F(\boldsymbol{v}, \boldsymbol{h}) =$

$|\boldsymbol{v}^*\boldsymbol{h}|^2$. For any $k \in \mathbb{N}_+$ the Haar integration formula (Fact 2.4) yields

$$\begin{aligned}\mathbb{E}_{|h\rangle}\left[|\boldsymbol{v}^*\boldsymbol{h}|^{2k}\right] =& \mathrm{tr}\left((\boldsymbol{v}\boldsymbol{v}^*)^{\otimes k}\,\mathbb{E}_{\boldsymbol{h}}\left[(\boldsymbol{h}\boldsymbol{h}^*)^{\otimes k}\right]\right)\\ =& \mathrm{tr}\left((\boldsymbol{v}\boldsymbol{v}^*)^{\otimes k}\binom{D+k-1}{k}^{-1}\boldsymbol{P}_{\vee^k}\right)\\ =& \binom{D+k-1}{k} \leq \frac{k!}{D^k}.\end{aligned} \tag{2.2}$$

This moment behavior indicates sub-exponential moment growth. We can now use some elementary tricks from probability theory to turn these moment bounds into an exponential concentration bound:

$$\begin{aligned}\Pr_{\boldsymbol{h}}\left[|\boldsymbol{v}^*\boldsymbol{h}|^2 \geq \tau\right] =& \Pr_{\boldsymbol{h}}\left[D|\boldsymbol{v}^*\boldsymbol{h}|^2/2 \geq D\tau/2\right]\\ =& \Pr_{|h\rangle}\left[\exp\left(D|\boldsymbol{v}^*\boldsymbol{h}|^2/2\right) \geq \exp\left(D\tau/2\right)\right]\\ \leq& \mathrm{e}^{-D\tau/2}\mathbb{E}\left[\exp\left(D|\boldsymbol{v}^*\boldsymbol{h}|^2/2\right)\right]\\ =& \mathrm{e}^{-D\tau/2}\sum_{k=0}^{\infty}\frac{1}{k!}\frac{D^k}{2^k}\mathbb{E}_{\boldsymbol{h}}\left[|\boldsymbol{v}^*\boldsymbol{h}|^{2k}\right]\\ \leq& \mathrm{e}^{-D\tau/2}\sum_{k=0}^{\infty}\frac{1}{2^k} = 2\mathrm{e}^{-D\tau/2}.\end{aligned}$$

The key step is Markov's inequality ($\Pr[S \geq \alpha] \leq \mathbb{E}[S]/\alpha$ for any nonnegative random variable S) in line three. ■

2.3.3 Step 3: union bound

We are now in a position to address quantum state complexity of a randomly drawn state vector $\boldsymbol{h}$. Recall that the state complexity in Definition 3.2 is implicitly defined by a maximization over (a lot of) candidate circuits $\boldsymbol{V}$ that map a designated starting state vector $\boldsymbol{\psi}_0$ to $\boldsymbol{v} = \boldsymbol{V}\boldsymbol{\psi}_0$. And we are looking for the circuit that is closest to $\boldsymbol{h}$ in fidelity. Our concentration result from above states that a high fidelity is doubly-exponentially unlikely for any fixed candidate $\boldsymbol{v}$. Hence, we expect to need a large amount of candidate state vectors $\boldsymbol{v}$ to counteract this tail suppression. The main result of this lecture follows from making this intuition precise.

Theorem 2.6 (Detailed restatement of Theorem 2.2). Let $C_\delta(\boldsymbol{h})$ with accuracy $\delta \in (0,1)$ denote the state complexity of a n-qubit state with respect to the elementary gate set $\boldsymbol{H}, \boldsymbol{T}$ and $\boldsymbol{CNOT}$. Then, a uniformly random n-qubit state vector $\boldsymbol{h} \sim \mathbb{S}^{D-1}$ is guaranteed to obey

$$\Pr_{\boldsymbol{h}\overset{unif}{\sim}\mathbb{S}^{D-1}}\left[C_\delta(\boldsymbol{h}) \leq R\right] \leq 2\,(2n)^{2R}\,\mathrm{e}^{-D(1-\delta^2)/2} \quad \text{for any } R \in \mathbb{N}.$$

This probability remains tiny until

$$R \approx \frac{D(1-\delta)}{4\log(2n)} = \frac{2^n(1-\delta)}{4\log(2n)}.$$

Proof. Insert the definition of state complexity, see Definition 3.2, and apply a union bound (Boole's inequality) to obtain

$$\begin{aligned}
\Pr_{\boldsymbol{h}\overset{unif}{\sim}\mathbb{S}^{D-1}}\left[C_\delta(\boldsymbol{h}) \leq R\right] &= \Pr_{\boldsymbol{h}\overset{unif}{\sim}\mathbb{S}^{D-1}}\left[\max_{\boldsymbol{V}\in\mathsf{V}_R}\left|\left(\boldsymbol{V}\boldsymbol{\psi}_0\right)^*\boldsymbol{h}\right|^2 \geq 1-\delta\right] \\
&\leq \sum\nolimits_{\boldsymbol{V}\in\mathsf{V}_R}\Pr_{\boldsymbol{h}\overset{unif}{\sim}\mathbb{S}^{D-1}}\left[\left|\left(\boldsymbol{V}\boldsymbol{\psi}_0\right)^*\boldsymbol{h}\right|^2 \geq 1-\delta\right] \\
&\leq \sum\nolimits_{\boldsymbol{V}\in\mathsf{V}_R} 2\mathrm{e}^{-D(1-\delta)/2} = |\mathsf{V}_R|\, 2\mathrm{e}^{-D(1-\delta)/2}.
\end{aligned}$$

The last inequality is curtesy of Lemma 2.5 which is valid for any $\boldsymbol{v} = \boldsymbol{V}\boldsymbol{\psi}_0$.

The claim now follows from a simple counting argument. There are at most $(4n^2)^R$ different n-qubits that one can construct by sampling R gates from the collection $\mathbb{I}, \boldsymbol{H}, \boldsymbol{T}$ and $\boldsymbol{CNOT}$ between any pair of qubits. Note that the inclusion of the identity as an elementary gate ensures that we also count circuits of strictly smaller size. ∎

3. quantum state complexity by design

Date: 3 April 2025

3.1 Motivation and statement of results

Agenda:

1. motivation and results
2. recapitulation
3. proof via (partial derandomization)
4. generalizations and follow-up work

The complexity of a *circuit* (quantum or classical) is defined as the minimal number of elementary steps needed to evaluate the function. This depends on the choice of model ('gate set'), but only in a mild way. It allows us to assert whether a given computational task is 'easy' (small complexity) or 'hard' (high complexity).

In quantum information and computation, the notion of complexity extends meaningfully to quantum states as well. State complexity measures the effort/time required to produce $|\psi\rangle = \boldsymbol{\psi}$ from a simple starting state $|\psi_0\rangle = \boldsymbol{\psi}_0$, e.g. the all-zero initialization $|0\cdots 0\rangle$.

Here are two basic facts about the analysis of complexity:

- *upper bounds are 'easy'*, because every circuit decomposition yields one for free. Certain circuit families also come with universal upper bounds, e.g. $2^{O(n)}$ for n-qubit quantum circuits and $O(n^2/\log(n))$ for n-qubit Clifford circuits (i.e. circuits generated by Hadamard ($\boldsymbol{H}$), CNOT and phase ($\boldsymbol{S}$)).
- *lower bounds are 'hard'*, because it requires us to rule out potential shortcuts. In classical Boolean logic, complexity captures the notion of optimal compilation. This problem sits in the second level of the polynomial hierarchy. Quantum circuit compilation is even harder.

Circuit complexity has long been a prominent foundational concept in (classical and quantum) computer science. Over the past years, *state complexity* has been identified as a useful concept in quantum physics. Here are a couple

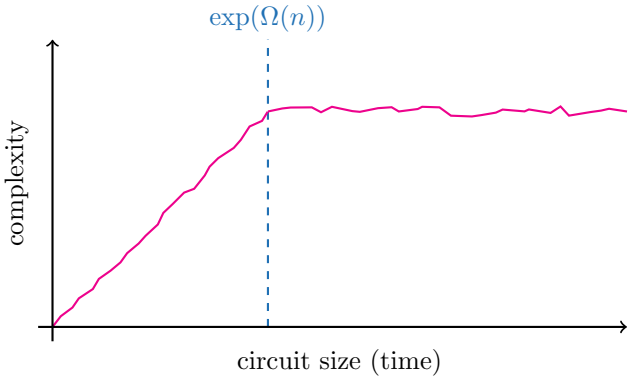


Figure 3.1 *Expected complexity growth in random circuits.* Conjecture 3.1 states that, for random quantum circuits acting on n qu$d$its, the circuit complexity grows linearly with circuit size (time) until it saturates at a value exponentially large in n.

of examples:

- **a1** *topological phases of matter (at zero temperature)* can be classified using the complexity of the ground state wave function;
- **a2** *chaotic Hamiltonians* produce long-time quantum evolutions that generate highly complex states;
- **a3** the *AdS/CFT-correspondence* posits that the complexity of a quantum state of the boundary theory corresponds to the volume in the bulk geometry, which is hidden behind the event horizon of a black hole.

It is extremely difficult to study complexity growth for concrete Hamiltonian evolutions. An alternative approach is to consider ensembles of circuits, and to derive lower bounds on complexity, which hold with high probability when samples are selected from these ensembles. This together with the AdS/CFT conjecture gave rise to the following conjecture:

the complexity of random circuits is conjectured to grow linearly with size

Conjecture 3.1 (Brown, Susskind [brown2018]). Most local (random) circuits of size T have a complexity that scales linearly in T for an exponentially long time.

This conjecture is visualized in Figure 3.1 Today, we will prove a related statement regarding the growth of state complexity under local random circuits on n qubits ($D = 2^n$). To achieve such a goal, we will work with the following standard notion of state complexity which we already saw last time.

formal definition of state complexity

Definition 3.2 (state complexity). Let V_R be the set of all size-R circuits on n qubits with gates from a universal gate set (e.g. Hadamard $\boldsymbol{H}$, T-gate $\boldsymbol{T}$ and CNOT gates), let $|\psi_0\rangle = \boldsymbol{\psi}_0$ be a fixed 'simple' state. For $\delta \in (0,1)$, we say that $|\psi\rangle = \boldsymbol{\psi} \in \mathbb{C}^D$ has δ-complexity at most R if

$$\max_{\boldsymbol{V} \in \mathsf{V}_R} F\left(\boldsymbol{\psi}, \boldsymbol{V}\boldsymbol{\psi}_0\right) \geq 1 - \delta^2.$$

If this is the case, we write $C_\delta(|\psi\rangle) \leq R$.

Based on this formal definition, we will prove the following rigorous lower bound on typical state complexity generated by local random quantum circuits. Such circuits arise step by step. At each step, two qubits are selected uniformly at random and a randomly selected elementary gate is applied to them.

'polynomial' state complexity growth

Theorem 3.3 ('polynomial' growth in state complexity, informal). 'Most' local random circuits of size T produce states with complexity (at least) $\Omega\left(T^{1/(5+o(1))}\right)$. This growth persists up to exponential circuit sizes $T \approx \sqrt{D} = 2^{n/2}$.

Similar statements are true for stronger notions of state and circuit complexity as well. We will briefly discuss some of them in Section 3.5. Remarkably, a recent result by Chi-Fang Chen, Jeongwan Haah, Jonas Haferkamp, Yunchao Liu, Tony Metger and Xinyu Tan improves this scaling from $\Omega\left(T^{1/(5+o(1))}\right)$ to $\Omega(T)$ for both the complexity of a quantum state and the complexity of a quantum circuit.

Theorem 3.4 ('linear' growth in circuit complexity, informal). 'Most' local random quantum circuits of size T cannot be accurately approximated by circuits of size sublinear in T. This growth persists up to exponential circuit sizes.

This proves the conjecture by Brown and Susskind mentioned earlier on.

3.2 Recapitulation: almost all states have high complexity

Let us start by recapitulating the complexity of generic (i.e. Haar-random) n-qubit states. These are states sampled uniformly from the complex unit sphere in $D = 2^n$ dimensions. There, concentration of measure together with a simple counting argument yield exponentially strong lower bounds on the state complexity.

exponential concentration for Haar-random states

Lemma 3.5 (Restatement of Lemma 2.5). Fix $|v\rangle \in \mathbb{C}^D$ $(D = 2^n)$ and let $|h\rangle = \boldsymbol{h} \overset{Haar}{\sim} \mathbb{C}^D$ be a Haar random state. Then,

$$mathrmPr\left[|\boldsymbol{v}^*\boldsymbol{h}|^2 \geq (1-\delta)\right] \leq 2\mathrm{e}^{-D(1-\delta)/2} \quad \text{for any } \delta \in (0,1).$$

This is a poor man's variant of a beautiful measure concentration phenomenon called Levy's lemma. It applies much more generally and is best proved via isoperometric inequalities. The argument presented here, however, does provide valuable guidance on how to deal with ensembles that are not quite Haar random.

Fact 3.6 (Haar integration; restatement of Fact 2.4). Let $|h\rangle \overset{Haar}{\sim} \mathbb{C}^D$ $(D = 2^n)$ be a Haar random state. Then, for all $k \in \mathbb{N}_+$

$$\mathbb{E}_{|h\rangle}\left[(|h\rangle\langle h|)^{\otimes k}\right] = \int_{\text{Haar}} (|h\rangle\langle h|)^{\otimes k}\, \mathrm{d}\mu(h) = \binom{D+k-1}{k}^{-1} P_{\vee k}, \qquad (3.1)$$

where $P_{\vee^k}$ is the projector onto the totally symmetric subspace of $\left(\mathbb{C}^D\right)^{\otimes k}$. ∎

Proof of Lemma 3.5. Let us start by computing the moments of the random variable $|\langle v, h\rangle|^2$. For any $k \in \mathbb{N}_+$ the Haar integration formula (Fact 2.4) yields

$$\begin{aligned}
\mathbb{E}_{|h\rangle}\left[|\langle v, h\rangle|^{2k}\right] =& \operatorname{tr}\left(\left(|v\rangle\langle v|\right)^{\otimes k} \mathbb{E}_{|h\rangle}\left[\left(|h\rangle\langle h|\right)^{\otimes k}\right]\right) \\
=& \operatorname{tr}\left(\left(|v\rangle\langle v|\right)^{\otimes k} \binom{D+k-1}{k}^{-1} P_{\vee^k}\right) \qquad (3.2) \\
=& \binom{D+k-1}{k} \leq \frac{k!}{D^k}. \qquad (3.3)
\end{aligned}$$

This moment behavior indicates sub-exponential moment growth. We can now use some elementary tricks from probability theory to turn these moment bounds into an exponential concentration bound:

$$\begin{aligned}
\Pr_{|h\rangle}\left[|\langle v, h\rangle|^2 \geq \tau\right] =& \Pr_{|h\rangle}\left[D|\langle v, h\rangle|^2/2 \geq D\tau/2\right] \\
=& \Pr_{|h\rangle}\left[\exp\left(D|\langle v, h\rangle|^2/2\right) \geq \exp\left(D\tau/2\right)\right] \\
\leq& \mathrm{e}^{-D\tau/2}\mathbb{E}\left[\exp\left(D|\langle v, h\rangle|^2/2\right)\right] \\
=& \mathrm{e}^{-D\tau/2}\sum_{k=0}^{\infty}\frac{1}{k!}\frac{D^k}{2^k}\mathbb{E}_{|h\rangle}\left[|\langle v, h\rangle|^{2k}\right] \\
\leq& \mathrm{e}^{-D\tau/2}\sum_{k=0}^{\infty}\frac{1}{2^k} = 2\mathrm{e}^{-D\tau/2}.
\end{aligned}$$

The key step is Markov's inequality ($\Pr[S \geq \alpha] \leq \mathbb{E}[S]/\alpha$ for any nonnegative random variable S) in line three. ■

As we saw last time, exponential concentration for Haar-random states implies a very strong claim about the complexity of generic states.

Note that the Haar measure is fair in the sense that it assigns the same infinitesimal weight to each D-dimensional quantum state. Laplace's definition of probability therefore allows us to interpret Eq. (2.2) as a bound on the relative volume of complexity-R states. This volume remains tiny until R approaches the overall Hilbert space dimension $D = 2^n$. In other words: *almost all states have almost maximal state complexity*.

3.3 Proof part 2: complexity by design

Our study of the complexity of Haar random states is a promising starting point. But it doesn't allow us to address less generic state ensembles. One solution is to apply a *partial derandomization* based on the Haar integration formula (3.1). The main result of last lecture (Haar-random states have exponentially high complexity) is contingent on the assumption that this formula is true for *all* tensor powers $k \in \mathbb{N}_+$ This allowed us to control all k moments of $|\langle \boldsymbol{v}, \boldsymbol{h}\rangle|^2$ and arrive at an exponentially strong concentration formula (Lemma 3.5). We

can relax these assumptions by assuming that the Haar integration formula is only approximately true for the first K tensor powers. Ensembles with this property are called ϵ-approximate K-designs.

(approximate) K-design

Definition 3.7 (approximate K-design). Fix $\epsilon \in (0,1)$ and a threshold $K \in \mathbb{N}_+$. We say that an ensemble $\{p_i, \boldsymbol{h}_i\} \subset \mathbb{S}^{D-1}$ of complex-valued unit vectors forms an ϵ-approximate K-design if

$$\left\| \mathbb{E}_{\boldsymbol{h}} \left[(\boldsymbol{h}\boldsymbol{h}^*)^{\otimes k} \right] - \binom{D+k-1}{k}^{-1} \boldsymbol{P}_{\vee^k} \right\|_1 \leq \epsilon \quad \text{for all } k = 1, \ldots, K.$$

Note that for $K = 1$, this requirement is met by any discrete vector distribution that is isotropic:

$$\mathbb{E}_{\boldsymbol{h}} \left[\boldsymbol{h}\boldsymbol{h}^* \right] = \sum\nolimits_i p_i \boldsymbol{h}_i \boldsymbol{h}_i^* = \frac{1}{D} \mathbb{I}.$$

One concrete example is sampling uniformly from an orthonormal basis of $\mathbb{C}^D$. Haar random vectors drawn uniformly from the sphere are another extreme case that occurs when we let K tend to infinity. Adjusting the design order K allows us to interpolate between those extremes. And, remarkably, the typical state complexity associated with such ensembles varies accordingly.

most K-design states have complexity linear in K

Proposition 3.8 (complexity by K-design). Suppose that $\boldsymbol{h} \in \mathbb{C}^D$ is sampled from an ϵ-approximate K-design. Then,

$$\Pr_{\boldsymbol{h}} \left[C_\delta(\boldsymbol{h}) \right] \leq (2n)^{2R} \left(\frac{K}{(1-\delta^2)D} \right)^K \quad \text{for all } R \in \mathbb{N}_+.$$

Disregarding constants and $(1-\delta^2)$, this probability remains tiny until

$$R \approx \frac{K\,(n - \log(K))}{\log(n)}.$$

Warning 3.9 Eq. (3.8) becomes vacuous once the design order K approaches the total system size $D = 2^n$. This puts an upper limit on the amount of progress we can make by letting K become larger and larger. ∎

Proof of Proposition 3.8. We will do the proof for the extreme case $\epsilon = 0$. An extension to $\epsilon > 0$ is relatively straightforward. The key ingredient is to replace the Haar concentration formula with a weaker concentration bound that only uses the first K moments. Fix $|v\rangle = \boldsymbol{v} \in \mathbb{C}^D$ arbitrary and $\tau \geq 0$. Markov's inequality then implies

polynomial concentration for K-designs

$$\begin{aligned} \Pr_{\boldsymbol{h}} \left[|\langle \boldsymbol{v}, \boldsymbol{h} \rangle|^2 \geq \tau \right] &= \Pr_{\boldsymbol{h}} \left[|\langle \boldsymbol{v}, \boldsymbol{h} \rangle|^{2K} \geq \tau^K \right] \\ &\leq \tau^{-K} \mathbb{E} \left[|\langle \boldsymbol{v}, \boldsymbol{h} \rangle|^{2K} \right] \\ &= \tau^{-k} \binom{D+K-1}{K}^{-1} \leq \left(\frac{K}{D\tau} \right)^K. \end{aligned}$$

Here, we have made use of the assumption that the K-design approximation is perfect ($\epsilon = 0$). This allows us to directly recycle the exact Haar integration from Eq. (3.3). The claim then readily follows from retracing the steps of the previous proof, but with this weaker polynomial concentration formula. ■

3.4 Proof part 3: connection to local random circuits

Proposition 3.8 highlights that the complexity of a randomly selected K-design state increases linearly with K until a certain threshold is met ($K \approx D = 2^n$). But, so far, K-designs have ben a rather abstract concept. The following deep result allows us to relate K-designs to random circuits of increasing size.

Fact 3.10 (local random circuits generate k-designs [Chen, Haah, Haferkamp, Liu, Metger, Tan).] Local random circuits of size $T = O\left(n\left(nK + \log(1/\epsilon)\right)\log^4(K)\right)$ produce state ensembles $U|\psi_0\rangle$ that form approximate K-designs with multiplicative error ϵ. ■

local random circuits form K-designs

This is a very recent and substantial improvement of a seminal result by Brandão, Horodecki and Haarow from 2016. Our main result is now an immediate consequence of Proposition 3.8 and Fact 3.10. The detailed conversion is a bit cumbersome, but here is the main gist.

Theorem 3.11 (linear growth in state complexity, informal). Fix $\boldsymbol{\psi}_0 = |\psi_0\rangle \in \mathbb{C}^D$ and let $\boldsymbol{U}$ be a local random circuit of size T. Then, with high probability

$$C_\delta\left(\boldsymbol{U}\boldsymbol{\psi}_0\right) = \Omega\left(\frac{T}{n^4}\right),$$

where we have suppressed the dependence on δ. This growth persists up to exponential circuit sizes $T = \mathcal{O}(\sqrt{D}) = \mathcal{O}\left(2^{n/2}\right)$.

linear state complexity growth (formal)

It is also possible to turn this probabilistic statement into a quantitative bound on the minimal number of high-complexity states that have this property. It must grow exponentially in circuit depth as $\exp\left(\Omega(T)\right)$. The trick is to exploit the fact that the weight distribution of a K-design cannot be too spiky. This then implies a direct relation between the probability of producing a high complexity state and the minimal number of high-complexity states within the entire ensemble.

3.5 Generalizations

The proof technique introduced above is very versatile and can be adjusted to cover stronger notions of complexity as well. Circuit complexity is one such example. We say that a unitary $\boldsymbol{U} \in \mathsf{U}(D)$ has δ-complexity at most R if

formal definition of circuit complexity

$$\min_{\boldsymbol{V}\in\mathsf{V}_R} \left\|\boldsymbol{U}\cdot\boldsymbol{U}^\dagger - \boldsymbol{V}\cdot\boldsymbol{V}^\dagger\right\|_\diamond \leq \delta. \tag{3.4}$$

Here, $\|\cdot\|_\diamond$ denotes the diamond distance of the two unitary channels involved. If Eq. (3.4) holds, we write $C_\delta(\boldsymbol{U}) \leq R$. It should be not surprising at this point that circuit complexity also grows with circuit size.

'polynomial' circuit complexity growth

Theorem 3.12 ('polynomial' growth in state complexity, informal). 'Most' local random circuits of size T produce unitaries with complexity (at least) $\Omega\left(T/n^4\right)$. This growth persists up to circuit sizes $T = \mathcal{O}\left(\sqrt{D}\right) = \mathcal{O}\left(2^{n/2}\right)$.

To prove this claim, it is helpful to first relate the diamond distance to another property that is easier to control:

$$\|\boldsymbol{U}\cdot\boldsymbol{U}^\dagger - \boldsymbol{V}\cdot\boldsymbol{V}^\dagger\|_\diamond \leq \delta \quad \Rightarrow \quad \left|\text{trace}\left(\boldsymbol{V}^\dagger\boldsymbol{U}\right)\right|^2 \geq D^2(1-\delta^2).$$

This necessary condition is much easier to control. In particular, we can use more general Haar integration techniques to bound moments and deduce polynomial concentration bounds. The rest of the proof is then almost identical to the state complexity case.

stronger/operational definitions of complexity

Next, we want to point out that it is possible to introduce *stronger/operational complexity notions* for both states and unitary circuits. These are based on the operational task of distinguishing the state/unitary in question from the most useless state/channel conceivable. For states this is the maximally mixed state $\boldsymbol{\tau} = \mathbb{I}/D$, while for channels this is the completely depolarizing channel $\mathcal{D}(\boldsymbol{\rho}) = \text{trace}(\boldsymbol{\rho})\boldsymbol{\tau}$. In both cases, the optimal single-shot distinguishability protocols are known. They give rise to the trace distance $\|\boldsymbol{\psi}\boldsymbol{\psi}^* - \boldsymbol{\tau}\|_1$ and the diamond distance $\|\boldsymbol{U}\cdot\boldsymbol{U}^\dagger - \mathcal{D}\|_\diamond$, respectively. But achieving these optimal values requires measurement procedures whose complexity mimics that of the state/unitary in question. This allows us to indirectly capture complexity by limiting the circuit size allowed for executing distinguishing measurements. A formal definition would go beyond the scope of this talk. Instead, we refer to Figure 3.2 (state complexity) and Figure 3.3 (circuit complexity) for visual illustrations. These stronger/operational notions of complexity imply the ones used so far, but the converse is not necessarily true, as the following example shows.

Example 3.13 Let $|h\rangle$ be a Haar-random state on $(n-1)$ qubits and define the n-qubit state $\boldsymbol{\psi} = \boldsymbol{h}\otimes\boldsymbol{e}_0 = |h\rangle\otimes|0\rangle$. Then, this state has exponential state complexity according to Definition 3.2. But it is actually very easy to distinguish this state vector from the maximally mixed state $\boldsymbol{\tau} = \mathbb{I}/D$. A computational basis measurement on the last qubit (and ignoring everything else) does the job with reasonable probability – especially if we replace qubits ($d = 2$) with higher-dimensional qu$d$its. ∎

This feature of strong/operational complexity delays the onset of complexity growth up to circuit sizes that cover all qu$d$its involved. Such a behavior accurately addresses physical effects like operator growth and the switchback effect in holography.

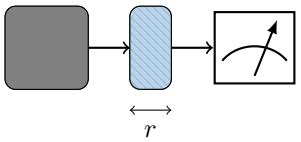


Figure 3.2 *Pictographic illustration of strong state complexity.* A black-box either outputs a (known) pure state $\rho = |\psi\rangle\langle\psi|$, or the maximally mixed state $\rho_0 = \frac{1}{d}\mathbb{I}$. The task is to correctly guess which one it produced by applying a pre-processing circuit V (blue line pattern) of limited size r and performing a simple measurement (right). We say that $|\psi\rangle$ has *strong/operational state complexity* at most r if the probability of correctly distinguishing both possibilities is close to optimal.

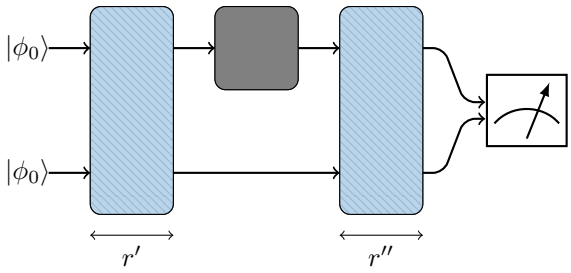


Figure 3.3 *Pictographic illustration of strong circuit complexity.* A black box (center) takes quantum states as inputs and applies either a unitary channel $\mathcal{U}(\rho) = U\rho U^\dagger$, or the depolarizing channel $\mathcal{D}(\rho) = \tau = \mathbb{I}/D$. The task is to correctly guess which evolution occurred. The rules of the game allow short pre- and post-processing circuits (blue line patterns) that may involve a quantum memory. The final guess must be based on a simple measurement (right). We say that U has *strong/operational circuit complexity* at most $r = r' + r''$ if the probability of correctly distinguishing both options is close to optimal.