Decoherence, NISQ computers and quantum supremacy

Introduction to Quantum Computing

Kenneth MAUSSANG

Université de Montpellier

Ecole Centrale Casablanca

2020 – 2021

Decoherence, NISQ computers and quantum supremacy

1. Density matrix - T_1 and T_2 times and decoherence
2. Quantum advantage, quantum supremacy
3. Quantum annealer

Introduction

Today's better quantum computers have error rates of typically 0.1% to 1%, with 50 to 100 qubits chips only.

As a comparison, a typical classical computer chip holds about $20 \cdot 10^9$ bits (or transistors) while the latest smartphone chips holds about $6 \cdot 10^9$ bits.

Classical digital computers are truly reliable at the bit level, with fewer than 1 error in 10^{24} operations. The far more common sources of error are software and mechanical malfunction.

Decoherence, NISQ computers and quantum supremacy

Decoherence, NISQ computers and quantum supremacy

I.1. Definition

Quantification of decoherence in quantum computers or quantum interferometer is of importance to quantify devices' performances.

A commonly used tool in that aim is the **density matrix** $\hat{\rho}$.

The later permit to define T_1 and T_2 times.

T_1 time refers to state population relaxation toward equilibrium (for example relaxation of spins in the direction of a magnetic field), while T_2 is a dephasing time between particles and refers to decoherence of a quantum superposition toward a classical superposition of states.

For example, if one considers an ensemble of N particles in the state

$$|\psi\rangle = \frac{1}{\sqrt{2^N}} \left(|0\rangle + |1\rangle\right)^{\otimes N},$$

one will measure $N/2$ atoms in state $|0\rangle$ and $N/2$ atoms in state $|1\rangle$. It is a pure quantum state.

Now one consider a classical statistical system made of $N/2$ atoms in state $|0\rangle$ and $N/2$ atoms in state $|1\rangle$. This ensemble will result in the same result after measurement but it is not the same quantum state!

I.1. Definition

One uses the density matrix formalism to describe a statistical ensemble of quantum particles.

A pure state is a system without any statistical superposition, that could be described by a single quantum state $|\psi\rangle$. The density matrix operator is then defined as follow

$$\boxed{\hat{\rho} = |\psi\rangle\langle\psi|}.$$

Let's consider an observable $\hat{\mathcal{O}}$. Since one has the following relation

$$\mathrm{Tr}\left(\hat{\rho}\hat{\mathcal{O}}\right) = \mathrm{Tr}\left(\hat{O}\left|\psi\right\rangle\left\langle\psi\right|\right) = \mathrm{Tr}\left(\left\langle\psi\right|\hat{O}\left|\psi\right\rangle\right) = \left\langle\hat{\mathcal{O}}\right\rangle.$$

So the density matrix $\hat{\rho}$ permits to evaluate the mean value of any operator

$$\left\langle\hat{\mathcal{O}}\right\rangle = \mathrm{Tr}\left(\hat{\rho}\hat{\mathcal{O}}\right).$$

I.1. Definition

In the case of a two level particle, the Bloch representation of any state is the following

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + \sin\left(\frac{\theta}{2}\right)e^{i\varphi}|1\rangle.$$

Geometrically, the state is represented on the Bloch sphere with a unitary vector $\vec{n}$ such that

$$\vec{n} = \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{pmatrix}.$$

A straightforward permits to demonstrate easily that in the case of two-level system

$$\boxed{\hat{\rho} = |\psi\rangle\langle\psi| = \frac{\hat{\mathbb{I}} + \vec{n}\cdot\hat{\vec{\sigma}}}{2}}, \quad \text{with} \quad \hat{\vec{\sigma}} = \hat{\sigma}_x\vec{u}_x + \hat{\sigma}_y\vec{u}_y + \hat{\sigma}_z\vec{u}_z.$$

I.1. Definition

A mixed state is a classical statistical superposition of an ensemble of orthogonal states. If $\{|\psi_1\rangle, ..., |\psi_N\rangle\}$ is an ensemble of N states, with statistical weight $\{w_i\}$ such that $\sum_i w_i = 1$, then the density matrix $\hat{\rho}$, of such a so-called mixed state, is defined as follow

$$\hat{\rho} = \sum_{i=1}^{N} w_i \, |\psi_i\rangle\langle\psi_i|.$$

Then, for an observable $\hat{\mathcal{O}}$,

$$\left\langle \hat{\mathcal{O}} \right\rangle = \mathrm{Tr}\left(\hat{\rho}\hat{\mathcal{O}}\right) = \sum_{i=1}^{N} w_i \mathrm{Tr}\left(|\psi_i\rangle\langle\psi_i|\hat{\mathcal{O}}\right) = \sum_{i=1}^{N} w_i \langle\psi_i|\hat{\mathcal{O}}\,|\psi_i\rangle,$$

as expected for a classical statistical superposition.

I.1. Definition

Now if one considers the thermal equilibrium and the basis $|i\rangle$ of eigenvectors of the Hamiltonian, the density matrix is provided by Maxwell statistical distribution

$$\hat{\rho} = \sum_{i=1}^{N} \frac{e^{-\beta E_i}}{Z} |i\rangle\langle i|,$$

with Z the partition function, E_i the energy of state $|i\rangle$ and $\beta = \frac{1}{k_B T}$. This state is known as the *Gibbs state* or *thermal state*, describing quantum systems at thermal equilibrium. Since $Z = \mathrm{Tr}\left(e^{-\beta \hat{H}}\right)$ for $\hat{H}$ the Hamiltonian, one may express it as

$$\hat{\rho} = \frac{e^{-\beta \hat{H}}}{\mathrm{Tr}\left(e^{-\beta \hat{H}}\right)}.$$

I.2. The Bloch ball

Now one resists oneself to the case of a statistical ensemble of two level particles, either in pure or mixed states. The density matrix $\hat{\rho}$ might be decomposed of Pauli's matrices basis. Moreover, it is straightforward to show that

$$\mathrm{Tr}\,(\hat{\rho}) = 1.$$

And since $\mathrm{Tr}\,\hat{\sigma}_x = \mathrm{Tr}\,\hat{\sigma}_y = \mathrm{Tr}\,\hat{\sigma}_z = 0$ but $\mathrm{Tr}\,\hat{\mathbb{I}} = 2$, $\exists \vec{n}$ such that

$$\hat{\rho} = \frac{\mathbb{I} + \vec{n} \cdot \hat{\vec{\sigma}}}{2}.$$

I.2. The Bloch ball

It is straightforward that the operator $\hat{\rho}$ is positive, thus with positive eignevalues. The eigenvalues of $\vec{n} \cdot \hat{\vec{\sigma}}$ are $\pm \|\vec{n}\|$. So the positivity of the operator is verified if and only if

$$\|\vec{n}\| \leq 1.$$

So any two level system might be represented in a Bloch ball, *i.e.* described by a vector $\vec{n}$ in a unitary ball (and no longer only a sphere), such that $\|\vec{n}\| \leq 1$.

$$\hat{\rho} \in \left\{ \frac{\mathbb{I} + \vec{n} \cdot \hat{\vec{\sigma}}}{2} \ , \ \|\vec{n}\| \leq 1 \right\}.$$

The vector $\vec{n}$ will be unitary ($\|\vec{n}\| = 1$) only in the case of a pure state.

The Schrödinger equation states that

$$i\hbar\frac{\partial}{\partial t}\left|\psi\right\rangle = \hat{H}\left|\psi\right\rangle,$$

and by taking the complex conjugate one gets

$$-i\hbar\frac{\partial}{\partial t}\left\langle\psi\right| = \left\langle\psi\right|\hat{H},$$

such that

$$\frac{\partial}{\partial t}\left|\psi\right\rangle\left\langle\psi\right| = -\frac{i}{\hbar}\left(\hat{H}\left|\psi\right\rangle\left\langle\psi\right| - \left|\psi\right\rangle\left\langle\psi\right|\hat{H}\right) = -\frac{i}{\hbar}\left[\hat{H},\hat{\rho}\right].$$

So the equation of motion for the density matrix operator evolves in time according to

$$\boxed{i\hbar\frac{\partial\hat{\rho}}{\partial t} = \left[\hat{H},\hat{\rho}\right]}.$$

Let consider a two level system, of basis $\{|0\rangle, |1\rangle\}$ of respecting energies 0 and $\hbar\omega_0$. A thermal equilibrium one has a probability $P_{|0\rangle}$ (resp. $P_{|1\rangle}$) to be in state $|0\rangle$ (resp. $|1\rangle$), with

$$P_{|0\rangle} = \frac{1}{1 + e^{-\beta\hbar\omega}} \quad \text{and} \quad P_{|1\rangle} = \frac{e^{-\beta\hbar\omega}}{1 + e^{-\beta\hbar\omega}}.$$

One introduces the density matrix written as

$$\hat{\rho} = \begin{pmatrix} \rho_{11} & \rho_{10} \\ \rho_{01} & \rho_{00} \end{pmatrix}.$$

I.4. Decoherence

At thermal equilibrium, the density matrix as the following expression

$$\hat{\rho}_{\text{thermal}} = \frac{1}{1 + e^{-\beta\hbar\omega}} \begin{pmatrix} e^{-\beta\hbar\omega} & 0 \\ 0 & 1 \end{pmatrix},$$

while for a Bell state such as $|\Psi+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$,

$$\hat{\rho}_{|\Psi+\rangle} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

The diagonal terms of a density matrix are called *populations*, and the non-diagonal terms are called *coherences*. A classical state has non coherence, *i.e.* no non-diagonal terms. A quantum superposition will relax to a classical state with a decrease of its non-diagonal terms.

I.4. Decoherence

Now if one consider an ensemble of N two level atoms initially in the state $|\Psi+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, there time evolution will be

$$|\psi(t)\rangle = \frac{|0\rangle + e^{i\omega_0 t}|1\rangle}{\sqrt{2}}.$$

But due to the interaction of the environment, the energy of the excited state might be shifted (fluctuation of magnetic field, etc...) from $\hbar\omega_0$ to $\hbar(\omega_0 + \delta\omega)$. It results in time evolutions differents depending on this energy shift. Due to the dispersion in the excited state energy, qubits will progressively undergo relative dephasing between one to each others

$$|\psi(t)\rangle = \frac{|0\rangle + e^{i\delta\omega t}e^{i\omega_0 t}|1\rangle}{\sqrt{2}}.$$

This is responsible of the decoherence, mixing up phases of superpositions and providing collapse of the non-diagonal terms of the density matrix.

The density matrix becomes

$$\hat{\rho}_{|\Psi+\rangle} = \frac{1}{2}\begin{pmatrix} 1 & e^{-i\omega_0 t}\int p(\delta\omega)e^{-i\delta\omega t}\mathrm{d}\delta\omega \\ e^{i\omega_0 t}\int p(\delta\omega)e^{i\delta\omega t}\mathrm{d}\delta\omega & 1 \end{pmatrix},$$

where $p(\delta\omega)$ is the statistical distribution of the energy shift.

Moreover, it is straightforward that

$$\lim_{t\to+\infty}\int p(\delta\omega)e^{i\delta\omega t}\mathrm{d}\delta\omega = 0.$$

Consequently, a pure state progressively collapse to a mixture state. But this simple model only explain the collapse of non-diagonal terms (phase decoherence). The interaction of the environment is also responsible of state rotation so that populations relaxes toward the thermal distribution.

The phase collapse and the population relaxation occurs with different timescales.

I.4. Decoherence

T_1 is the population relaxation time constant on which diagonal terms of the density matrix exponentially relaxes (as e^{-t/T_1}). It might be modelized by an additional term in the equation of evolution of the following form

$$\left.\frac{\partial\hat{\rho}}{\partial t}\right)_{\text{relaxation}} = -\frac{1}{T_1}\left(\hat{\rho} - \hat{\rho}_{\text{thermal}}\right).$$

T_2 is the coherence relaxation time constant (or dephasing time) on which non-diagonal terms of the density matrix exponentially relaxes (as e^{-t/T_2}). It might be modelized by an additional term in the equation of evolution of the following form

$$\left.\frac{\partial\hat{\rho}}{\partial t}\right)_{\text{dephasing}} = -\frac{1}{T_2}\begin{pmatrix} 0 & \rho_{10} \\ \rho_{01} & 0 \end{pmatrix}.$$

I.4. Decoherence

To model the global dynamic, two additional terms are empirically added to the time evolution equation of the density matrix as follow

$$\boxed{\frac{\partial \hat{\rho}}{\partial t} = -\frac{i}{\hbar}\left[\hat{H}, \hat{\rho}\right] - \frac{1}{T_1}\left(\hat{\rho} - \hat{\rho}_{\text{thermal}}\right) - \frac{1}{T_2}\begin{pmatrix} 0 & \rho_{10} \\ \rho_{01} & 0 \end{pmatrix}}.$$

Decoherence, NISQ computers and quantum supremacy

1. Density matrix - T_1 and T_2 times and decoherence

2. Quantum advantage, quantum supremacy
   - Definitions
   - Demonstration (or not ?) of quantum supremacy by Google in 2019
   - IBM's answer to Google

3. Quantum annealer

Decoherence, NISQ computers and quantum supremacy

1. Density matrix - T_1 and T_2 times and decoherence

2. Quantum advantage, quantum supremacy
   - Definitions
   - Demonstration (or not ?) of quantum supremacy by Google in 2019
   - IBM's answer to Google

3. Quantum annealer

II.1. Definitions

Both quantum supremacy and quantum advantage are conceptual criteria, more related to theoretical computer sciences, regardless to the usefulness of the problem considered.

Quantum advantage

It consists in demonstrating that a quantum device can solve a problem faster than classical computers.

Quantum supremacy

It is the demonstration that a programmable quantum device can solve a problem that a classical computer practically can not due to its complexity ("hard problems" with important computation time).

Decoherence, NISQ computers and quantum supremacy

1. Density matrix - T_1 and T_2 times and decoherence

2. Quantum advantage, quantum supremacy
   - Definitions
   - Demonstration (or not ?) of quantum supremacy by Google in 2019
   - IBM's answer to Google

3. Quantum annealer

II.2. Demonstration (or not ?) of quantum supremacy by Google in 2019

Sampling the output distributions of random quantum circuits is believed to be hard for classical computers based on reasonable complexity assumptions.

Google has announced to work on demonstration of quantum supremacy before the end of 2017 by solving this problem with an array of 49 superconducting qubits. But the main challenge was to develop such a chip, with acceptable error rates.

In October 2017, IBM demonstrated the simulation of 56 qubits on a conventional supercomputer, increasing the number of qubits needed for quantum supremacy. Until January 2018, only Intel has officially announced to have produced such a chip, with indeed the goal to demonstrate quantum supremacy.

II.2. Demonstration (or not ?) of quantum supremacy by Google in 2019

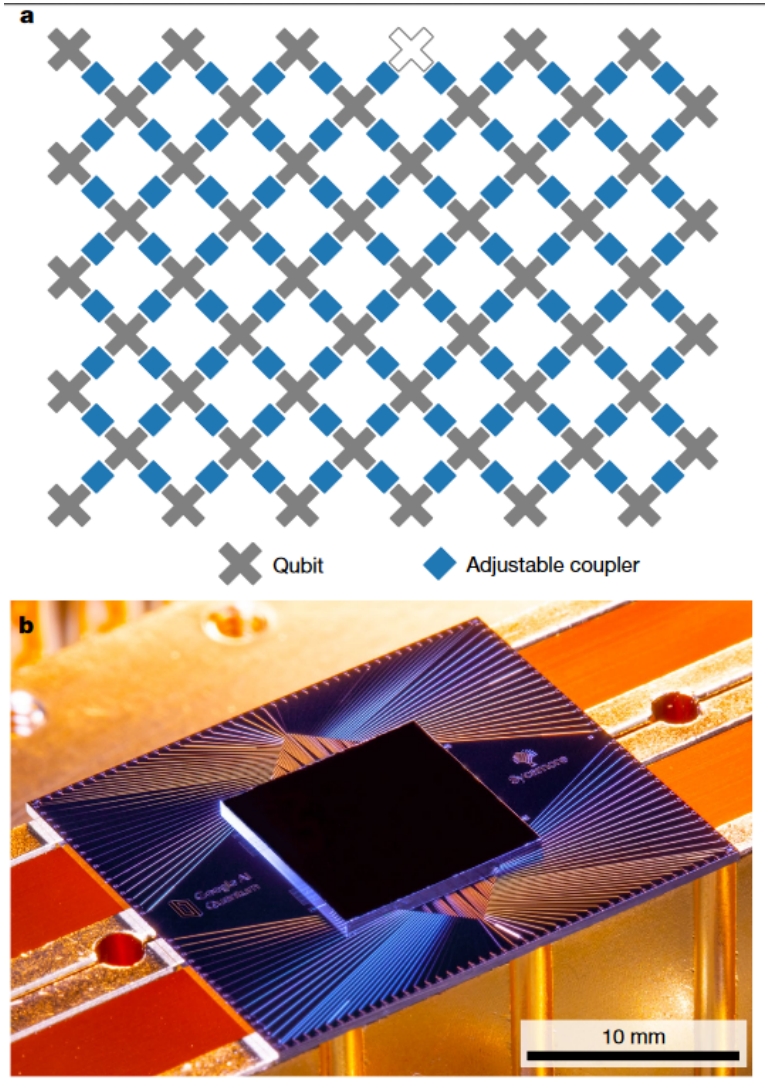


Fig. 1 | The Sycamore processor. a, Layout of processor, showing a rectangular array of 54 qubits (grey), each connected to its four nearest neighbours with couplers (blue). The inoperable qubit is outlined. **b**, Photograph of the Sycamore chip.

In October 2019, Google has published an article in *Nature*, claiming its quantum computing research team has demonstrated quantum supremacy with a 53 qubits quantum computer, based on superconducting qubits.

Quantum supremacy using a programmable superconducting processor, Nature volume 574, pages 505–510 (2019)

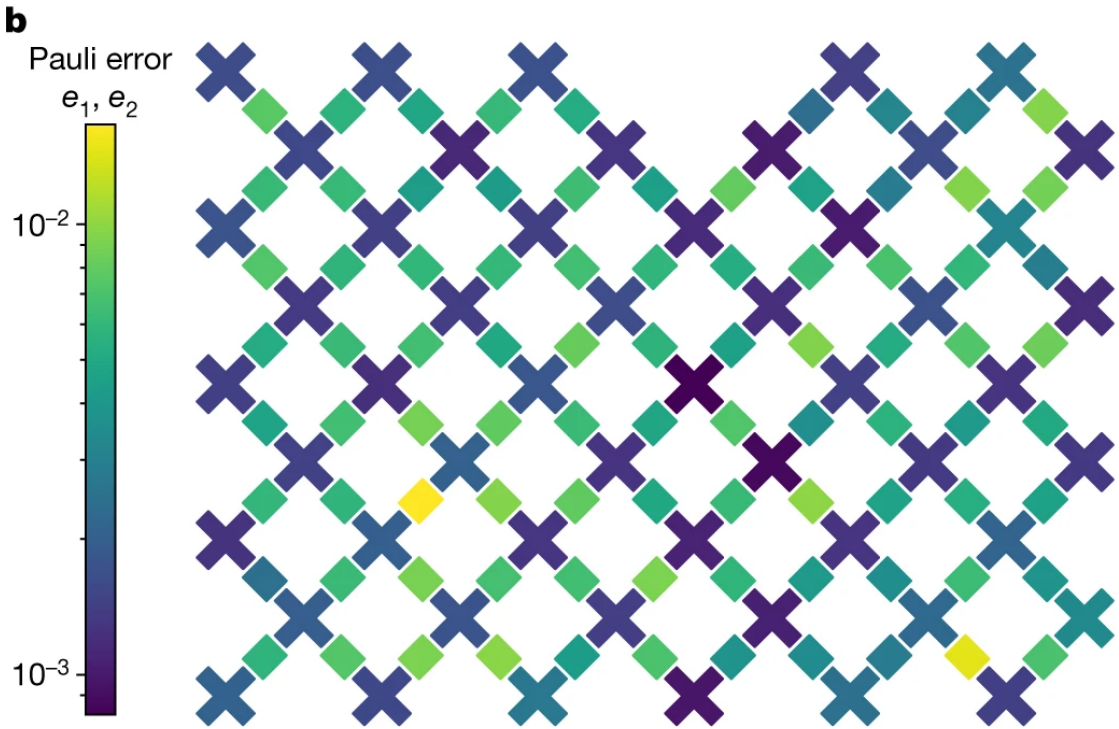
II.2. Demonstration (or not ?) of quantum supremacy by Google in 2019

- chip's name: *Sycamore*;
- programmable processor;
- typical single qubit gate error rate: about 0.15%;
- 2-qubits gates error rate: 0.6%;
- execution time of 2-qubits gates: 12 ns;
- estimated global circuit fidelity: $\mathcal{F}$=0.2%
  (circuits with 20 cycles of 2-qubits gates - 430 two-qubits gates and 1113 single-qubit gates);
- measurement error rates: typically 3-4%;
- 53 qubits
  computational state-space of dimension 2^{53} ($\sim 10^{16}$ states).

II.2. Demonstration (or not ?) of quantum supremacy by Google in 2019

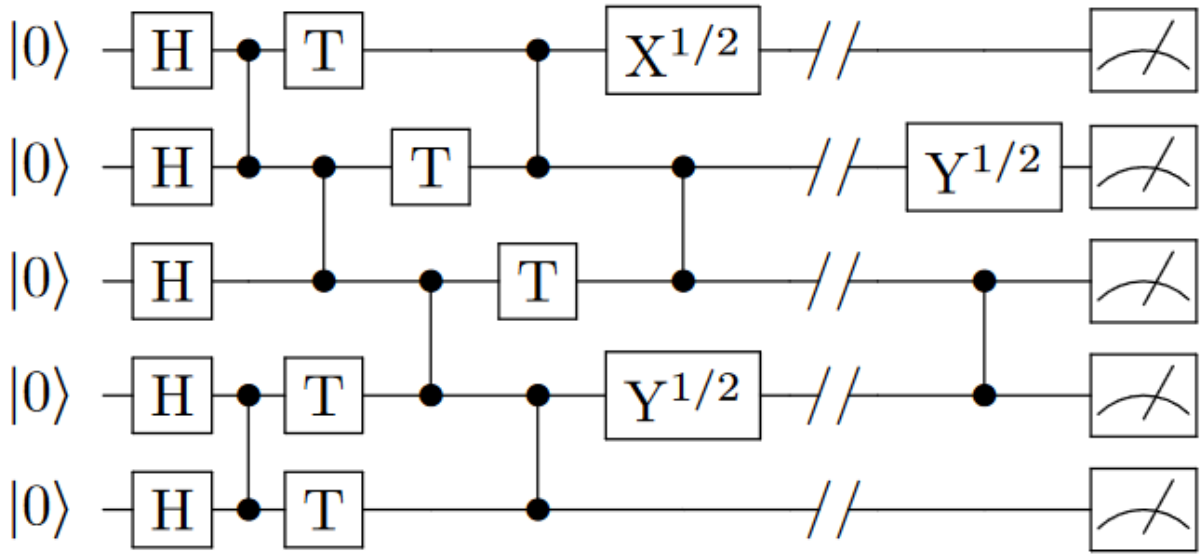
Pauli and measurement errors

Average error	Isolated	Simultaneous
Single-qubit (e_1)	0.15%	0.16%
Two-qubit (e_2)	0.36%	0.62%
Two-qubit, cycle (e_{2c})	0.65%	0.93%
Readout (e_r)	3.1%	3.8%



Quantum supremacy using a programmable superconducting processor, Nature volume 574, pages 505–510 (2019)

A specific algorithm has been chosen, constructed specifically to demonstrate quantum supremacy. The algorithm is based on a circuit with fixed two-qubits gates and randomly-chosen single qubit gates is chosen. The circuit is a sequence of d clock cycles of one- and two-qubits gates, with gates applied to different qubits in the same cycle.



Example of a random circuit used in the case of a 1D geometry of qubits.

Random quantum circuits with gates sampled from a universal gate set are examples of quantum chaotic evolutions that naturally lend themselves to the quantum computational framework.

The algorithm is based on a circuit with fixed two-qubits gates and randomly-chosen single qubit gates is chosen. The circuit is a sequence of d clock cycles of one- and two-qubits gates, with gates applied to different qubits in the same cycle.

The corresponding circuit is executed millions of times for $d = 20$ cycles.

Each time, all qubits are measured, generating a 53-bit string. The collected sample of 53-bit strings is not uniformly distributed.

Comparing with classical simulations one can verify "heavy output generation" - that the average probability of strings in the sample is greater than 2^{-n}. Because a random circuit has no structure, and the Hilbert space is exponentially large in n, simulation using a classical supercomputer is hard. Experiment verifies that the hardware is working well enough to produce meaningful results in a regime where classical simulation is very difficult.

II.2. Demonstration (or not ?) of quantum supremacy by Google in 2019

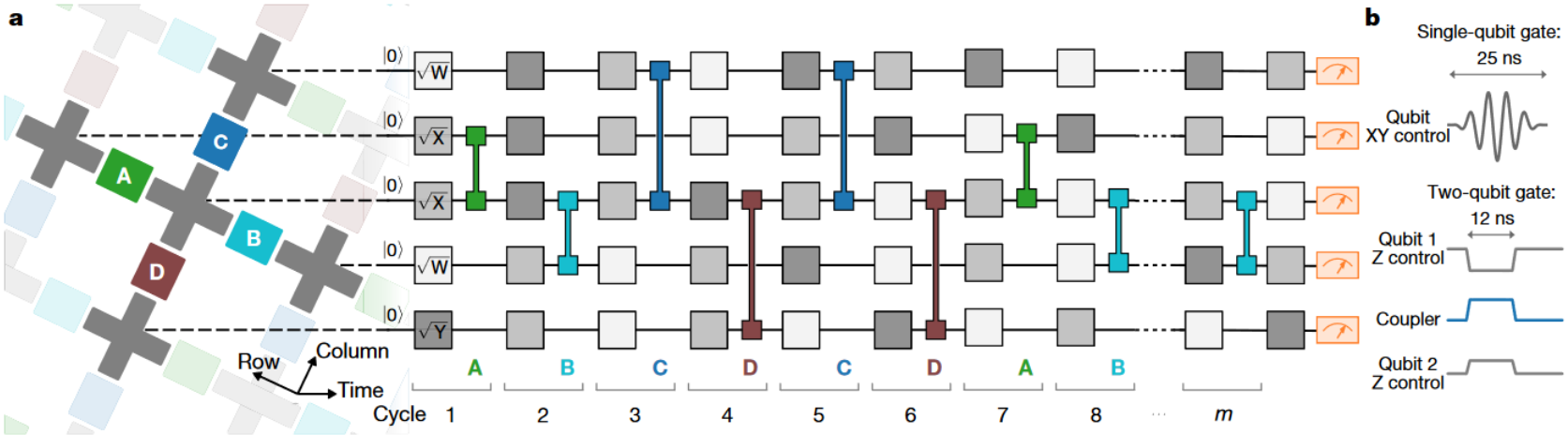


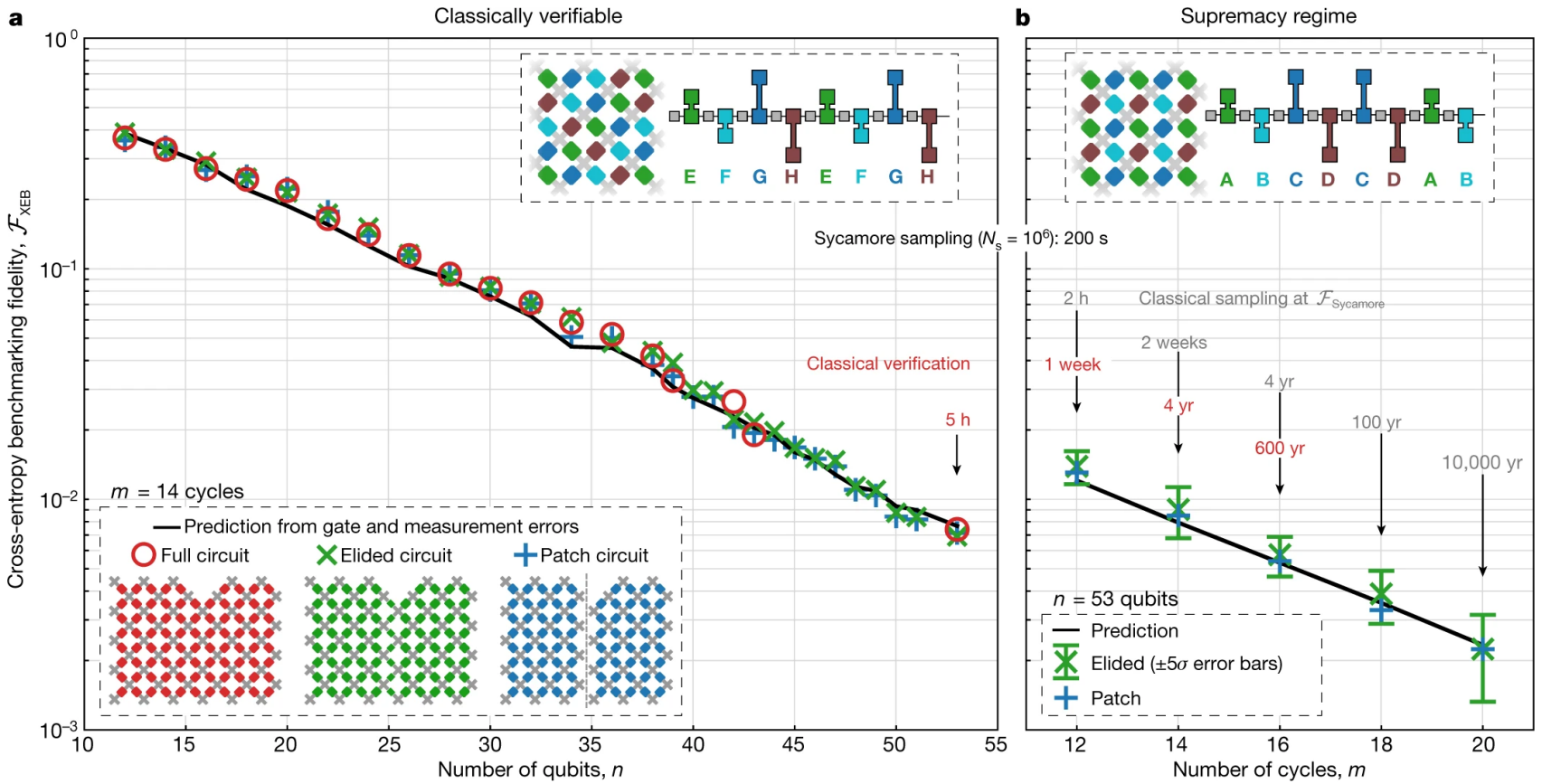
Fig. 3 | Control operations for the quantum supremacy circuits. a, Example quantum circuit instance used in our experiment. Every cycle includes a layer each of single- and two-qubit gates. The single-qubit gates are chosen randomly from $\{\sqrt{X}, \sqrt{Y}, \sqrt{W}\}$, where $W = (X+Y)/\sqrt{2}$ and gates do not repeat sequentially. The sequence of two-qubit gates is chosen according to a tiling pattern, coupling each qubit sequentially to its four nearest-neighbour qubits. The couplers are divided into four subsets (ABCD), each of which is executed simultaneously across the entire array corresponding to shaded colours. Here we show an intractable sequence (repeat ABCDCDAB); we also use different coupler subsets along with a simplifiable sequence (repeat EFGHEFGH, not shown) that can be simulated on a classical computer. **b**, Waveform of control signals for single- and two-qubit gates.

Random circuit used.

Measurements from repeated experiments sample the resulting probability distribution, which was verified then using classical simulations. **It tooks to the Sycamore processor about 200 seconds to sample one instance of a quantum circuit, while a million times—our benchmarks currently indicate that the equivalent task for a state-of-the-art classical supercomputer would take approximately 10,000 years!**

The strategy of Google consisted in choosing an algorithm adapted to the design of the quantum chip they used. This permits to avoid the use of swap gates to couples non-adjacent qubits and reduce the number of quantum gates requires. The algorithm used was not of practical interest but well adapted to a quantum chip and not at all for a classical calculation.

II.2. Demonstration (or not ?) of quantum supremacy by Google in 2019



Quantum supremacy using a programmable superconducting processor, Nature volume 574, pages 505–510 (2019)

This results is a milestone is the sense that it demonstrates that a 53 qubits programmable quantum computers might be used to realize a quantum calculation, even if qubits manipulations are not perfect, with good but not negligeable error rates.

John Preskill, professor of Theoretical Physics at CalTech, has qualified such type of devices: *Noisy Intermediate-Scale Quantum* (NISQ).

NISQ devices cannot be simulated by brute force using the most powerful currently existing supercomputers. But noise limits the computational power of such NISQ devices. **NISQ devices are not disruptive technologies by themselves**, but rather a step toward more powerful quantum technologies in the future.

For Preskill, the next steps consist in

- finding real world applications;
- dramatically extent qubits lifetimes either with quantum error correction codes or with qubits fabrication process improvements (or technological paradigm shift for qubits implementation);
- improve significantly the two-qubits gates fidelity;
- develop chips with more qubits and better gates.

Decoherence, NISQ computers and quantum supremacy

1. Density matrix - T_1 and T_2 times and decoherence

2. Quantum advantage, quantum supremacy
   - Definitions
   - Demonstration (or not ?) of quantum supremacy by Google in 2019
   - IBM's answer to Google

3. Quantum annealer

II.3. IBM's answer to Google

The quantum supremacy claimed has been contested by IBM researchers. They have discovered a classical algorithm to simulate the execution of Google's quantum algorithm, which can be implemented on a classical supercomputer (Summit supercomputer at Oak Ridge National Laboratories).

They are able to simulate 53- and 54-qubits Sycamore circuits with high fidelity to arbitrary depth (number of cycle d). They have demonstrated that for 53-qubits with depth of 20 cycles, the result of the quantum calculation can be obtained with their supercomputer in only 2.5 days! While Google claimed classical computers required 10,000 years.

It is still larger than the 200 s obtained with the quantum computer, but it is clearly less obvious to claim quantum supremacy. IBM stressed out that it was only a preliminary estimation of the calculation time, with the most pessimist estimation.

II.3. IBM's answer to Google

At the end, it could be possible to obtain the same result than the quantum computer with a calculation time even shorter. But the implementation on the Oak Ridge's supercomputer has not been yet realized.

A calculation time of 200 s is certainly shorter than a time of the order of one day, but if IBM is right, quantum supremacy is still not achieved.

But Google's result with Sycamore remains a milestone: the achievement of a quantum calculation with 53 qubits on a real quantum computer (with noisy qubits).

III.1. Quantum annealing processor

Quantum annealing is a method for finding the global minimum of a given function over a given set of candidate solutions, thanks to a process using quantum fluctuations. It is adapted for finding extrema of multidimensional functions.

A quantum annealing processor naturally returns low-energy solutions of a given potential. Quantum annealing processors are chips made out of a lot of qubits but **not programmable**, in the sense they can't implement a circuit diagram with quantum gates.

Quantum annealing is used mainly for combinatorial optimization problems (with many local minima).

Quantum annealing starts initializing an ensemble of qubits in a superposition of all possible states (candidate solutions) with equal weights. Then the system evolves, governed by Schrödinger equation.

III.2. Optimization problems

Physics can help solve these sorts of problems because we can frame them as energy minimization problems. That is the key point of a quantum annealer: it converts an optimization problem into a fundamental state search of a quantum system. A fundamental rule of physics is that everything tends to seek a minimum energy state.

Quantum annealing simply uses quantum physics to find low-energy states of a problem and therefore the optimal or near-optimal combination of elements.

D-Wave Systems is a Canadian company D-Wave Systems which provide quantum machines dedicated to perform doing quantum annealing.

Decoherence, NISQ computers and quantum supremacy

III.3. D-wave quantum processors

- In 2011, Lockheed-Martin purchased a D-Wave One model for about $ 10 million.
- In May 2013, Google purchased a D-Wave Two with 512 qubits.

As of now, the question of whether the D-Wave processors offer a speedup over a classical processor is still unanswered. Tests performed by researchers at Quantum Artificial Intelligence Lab (NASA), USC, ETH Zurich, and Google show that until now there is no evidence of a quantum advantage.

Being the only kind of quantum computer available for actual sale (assuming you have $10 million to $15 million to spare) has made D-Wave unique for several years. D-Wave's recently launched real-time cloud platform, called Leap.

III.4. The user's view

A D-Wave quantum annealer is constituted by two main elements:

1. a Quantum Processor Unit (QPU) that implement the quantum annealing algorithm;
2. a conventional computer containing a front end server (solver application programming interface, SAPI), and a back end system that communicates with the QPU.

III.4. The user's view

An objective function is a mathematical expression of the energy of a system as a function of binary variables representing the qubits.

In most cases, the lower is the energy of the objective function, the better the solution. Sometimes any low-energy state is an acceptable solution to the original problem; for other problems, only optimal solutions are acceptable. The best solutions typically correspond to the global minimum energy in the solution space.

III.4. The user's view

The NP problem to solve has to be translated to an input for an equivalent Ising model. Then, it is transformed into a so-called *native* problem that matches the qubits connections topology of the D-Wave processor.

Binary objective functions can be represented as graphs.
The Ising model problem is defined as follows: given a graph $G = (V, E)$ with weights h_i (called *fields*) on vertices and J_{ij} (called *couplers*) on edges, find an assignment of *spins* $S = (s_1, ..., s_n)$ to vertices, with $s_i \in \{-1, +1\}$ so they minimize the *energy function* $H(S)$

$$H(S) = \sum_{i \in V} h_i s_i + \sum_{(i,j) \in E} J_{ij} s_i s_j.$$

III.4. The user's view

Quadratic unconstrained binary optimization (QUBO) is a pattern matching technique, common in machine learning applications.

QUBO is an NP hard problem. Examples of problems that can be formulated as QUBO problems are the maximum cut, graph coloring and the partition problem. For a QUBO problem, variables are TRUE and FALSE (corresponding qubits states correspond to 1 and 0 values).

A QUBO problem is defined using an upper-diagonal matrix Q, which is an $N \times N$ upper-triangular matrix of weights, and x a vector of binary variables. This matrix is used to define the function $f(x)$ to be minimized

$$f(x) = \sum_i Q_{ii} x_i + \sum_{i<j} Q_{ij} x_i x_j,$$

where the diagonal terms Q_{ii} are the linear coefficients and the nonzero off-diagonal terms are the quadratic coefficients Q_{ij}.

This minimization of $f(x)$ is equivalent to

$$\min_{x \in \{0,1\}^n} x^t Q x.$$

The objective function of the QUBO problem might be reformulate as the following expression in scalar notation

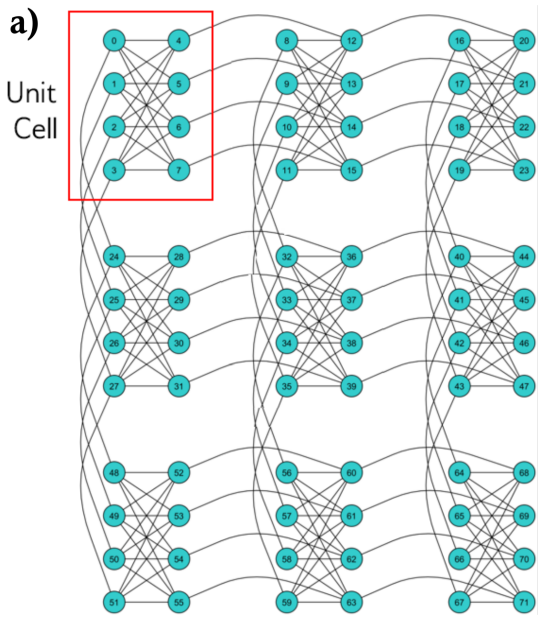
$$H_{\text{QUBO}}(a_i, b_{ij}, q_i) = \sum_i a_i q_i + \sum_{i<j} b_{ij} q_i q_j.$$

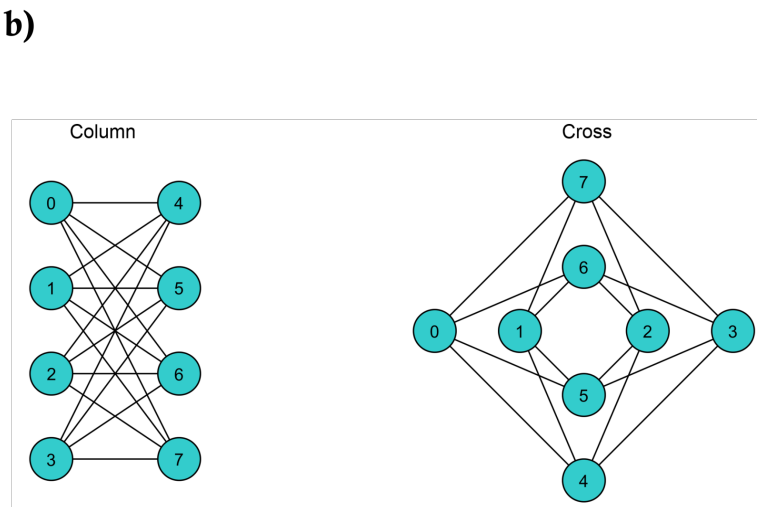
A QUBO problem might be transformed to an Ising problem with the following correspondence

$$s = 2q - 1.$$

III.4. The user's view

The D-Wave QPU is a lattice of interconnected qubits. While some qubits connect to others via couplers, the D-Wave QPU is not fully connected. Qubits interconnect in an architecture known as Chimera. D-Wave's most recent QPU, 2000Q, supports a $C16$ Chimera graph, with 2048 qubits logically mapped into a 16×16 matrix of unit cells made of 8 qubits.





III.5. The quantum annealing algorithm

The Ising model implemented in the QPU is described by the following Hamiltonian

$$\mathcal{H}_p = \sum_i h_i \sigma_i^z + \sum_{i<j} J_{ij} \sigma_i^z \sigma_j^z,$$

where σ_i^z is the Pauli matrix z acting on spin i, h_i is the magnetic field on spin i and J_{ij} the coupling strength between spins i and j. The ground state of $\mathcal{H}_p$ corresponds to a spin configuration $S = (s_1, ..., s_n) \in \{-1, +1\}^n$ that minimises the Ising energy function.

III.5. The quantum annealing algorithm

Quantum annealing uses an analog process to find optimal and near-optimal solutions to the energy function $H(S)$. A quantum annealing algorithm consists in four components

- an *initial Hamiltonian* $\mathcal{H}_i$, which describes initial conditions;
- the *problem Hamiltonian* $\mathcal{H}_p$ described above;
- a pair of *path functions* $A(s)$ and $B(s)$ that control the transition from $\mathcal{H}_i$ to $\mathcal{H}_p$ over a time interval $s : 0 \rightarrow 1$ (in current D-wave's systems, theses functions are related by $B(s) = 1 - A(s)$;
- a parameter t_a that specifies the total time for the transition (in the microsecond range).

III.5. The quantum annealing algorithm

Quantum annealing uses an adiabatic quantum evolution approach to approximate solutions of the energy function $H(S)$. This is done by traversing from the ground state of an initial Hamiltonian $\mathcal{H}_i$ to a ground state of a final Hamiltonian $\mathcal{H}_p$. According to this scheme, a time dependent Hamiltonian is defined as

$$\mathcal{H}(t) = A(\tau)\mathcal{H}_i + B(\tau)\mathcal{H}_p,$$

where $\tau = t/t_a$ for $0 \leq t \leq t_a$ and t_a is the total annealing time. Usually, the ground state of the initial state $\mathcal{H}_i$ is easy to prepare and the ground state of the final Hamiltonian $\mathcal{H}_p$ codifies the solution of our problem. $A(s)$ and $B(s)$ are chosen such that at time $\tau = 0$, $\mathcal{H}_i$ is predominant in $\mathcal{H}(0)$. As time evolution goes from $\tau = 0$ to $\tau = 1$, the influence of $\mathcal{H}_p$ increases while $\mathcal{H}_i$ fades away.

III.5. The quantum annealing algorithm

The mean idea of quantum annealing is that it is possible to prepare qubits in the ground state of the known and chosen Hamiltonian $\mathcal{H}_i$. Then, parameters defining $\mathcal{H}(t)$ evolves in time, slow enough so the global wavefunction of the qubits evolves adiabatically, and stay in the ground state of the instantaneous Hamiltonian $\mathcal{H}(t)$. At the end of the process, the wavefunction obtained is expected to be the ground state: one just has to measure the state of each qubit to reconstruct the corresponding state.

An important question is how slow the evolution needs to be in order to assure adiabaticity. According to the adiabatic theorem, a quantum mechanical system subjected to gradually changing external conditions adapts its functional form changes occurs slowly enough.

III.5. The quantum annealing algorithm

For a non-degenerate spectrum with a gap between the ground state and first excited state, the adiabatic evolution is assured if the evolution time τ satisfies the following condition

$$\tau \gg \frac{\max_{0\leq t\leq\tau}\left[\left|\left\langle\phi_0(t)\left|\frac{\mathrm{d}\mathcal{H}(t)}{\mathrm{d}t}\right|\phi_1(t)\right\rangle\right|\right]}{\min_{0\leq t\leq\tau}\left[\Delta^2(t)\right]},$$

where $|\phi_0(t)\rangle$ and $|\phi_1(t)\rangle$ being respectively the instantaneous ground state and the first excited state of the total Hamiltonian $\mathcal{H}(t)$ and $\Delta(t)$ the instantaneous gap between the ground state and the first excited state energies.

III.6. Quantum annealing based calculation

Quantum annealing is the key component of a computation made of four steps

Programming/Initialisation

The weights (h_i, J_{ij}) are loaded onto the control system and qubits are placed in an initial superposition state according to $\mathcal{H}_i$.

Annealing

Adiabatic transition from $\mathcal{H}_i$ to $\mathcal{H}_p$ over a time t_a.

III.6. Quantum annealing based calculation

Readout

At the end of the transition, qubits have states according to $\mathcal{H}_p$ which matches $E(s)$. Qubits values are read, yielding solution S to the input.

Resampling

Since any quantum computation is probabilistic, there is always a non negligible probability that the computation does not finish in the ground state of $\mathcal{H}_p$. Given the relatively high initialisation times, it is cost-effective to repeat the anneal-readout cycle many times per input.

III.6. Quantum annealing based calculation

In D-wave systems, the initial Hamiltonian is fixed, but the problem Hamiltonian, annealing time t_a and the number of resampling steps R are supplied by the user. The total calculation time $T(R)$ required by the QPU to return a sample of R solutions to one input instance is then

$$T(R) = t_{\text{program}} + R(t_a + t_{\text{read}})$$

For a D-wave 2000Q systems, typical values are

- $t_{\text{program}} = 9$ ms
- $t_a \geq 5\,\mu$s
- $t_{\text{read}} = 120\,\mu$s

It results in a calculation time for 1000 solutions of typically 149 ms. The annealing step is just a tiny fraction of $T(R)$.

Decoherence, NISQ computers and quantum supremacy

1. Density matrix - T_1 and T_2 times and decoherence
2. Quantum advantage, quantum supremacy
3. Quantum annealer
   - Quantum annealing processor
   - Optimization problems
   - D-wave quantum processors
   - The user's view
   - The quantum annealing algorithm
   - Quantum annealing based calculation
   - Advantages and limitations of quantum annealing
   - The traveling Salesman Problem

III.7. Advantages and limitations of quantum annealing

- Quantum annealer isn't a general-purpose computer, in that it can only solve a set of problems that can be structured as energy minimizations. Contrary to universal quantum computers, they are limited to specific application, which might be described formally as an Ising problem.
- Quantum annealing chips are not design to achieve multiple qubits gates manipulations and consequently they are not limited by the corresponding error rate.
- Due to its adiabatic nature, and given that a quantum system naturally relaxes towards its ground states, quantum annealing is for more tolerant to noise and errors on qubits. That is why QPU have quite more qubits in the case of quantum annealing, but it is not possible to programm a quantum circuit on such chip.

III.7. Advantages and limitations of quantum annealing

By the end of 2019, the most powerful system commercially available from D-wave is the D-wave 2000Q system, released in January 2017.

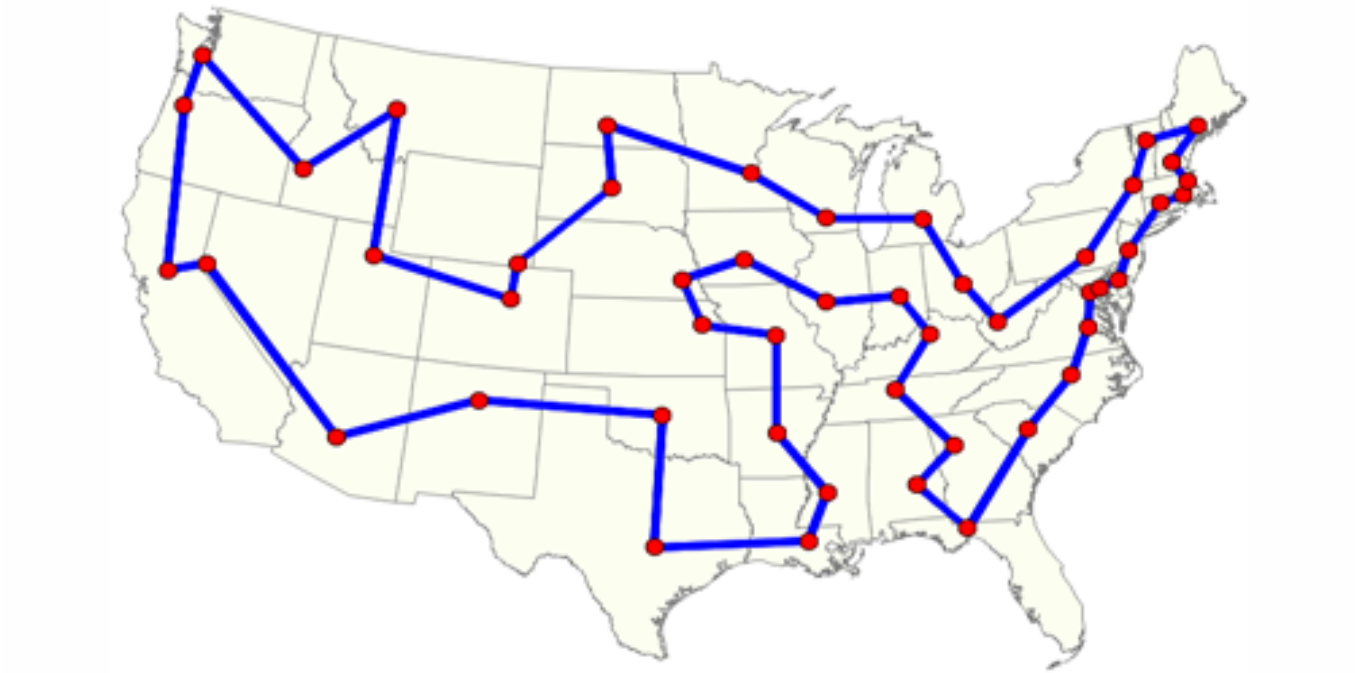
- 2048 qubits (while programmable computers are still in the 50-100 qubits range), and 6,016 couplers;
- 128,472 Josephson junctions in total on the chip, with 200 I/O lines;
- Operation in a dilution He cryostat (15 mK), total power consumption of 25 kW.

III.7. Advantages and limitations of quantum annealing

In February 2019 D-Wave announced their next-generation Pegasus P16 quantum processor chip, announcing that it would be "the world's most connected commercial quantum system." The next-generation system would use the Pegasus P16 chip.

- 5,640 qubits and reduced noise;
- 40,484 couplers and 1,030,000 Josephson junctions;
- 15 connections per qubit instead of 6;
- should be available in mid-2020.

III.8. The traveling Salesman Problem



Quantum speedup of the traveling-salesman problem for bounded-degree graphs, Dominic J. Moylett, Noah Linden, and Ashley Montanaro Phys. Rev. A 95, 032323 (2017)

`https://journals.aps.org/pra/abstract/10.1103/PhysRevA.95.032323`

III.8. The traveling Salesman Problem

An instance of the problem is given by a set

$$\{d_{ij};\ i,j \in [\![1,N]\!]\}$$

where d_{ij} corresponds to the distance between the i-th and the j-th city. One assumes $d_{ij} = d_{ji}$.

The problem can be cast into the form where one minimizes an Ising Hamiltonian under some constraints. A tour can be represented by an $N \times N$ matrix $\mathcal{T}$ with elements either 0 or 1. In a given tour, if the city j is visited immediately after visiting city i, then $\mathcal{T}_{ij} = 1$, otherwise $\mathcal{T}_{ij} = 0$. Generally an additional constraint is imposed that one city has to be visited once and only once in a tour. Any valid tour with the above restriction may be represented by a $\mathcal{T}$ matrix whose each row and each column has one and only one element equal to 1 and rest all are 0s.

III.8. The traveling Salesman Problem

For a symmetric metric, a tour and its reverse tour have the same length, and it is more convenient to work with an undirected tour matrix

$$\mathcal{U} = \frac{1}{2}\left(\mathcal{T} + \mathcal{T}^T\right)$$

where $\mathcal{T}^T$, the transpose of $\mathcal{T}$, represents the reverse of the tour given by $\mathcal{T}$.
In terms of $\mathcal{U}_{ij}$s, the length of a tour can be represented by

$$\mathcal{H} = \frac{1}{2}\sum_{i,j=1}^{N} d_{ij}\mathcal{U}_{ij}.$$

III.8. The traveling Salesman Problem

One can rewrite the above Hamiltonian in terms of Ising spins $\mathcal{S}_{ij}$s as

$$\mathcal{H}_{\text{TSP}} = \frac{1}{2} \sum_{i,j=1}^{N} d_{ij} \frac{1 + \mathcal{S}_{ij}}{2}.$$

where $\mathcal{S}_{ij} = 2\mathcal{U}_{ij} - 1$ are the Ising spins. The Hamiltonian is similar to that of a non-interacting Ising spins on a $N \times N$ lattice, with random fields d_{ij} on the lattice points $\{i, j\}$.