Tutorial Webpage

Quantum Data Management

From Theory to Opportunities

Rihan Hai, Shih-Han Hung, Sebastian Feld

s2.smu.edu/~mitch/class/5395/papers/feynman-quantum-1981.pdf
en.wikipedia.org/wiki/Richard_Feynman#/media/File:Richard_Feynman_Nobel.jpg

Where it all started

Simulating Physics with Computers

Richard P. Feynman

Department of Physics, California Institute of Technology, Pasadena, California 91107

Received May 7, 1981

1. INTRODUCTION

On the program it says this is a keynote speech—and I don't know what a keynote speech is. I do not intend in any way to suggest what should be in this meeting as a keynote of the subjects or anything like that. I have my own things to say and to talk about and there's no implication that anybody needs to talk about the same thing or anything like it. So what I want to talk about is what Mike Dertouzos suggested that nobody would talk about. I want to talk about the problem of simulating physics with computers and I mean that in a specific way which I am going to explain. The reason for doing this is something that I learned about from Ed Fredkin, and my entire interest in the subject has been inspired by him. It has to do with learning something about the possibilities of computers, and also something about possibilities in physics. If we suppose that we know all the physical laws perfectly, of course we don't have to pay any attention to computers. It's interesting anyway to entertain oneself with the idea that we've got something to learn about physical laws; and if I take a relaxed view here (after all I'm here and not at home) I'll admit that we don't understand everything.

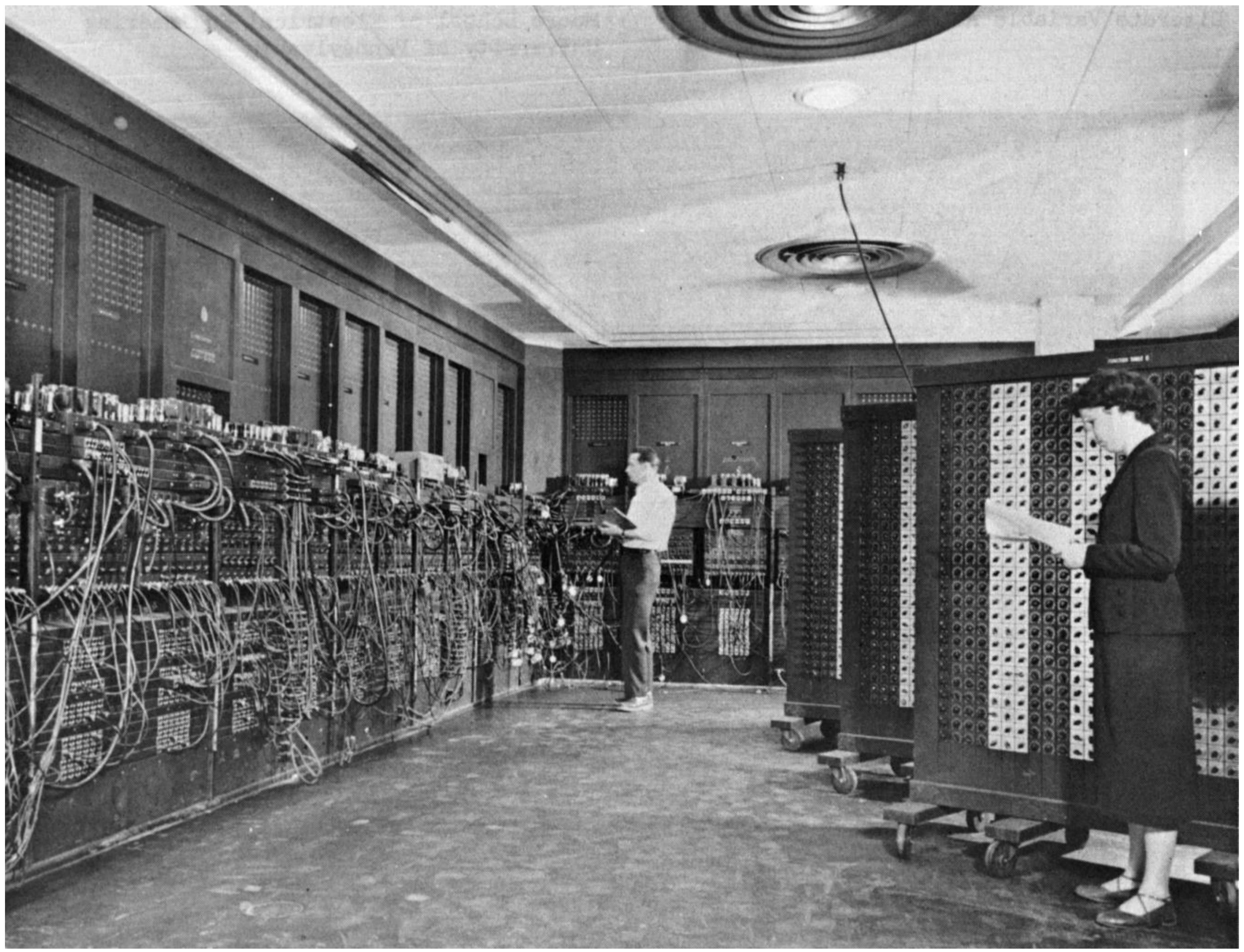
The first question is, What kind of computer are we going to use to simulate physics? Computer theory has been developed to a point where it realizes that it doesn't make any difference; when you get to a *universal computer*, it doesn't matter how it's manufactured, how it's actually made. Therefore my question is, Can physics be simulated by a universal computer? I would like to have the elements of this computer *locally interconnected*, and therefore sort of think about cellular automata as an example (but I don't want to force it). But I do want something involved with the

0020-7748/82/0600-0467$03.00/0 © 1982 Plenum Publishing Corporation

be understood very well in analyzing the situation. And I'm not happy with all the analyses that go with just the classical theory, because nature isn't classical, dammit, and if you want to make a simulation of nature, you'd better make it quantum mechanical, and by golly it's a wonderful problem, because it doesn't look so easy. Thank you.

en.wikipedia.org/wiki/ENIAC#/media/File:Glen_Beck_and_Betty_Snyder_program_the_ENIAC_in_building_328_at_the_Ballistic_Research_Laboratory.jpg

Back to the future



nature.com/articles/s41586-019-1666-5
journals.aps.org/pra/abstract/10.1103/PhysRevA.106.032428
writings.stephenwolfram.com/2018/04/buzzword-convergence-making-sense-of-quantum-neural-blockchain-ai/

Lots of activities

Article | Published: 23 October 2019

Quantum supremacy using a programmable superconducting processor

Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerin, ... John M. Martinis + Show authors

Nature **574**, 505–510 (2019) | Cite this article

Simulating key properties of lithium-ion batteries with a fault-tolerant quantum computer

Alain Delgado, Pablo A. M. Casares, Roberto dos Reis, Modjtaba Shokrian Zini, Roberto Campos, Norge Cruz-Hernández, Arne-Christian Voigt, Angus Lowe, Soran Jahangiri, M. A. Martin-Delgado, Jonathan E. Mueller, and Juan Miguel Arrazola

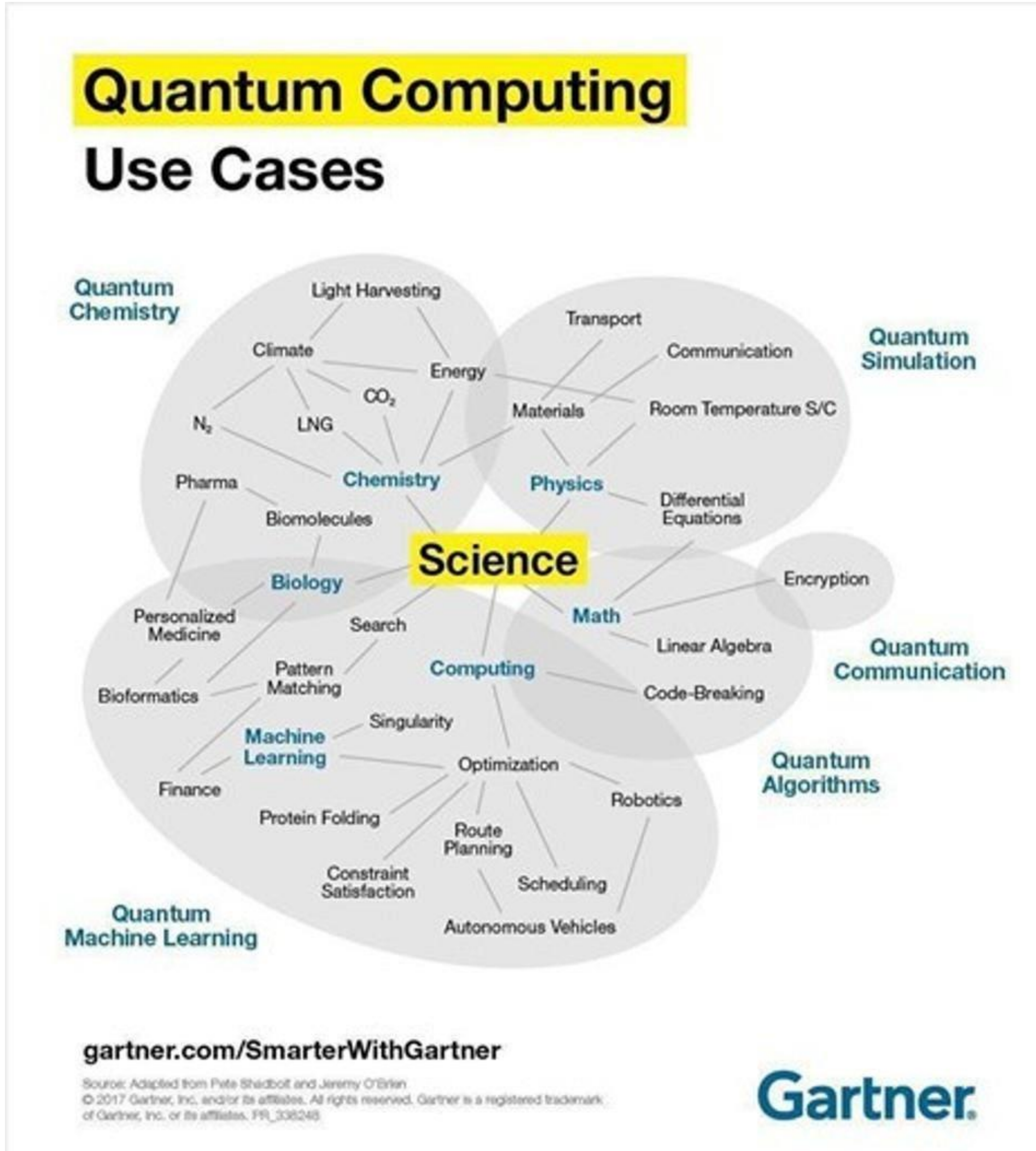
Phys. Rev. A **106**, 032428 – Published 26 September 2022

Buzzword Convergence: Making Sense of Quantum Neural Blockchain AI

April 1, 2018

forbes.com/sites/chuckbrooks/2021/03/21/the-emerging-paths-of-quantum-computing/?sh=66cac7936613

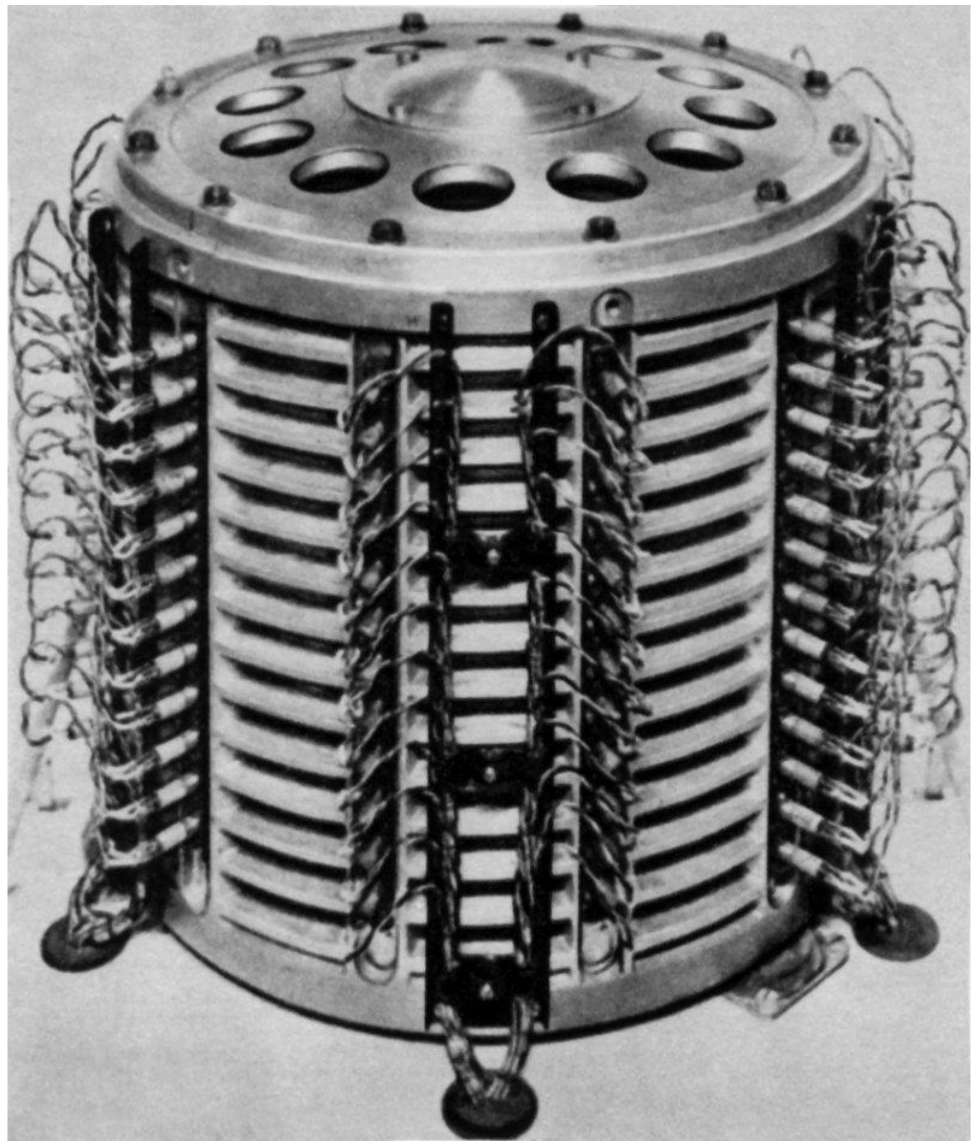
Lots of (potential) use cases



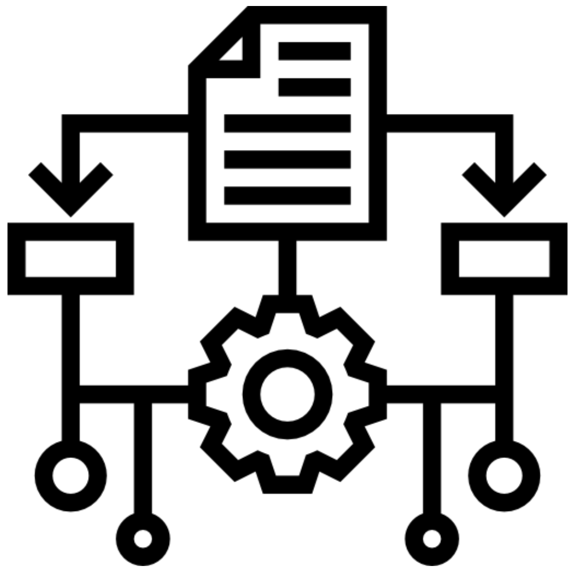
de.wikipedia.org/wiki/Trommelspeicher#/media/Datei:Pamiec_bebnowa_1.jpg
flaticon.com/de/kostenloses-icon/algorithmus_1119005
forbes.com/sites/tiriasresearch/2023/11/28/quantum-computing-is-coming-faster-than-you-think/

Menu for today

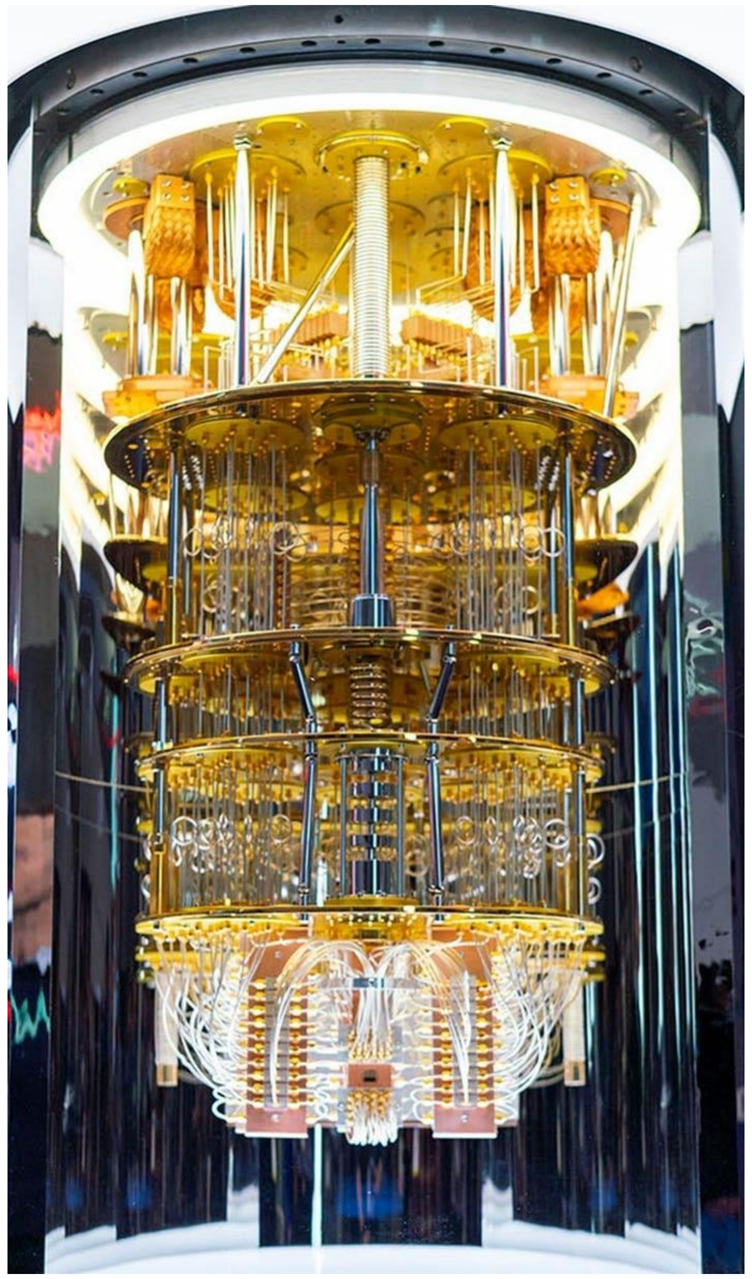
Entering the era of quantum computing, what is the future of data management?



Data management problem



Algorithm



Quantum computer

Contents and instructors

Structure
- Introduction [~5‘]
- Fundamentals of quantum computing [~20‘]
- Data management using quantum computers [~35‘]
- Data management via quantum internet [~30‘]

Rihan Hai
TU Delft, NL

Shih-Han Hung
Academica Sinica, TW

Sebastian Feld
TU Delft, NL

Fundamentals of quantum computing

Fundamentals of quantum computing

- Overview
- Quantum Gate Model
- Quantum Annealing
- Conclusion

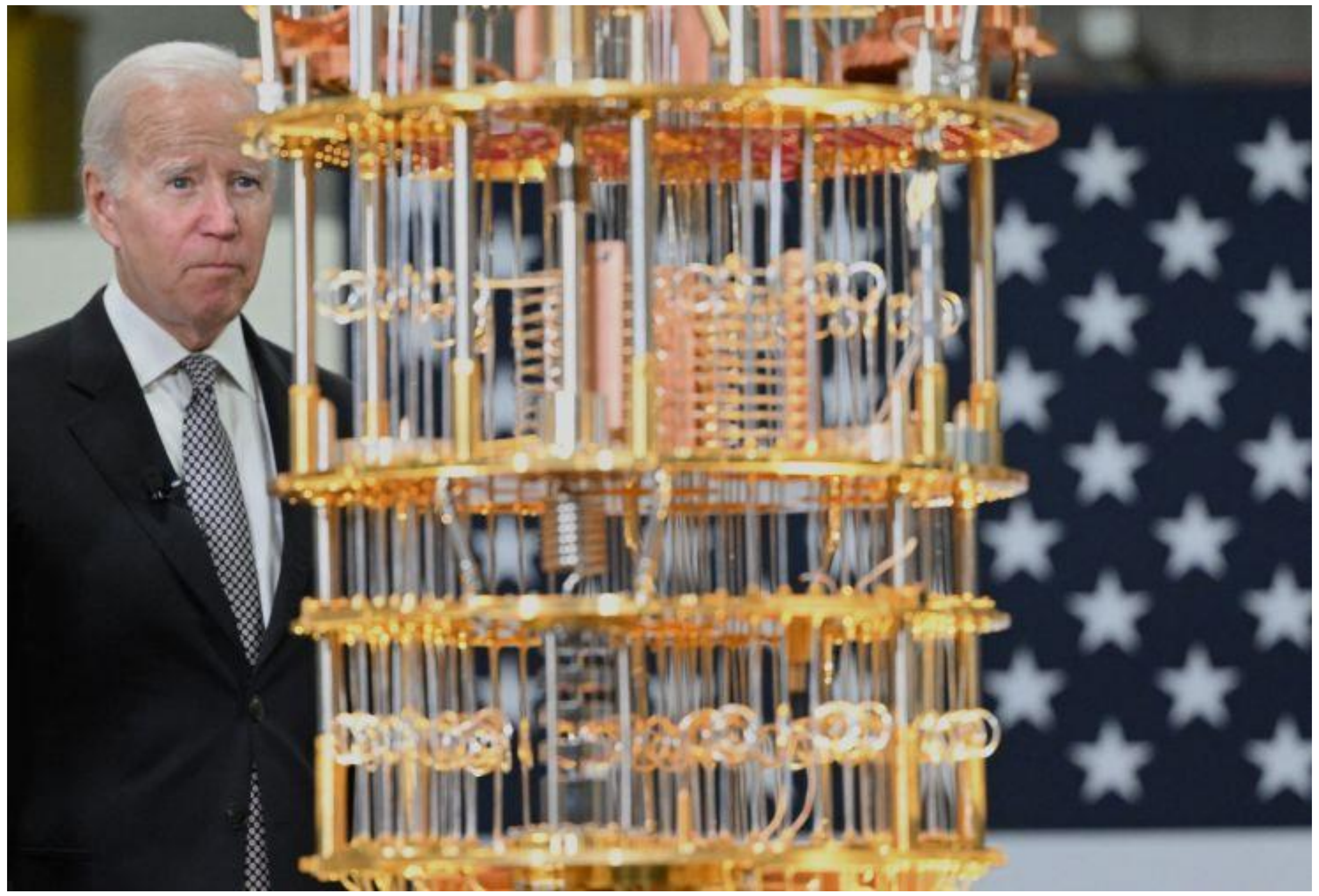
Fundamentals of quantum computing

- **Overview**
- Quantum Gate Model
- Quantum Annealing
- Conclusion

twitter.com/IBMDACH/status/1189194971925975041
em-press.nl/image/I0000k6zUoyh5gdA
foreignpolicy.com/2023/07/31/us-quantum-technology-china-competition-security/

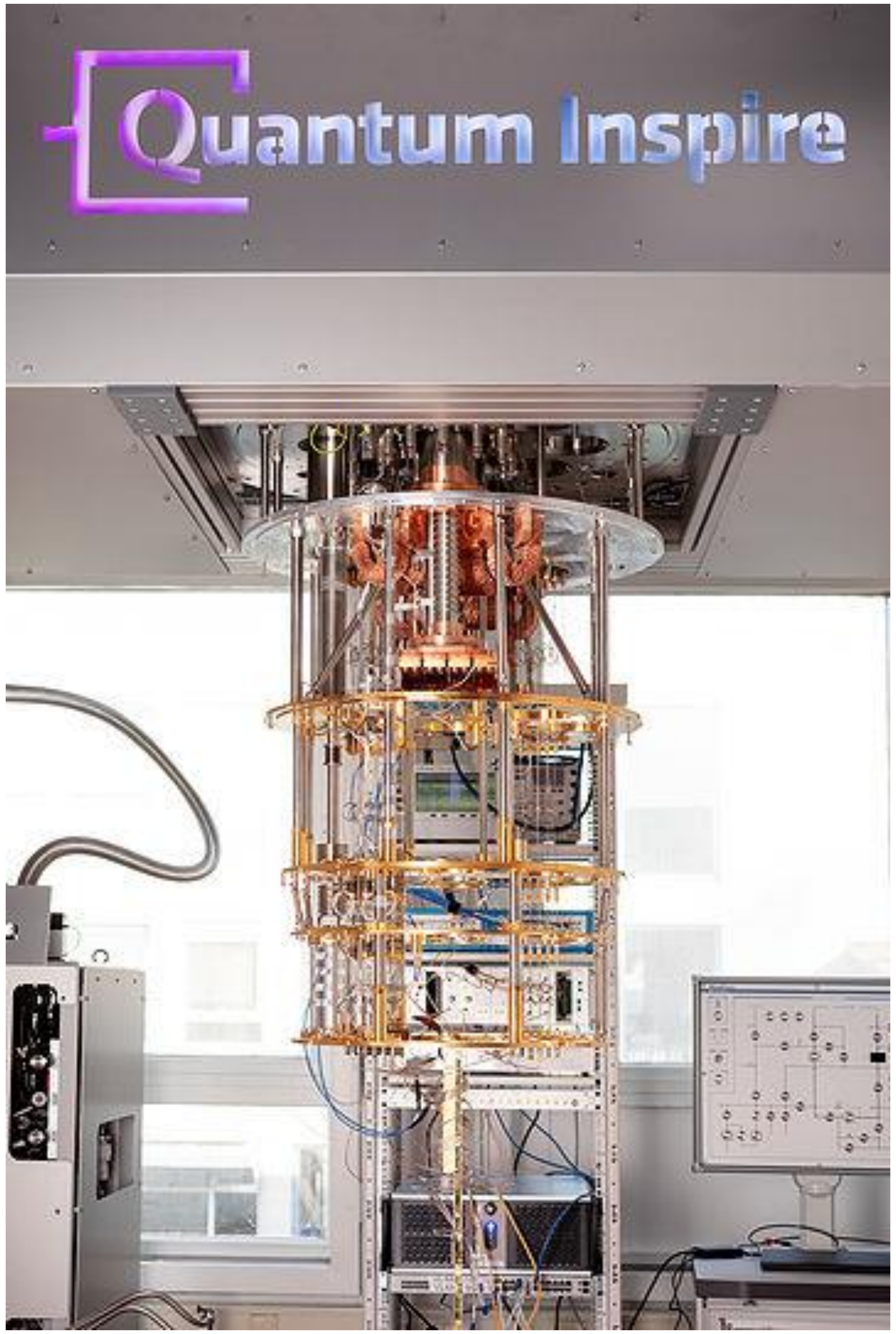
What exactly is a quantum computer?

- Golden chandelier plus VIP



What exactly is a quantum computer?

- Jokes aside: what is included and what is not?



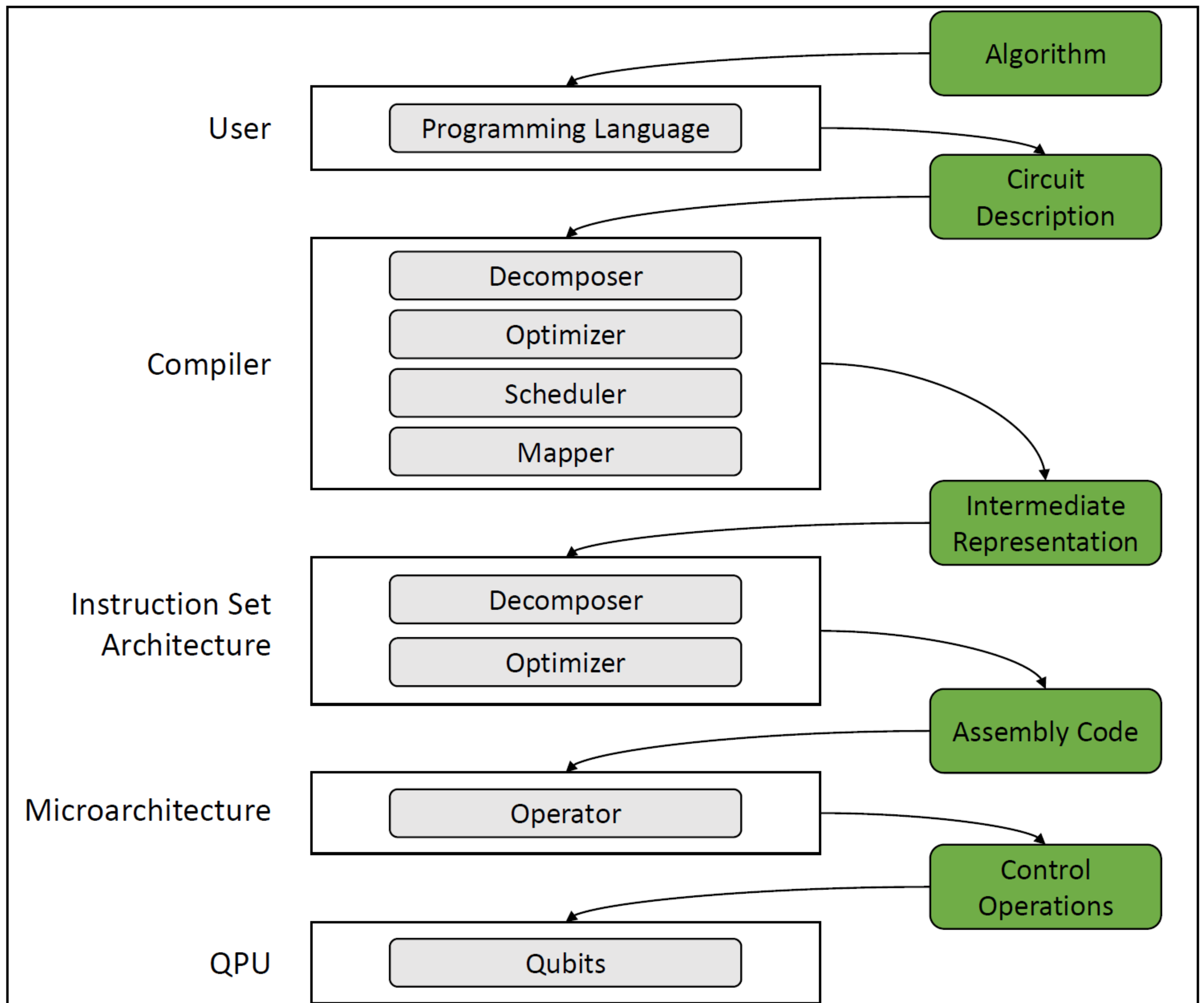
en.wikipedia.org/wiki/Quantum_computing#/media/File:IBM_Q_system_(Fraunhofer_2).jpg

What exactly is a quantum computer?

- Often hidden behind a service

What exactly is a quantum computer?

- Full-stack quantum computation

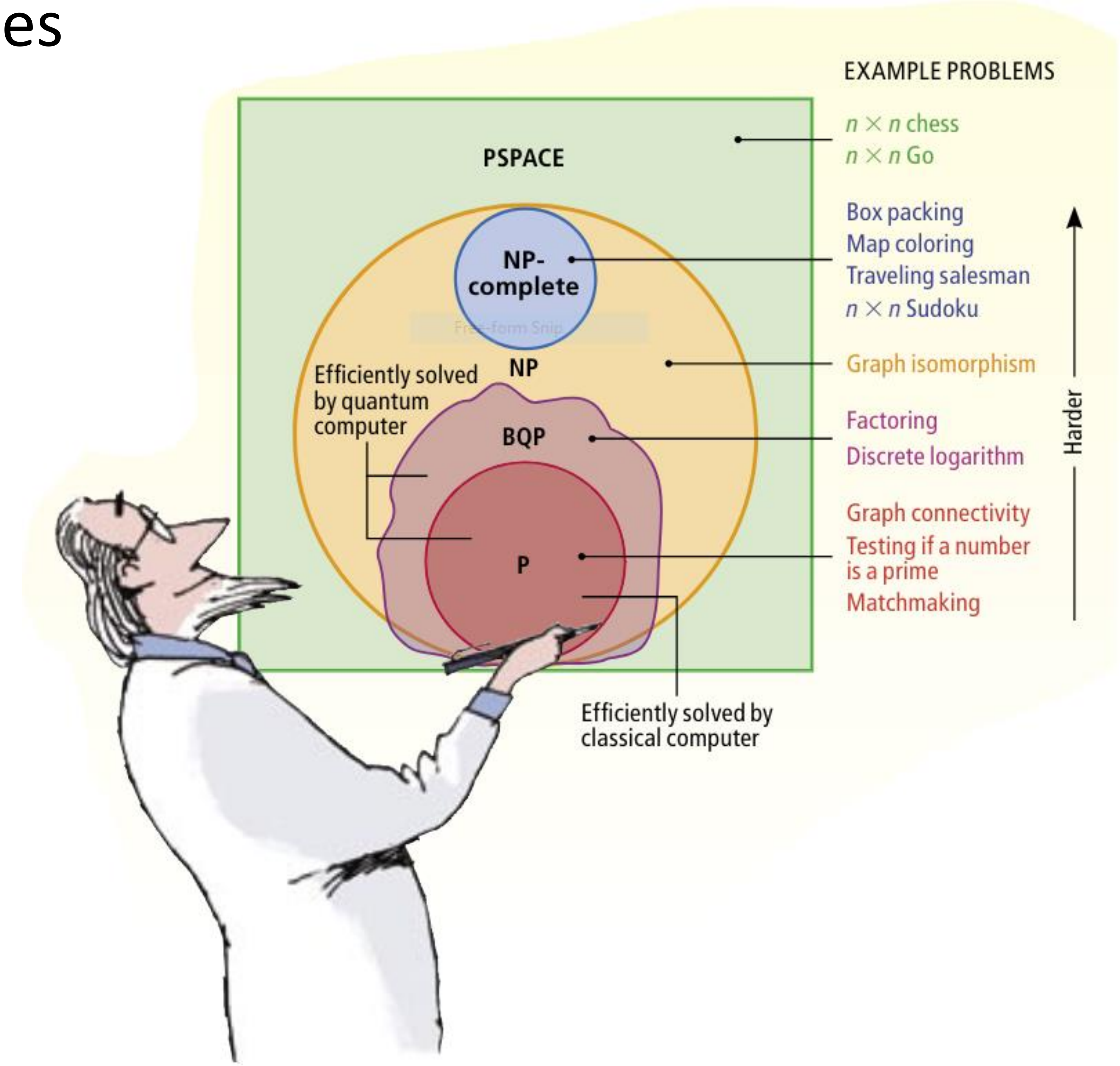


What exactly is a quantum computer?

- It‘s not a magical wonder machine
- Quantum computing is turing-complete

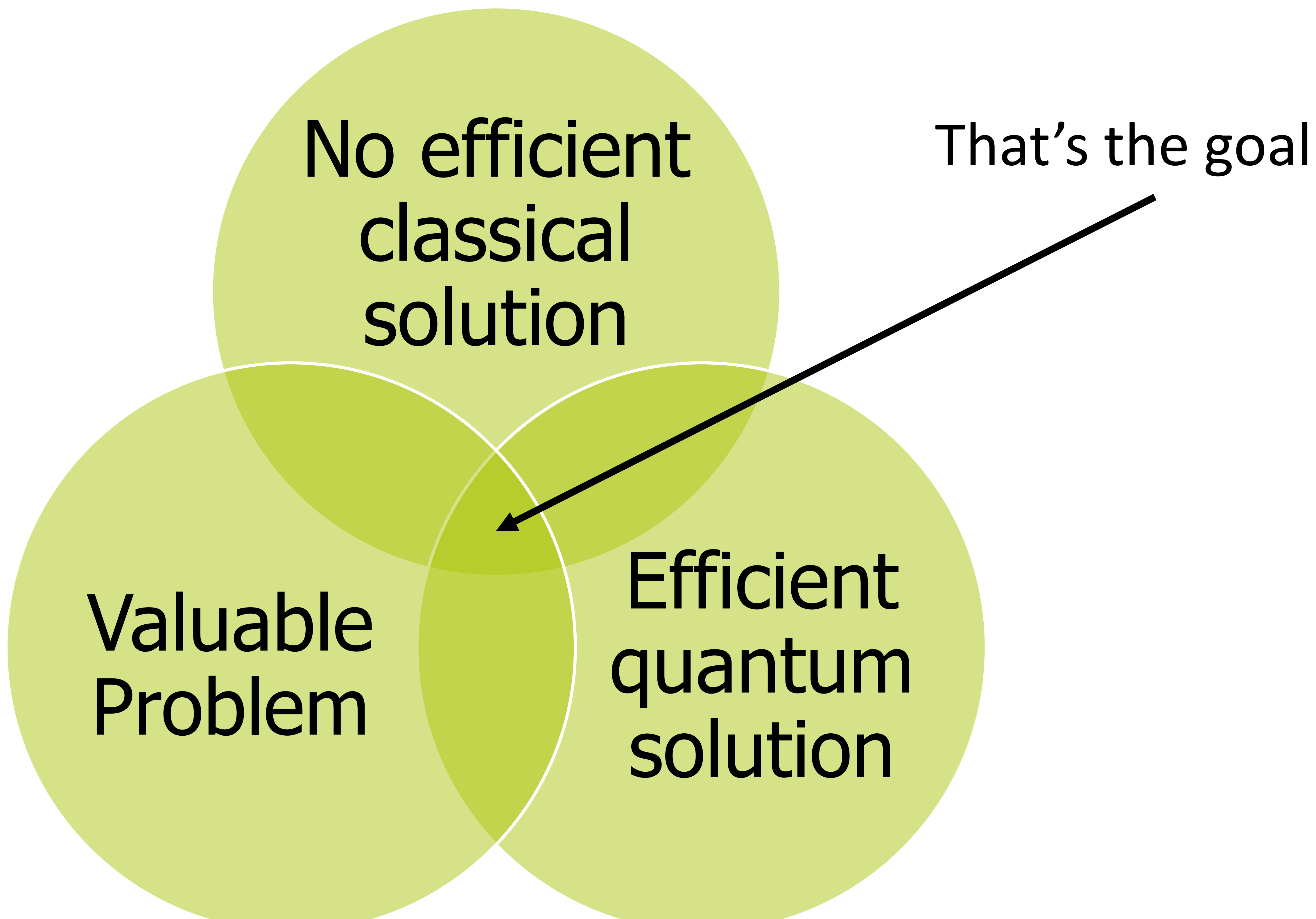
Know your limits

- Quantum computing is powerful, and we don‘t even know the boundaries



The holy grail

- At least three things need to come together

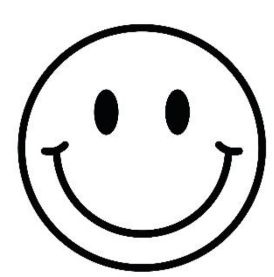


Two different flavors

Quantum Gate Model

TU Delft

Google IBM Microsoft …

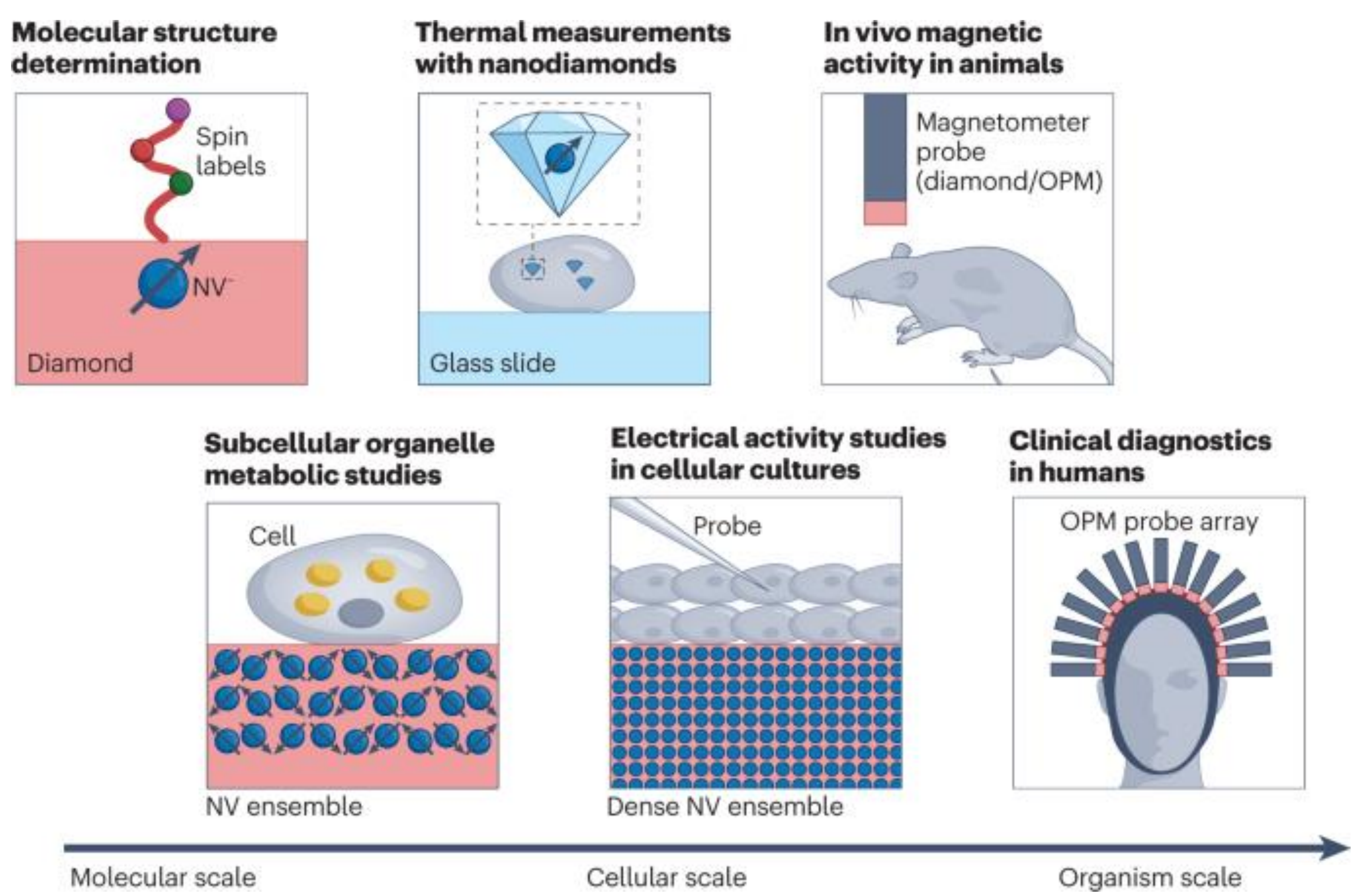


Quantum Annealing

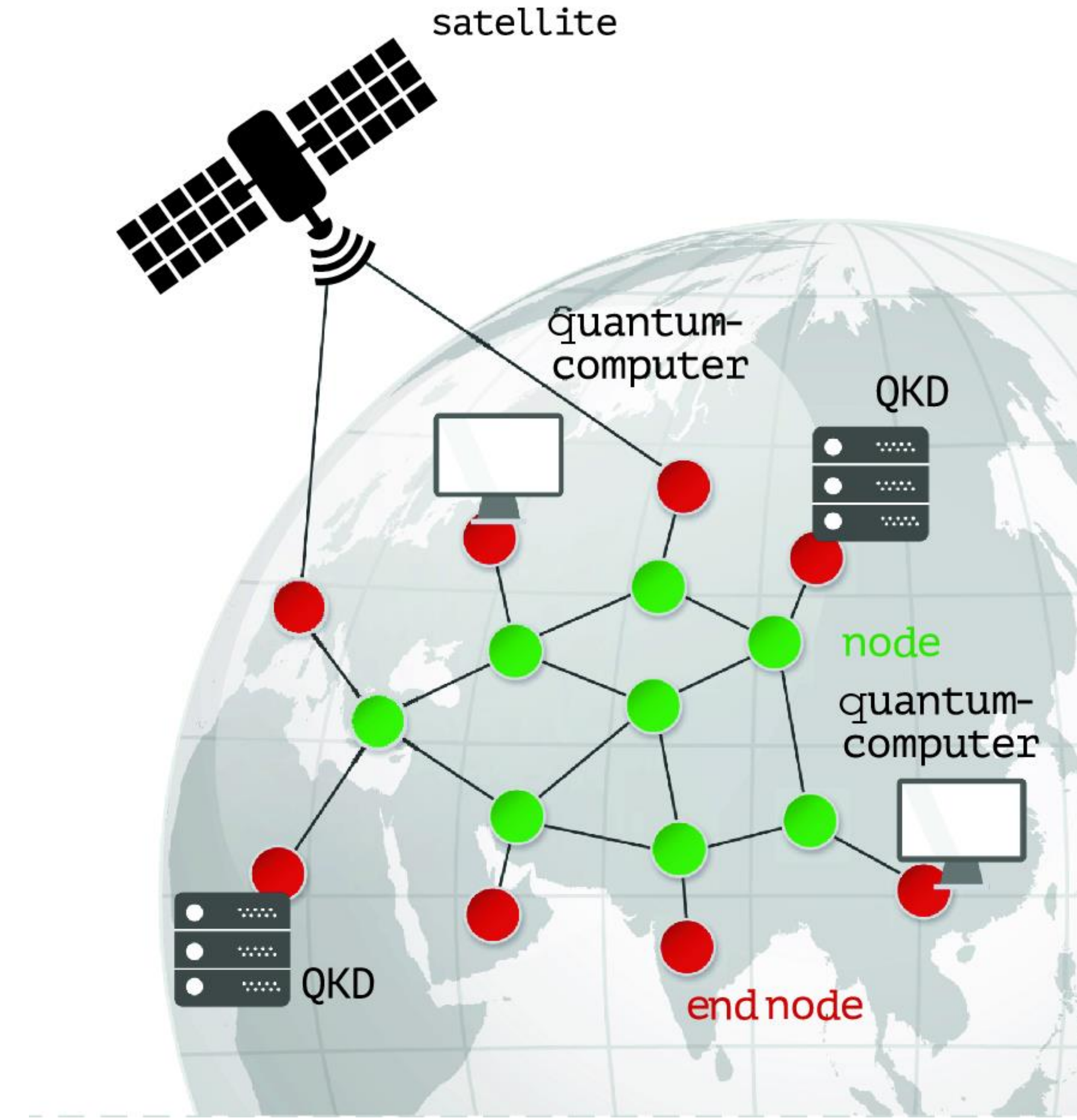
D-Wave
The Quantum Computing Company™

And what else?

Quantum Sensing



Quantum Internet



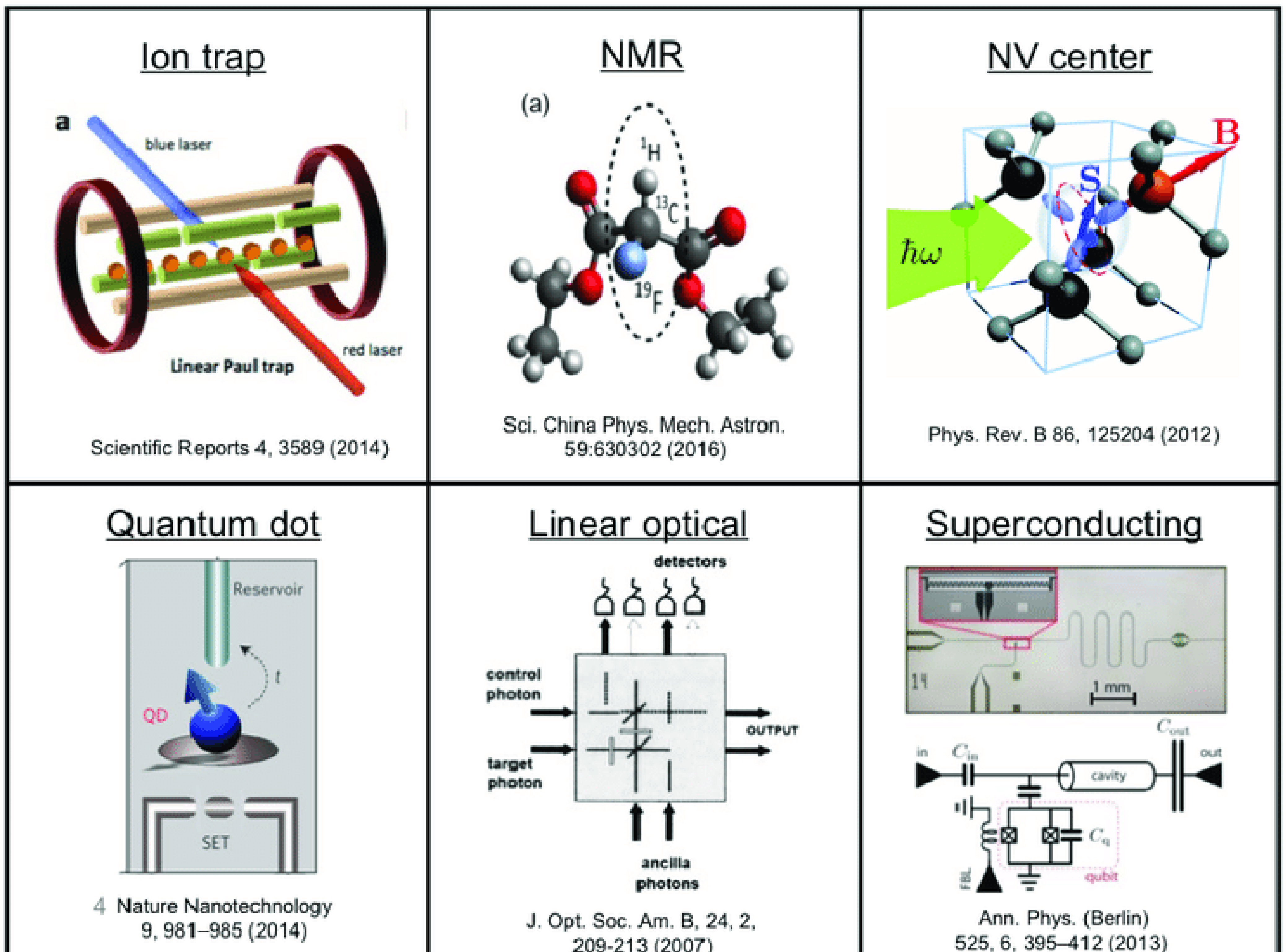
Fundamentals of quantum computing

- Overview
- **Quantum Gate Model**
- Quantum Annealing
- Conclusion

epj-conferences.org/articles/epjconf/pdf/2019/19/epjconf_chep2018_09010.pdf

Many different qubit technologies

- Quantum bit (qubit), the main building block



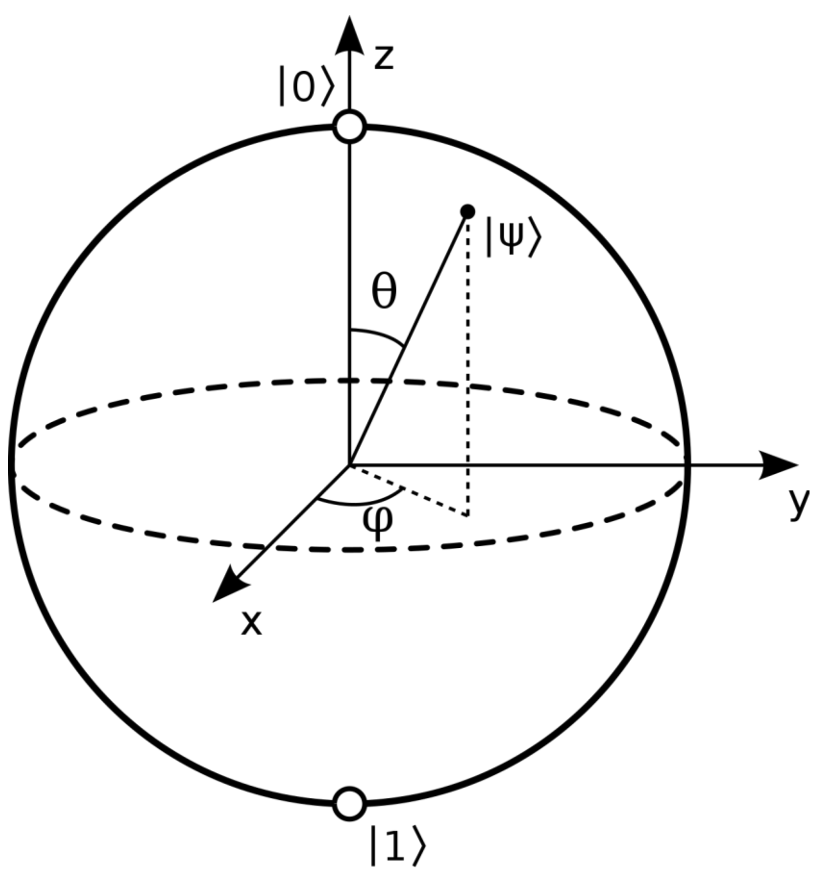
…
many
more

What is a qubit (for us)?

- A qubit is a linear combination of basis states

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

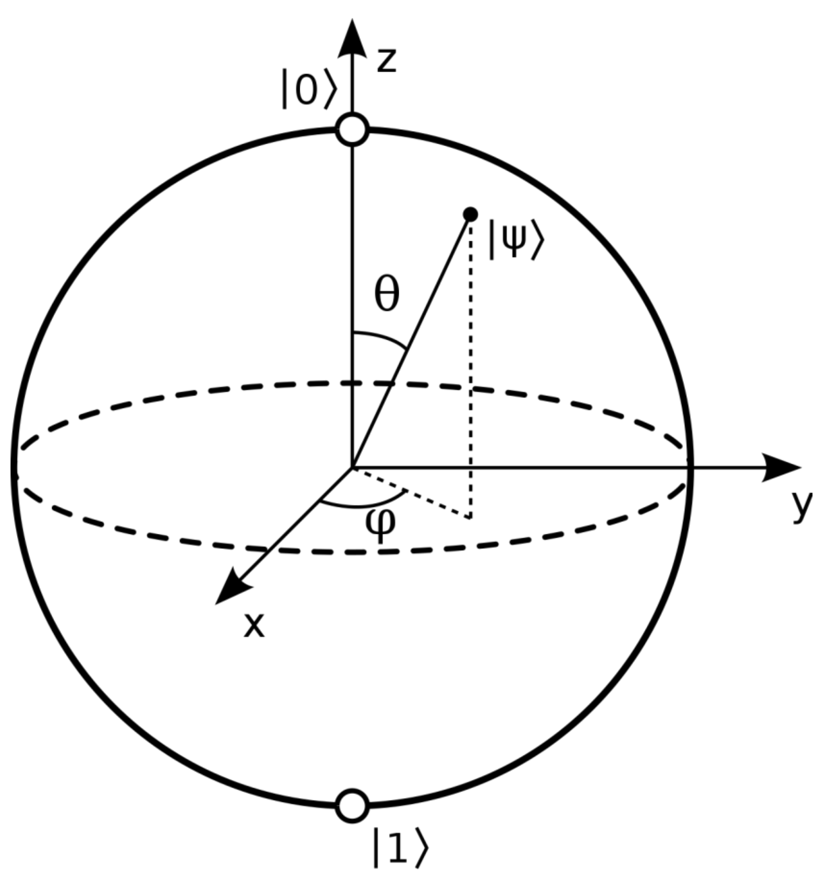
$$\alpha, \beta \in \mathbb{C} \text{ with } |\alpha|^2 + |\beta|^2 = 1$$



en.wikipedia.org/wiki/Bloch_sphere
en.m.wikipedia.org/wiki/File:Qcircuit_measure-arrow.svg

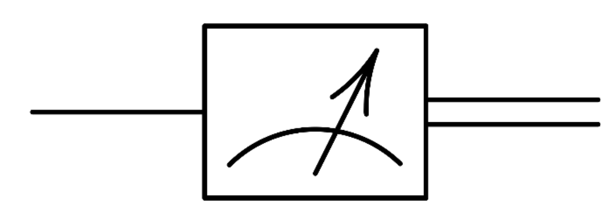
Measurement gives classical information

- α, β are called probability amplitudes
- When measuring, $|\alpha|^2$ is probability of finding qubit in state $|0\rangle$



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

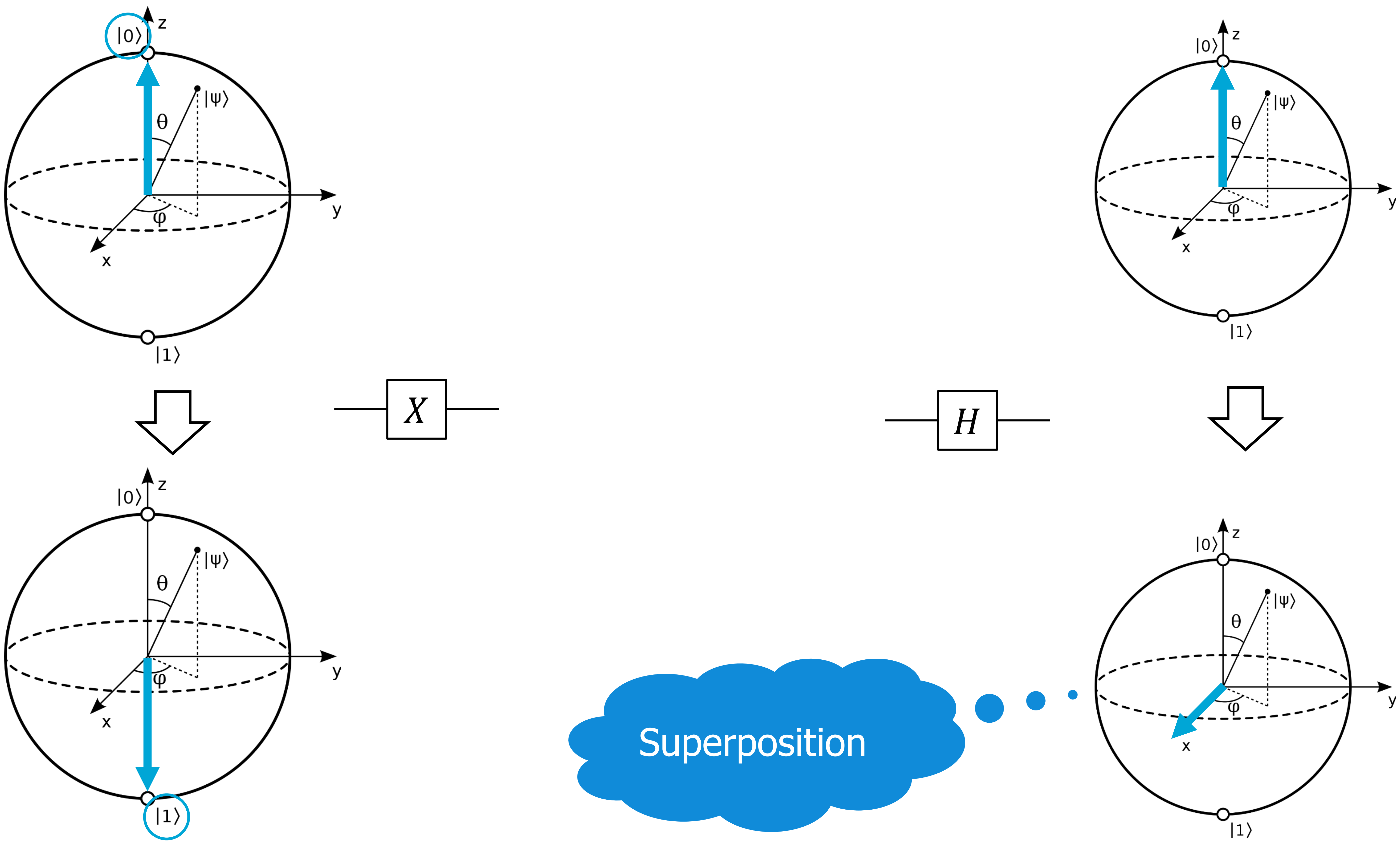
$$\alpha, \beta \in \mathbb{C} \text{ with } |\alpha|^2 + |\beta|^2 = 1$$

$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  $|0\rangle$ or $|1\rangle$

en.wikipedia.org/wiki/Bloch_sphere

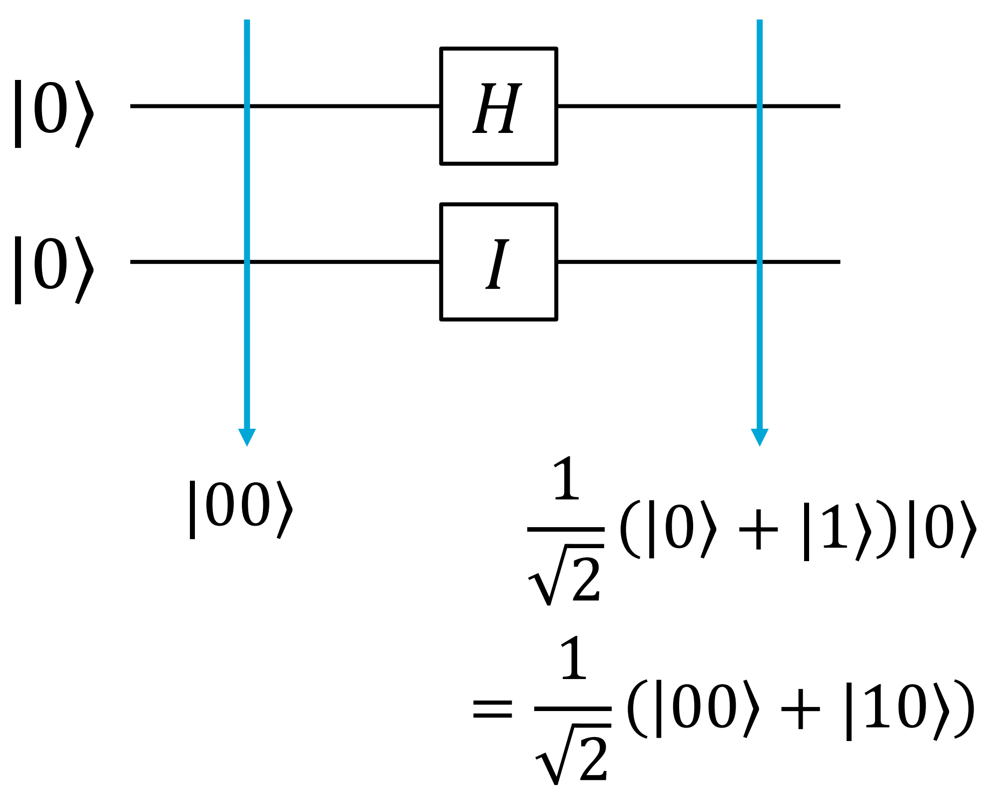
Gates

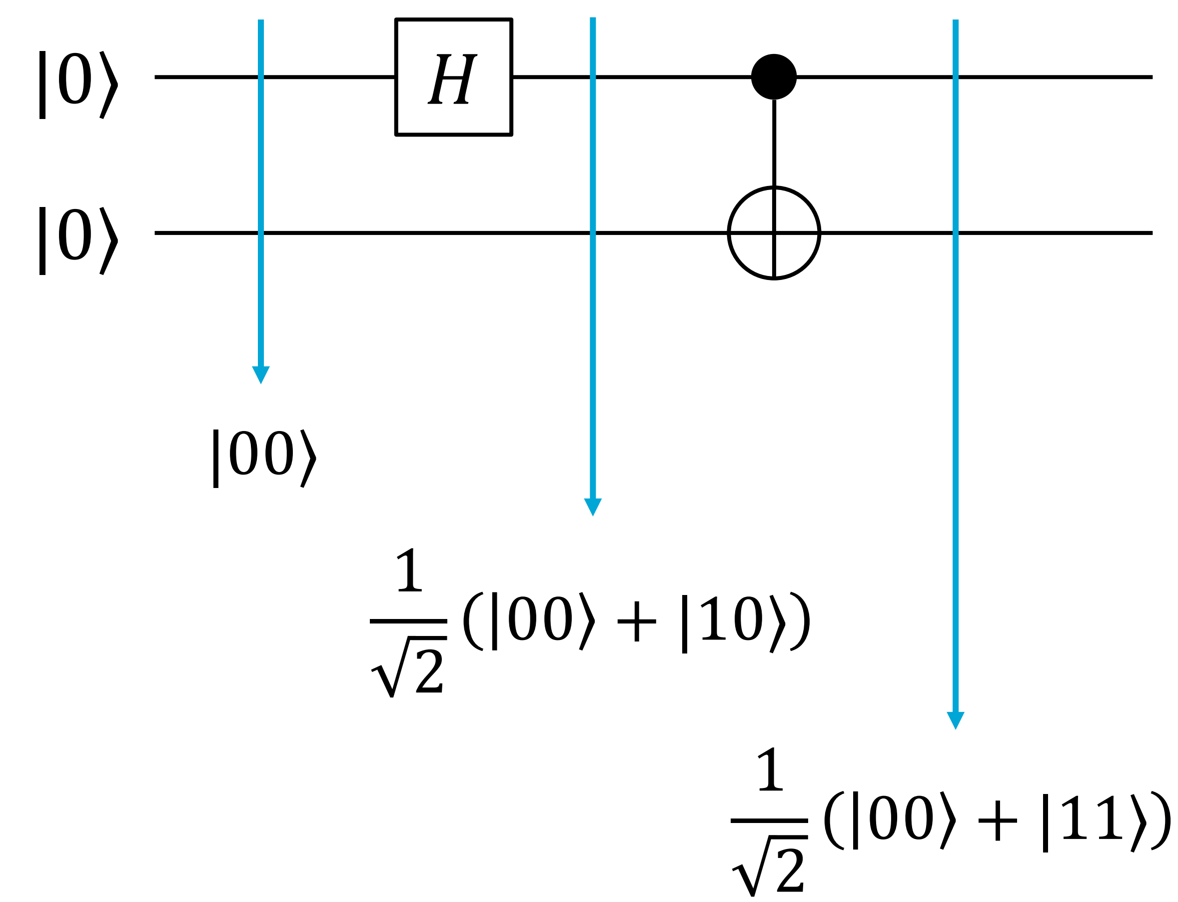
- Quantum gates operate on a quantum state
- Rotations around an axis, but also controlled operations

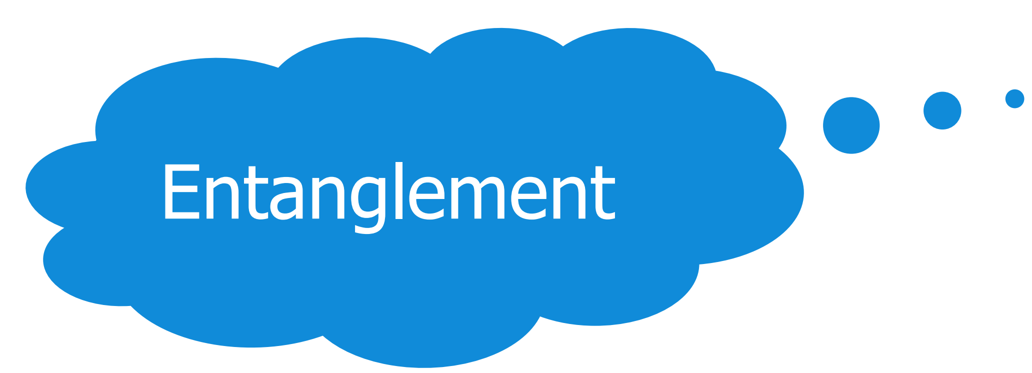


Gates

- Quantum gates operate on a quantum state
- Rotations around an axis, but also controlled operations

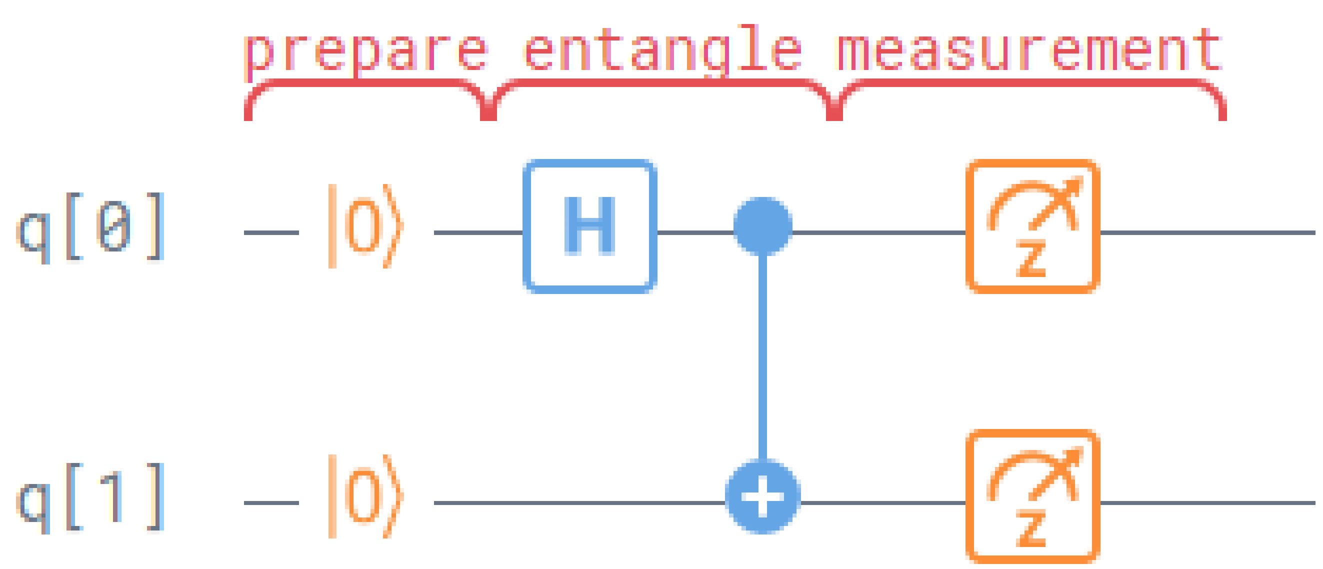






Circuits

- Back to assembly programming → Quantum Assembly (QASM)
- High-level programming languages available, though

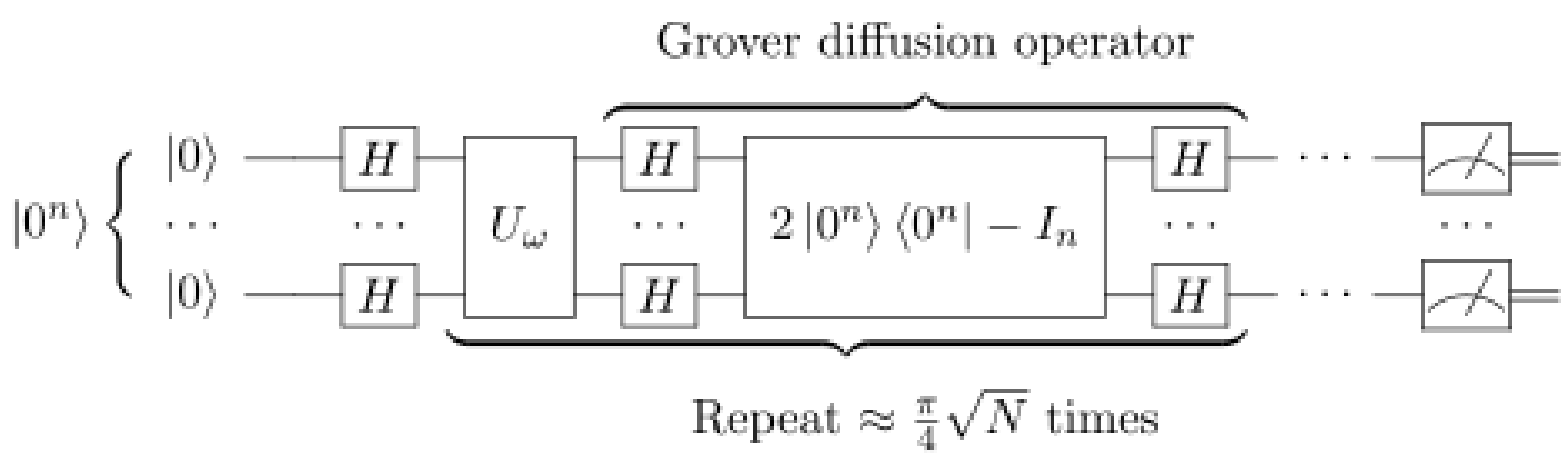


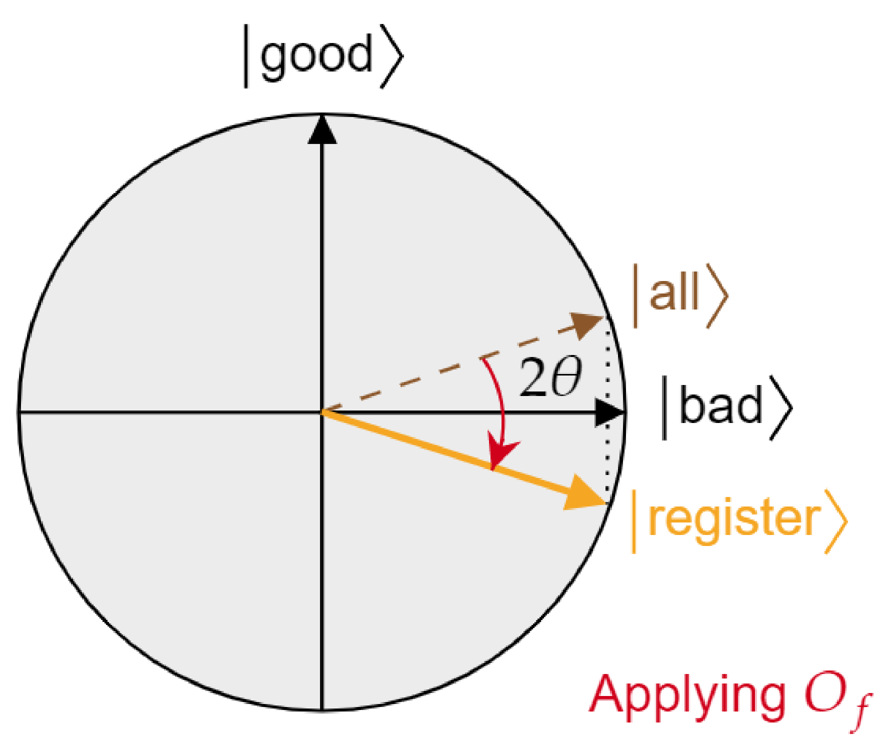

```
version 1.0

qubits 2

.prepare
    prep_z q[0:1]

.entangle
    H q[0]
    CNOT q[0], q[1]

.measurement
    measure_all
```

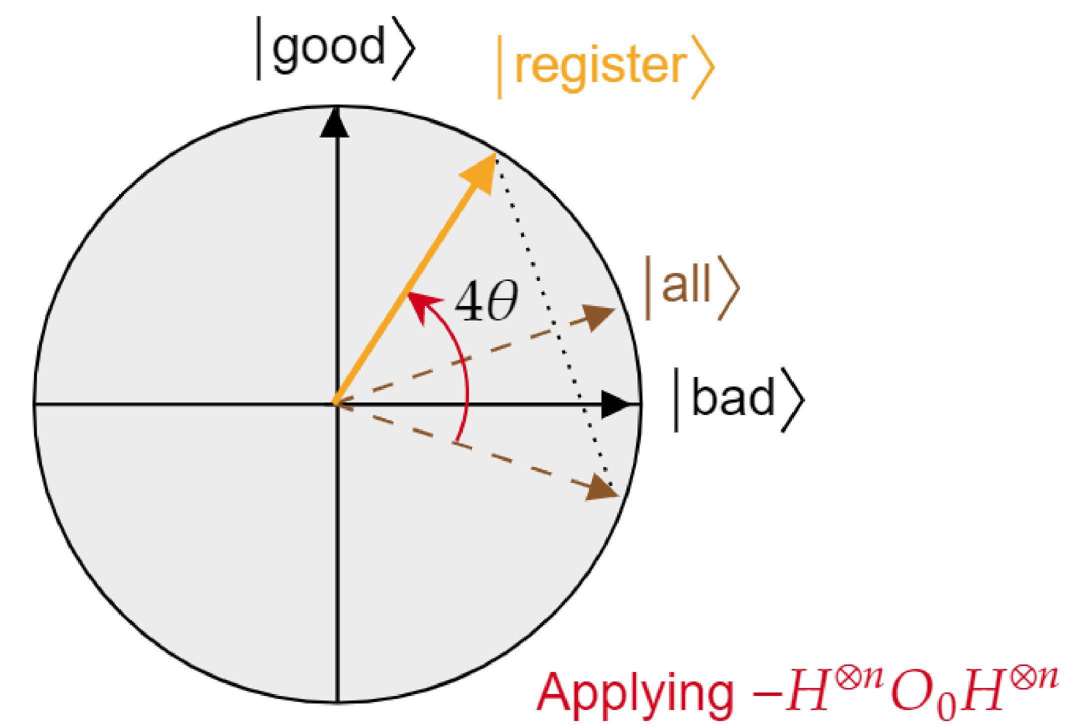
en.wikipedia.org/wiki/Grover's_algorithm
learn.microsoft.com/en-us/azure/quantum/concepts-grovers

General idea of Grover‘s algorithm

- Manipulate the amplitudes, such that the probability of our solution state is maximized

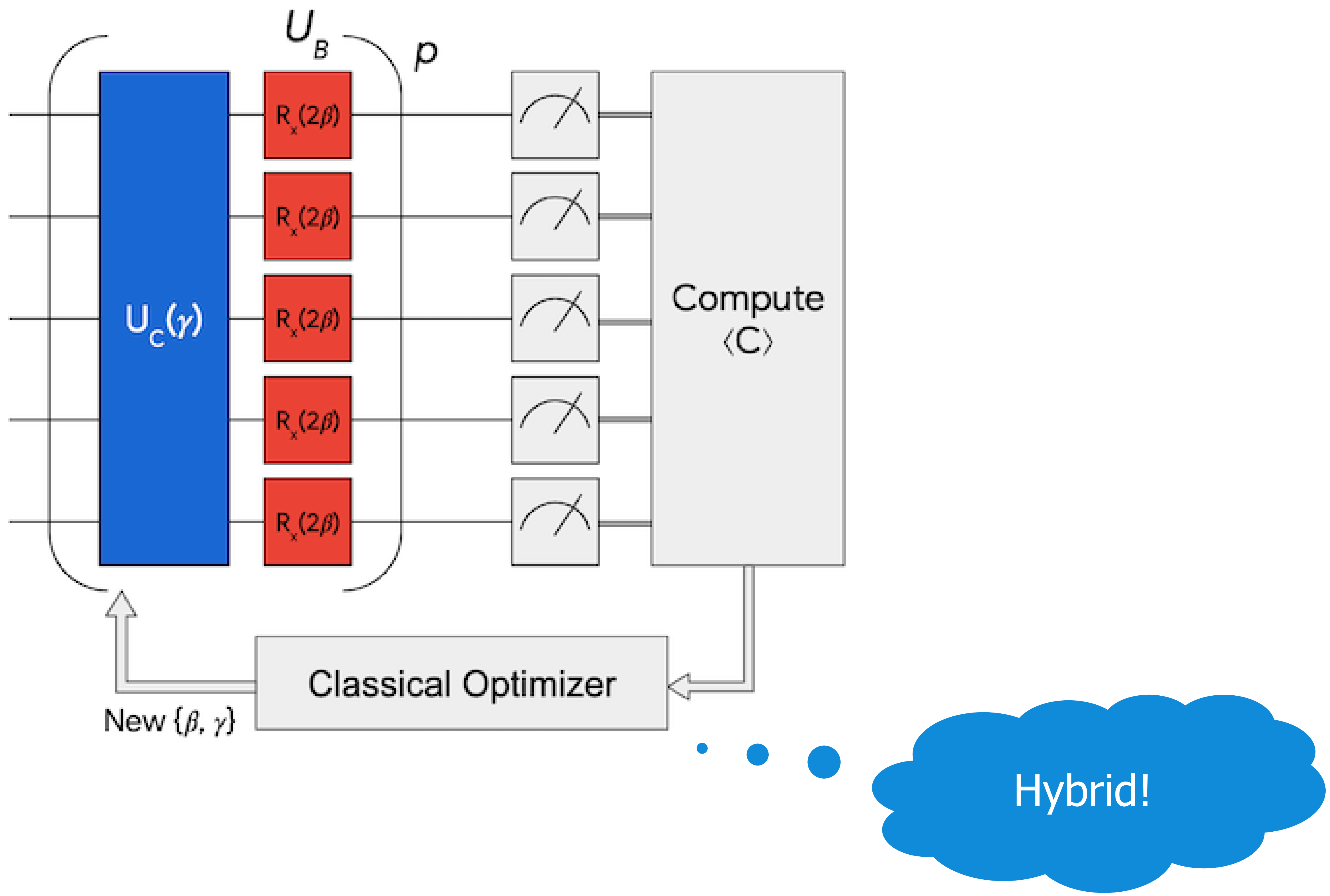






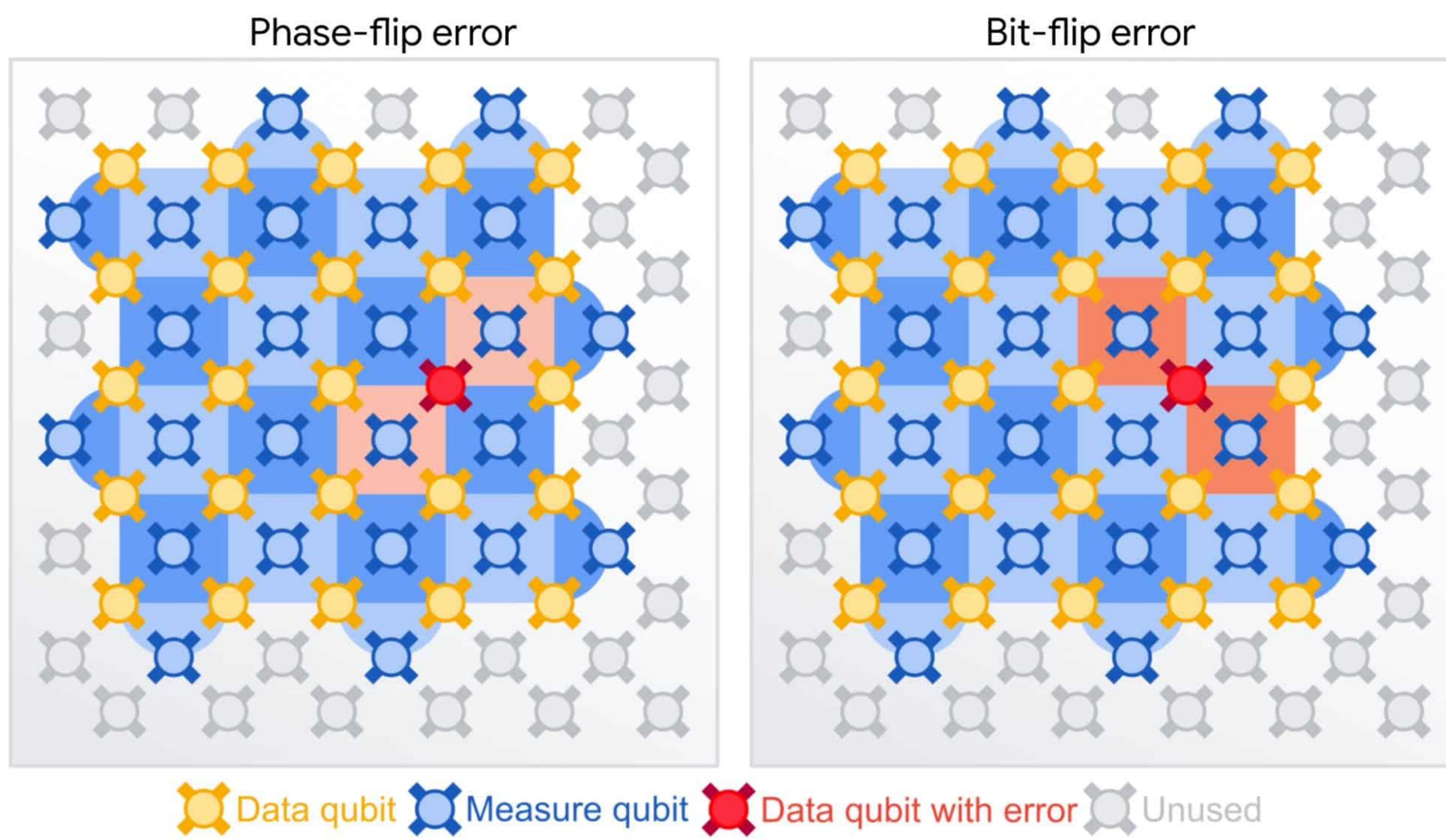
General idea of QAOA (and VQE)

- Minimize a problem-specific quantum operator C by parametrizing β and γ



Main problems

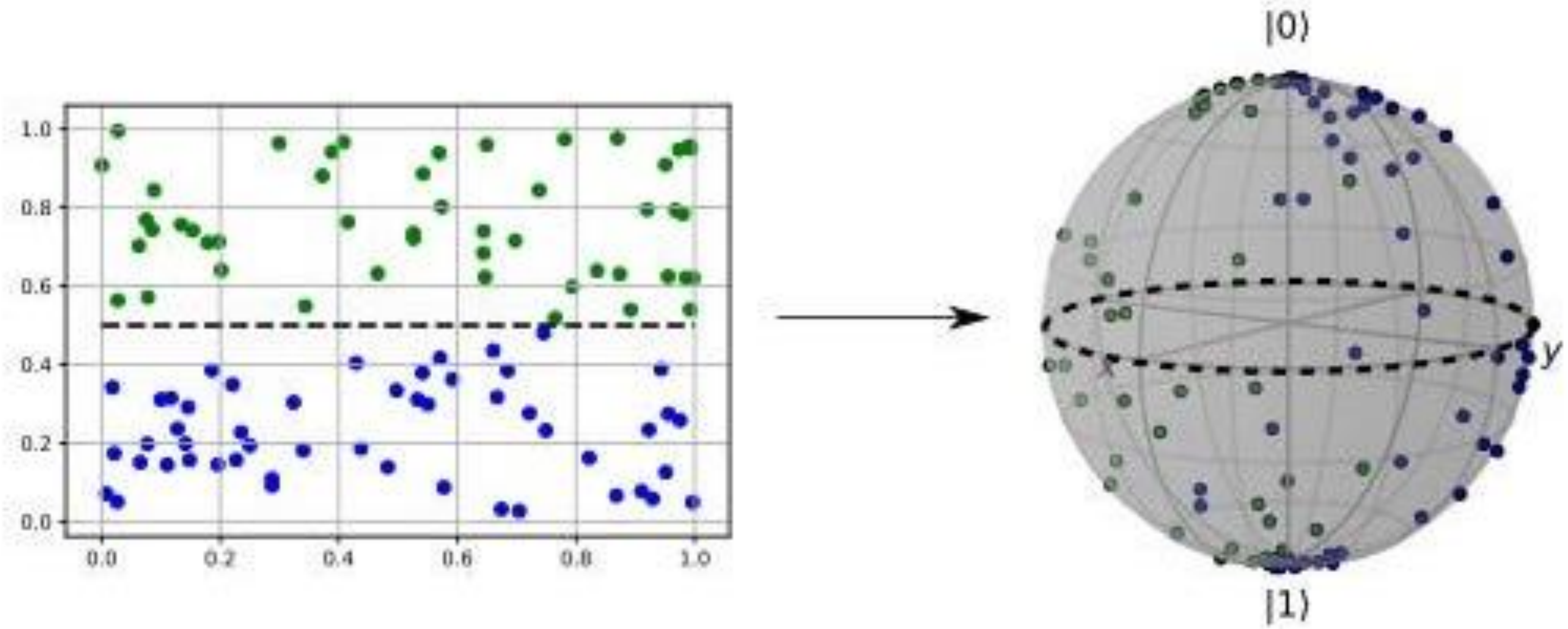
- High noise levels make extensive error correction protocols necessary

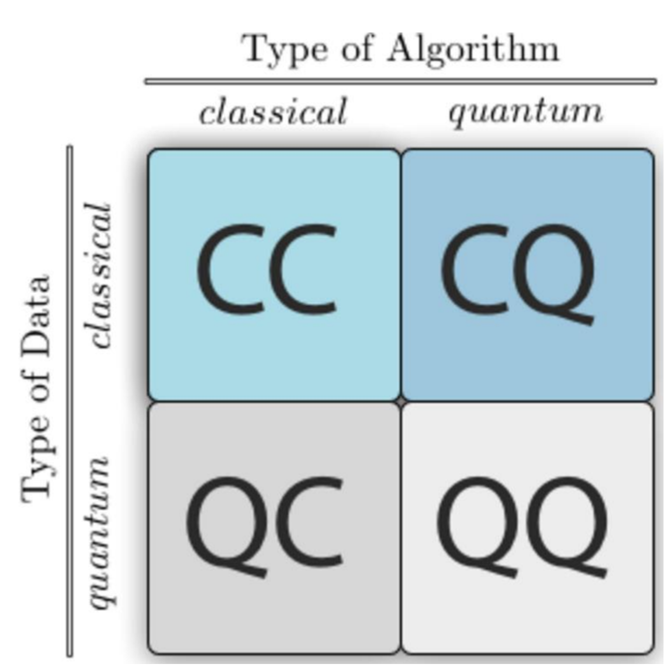


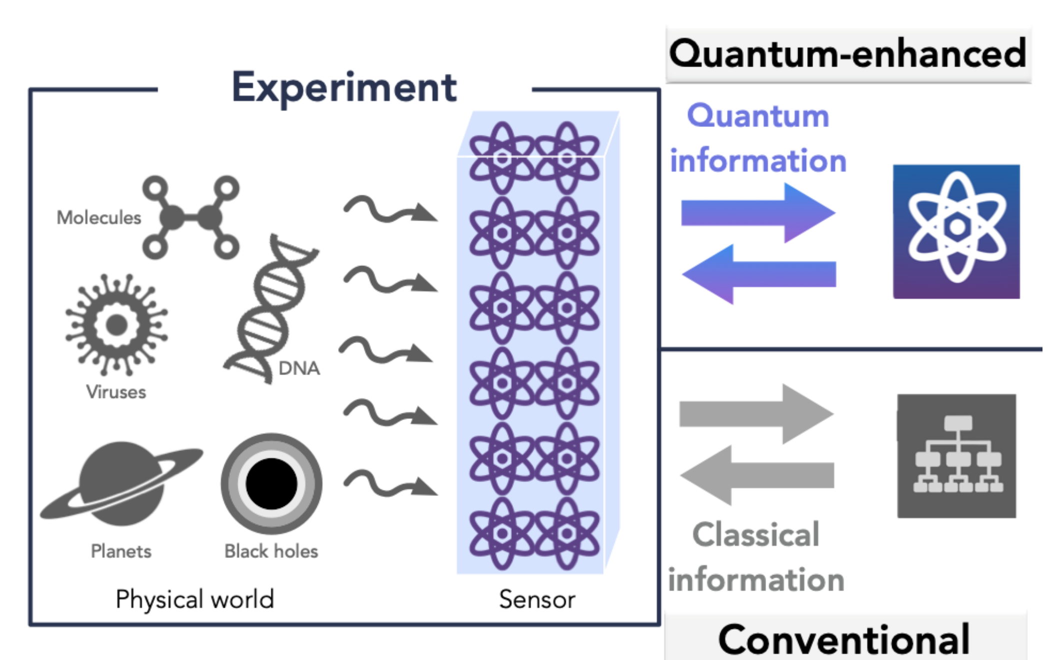
shafi-syed.medium.com/quantum-data-and-its-embeddings-1-3b022b2f1245
en.wikipedia.org/wiki/Quantum_machine_learning
science.org/doi/10.1126/science.abn7293

Main problems

- Data loading, as we basically perform in-memory computing

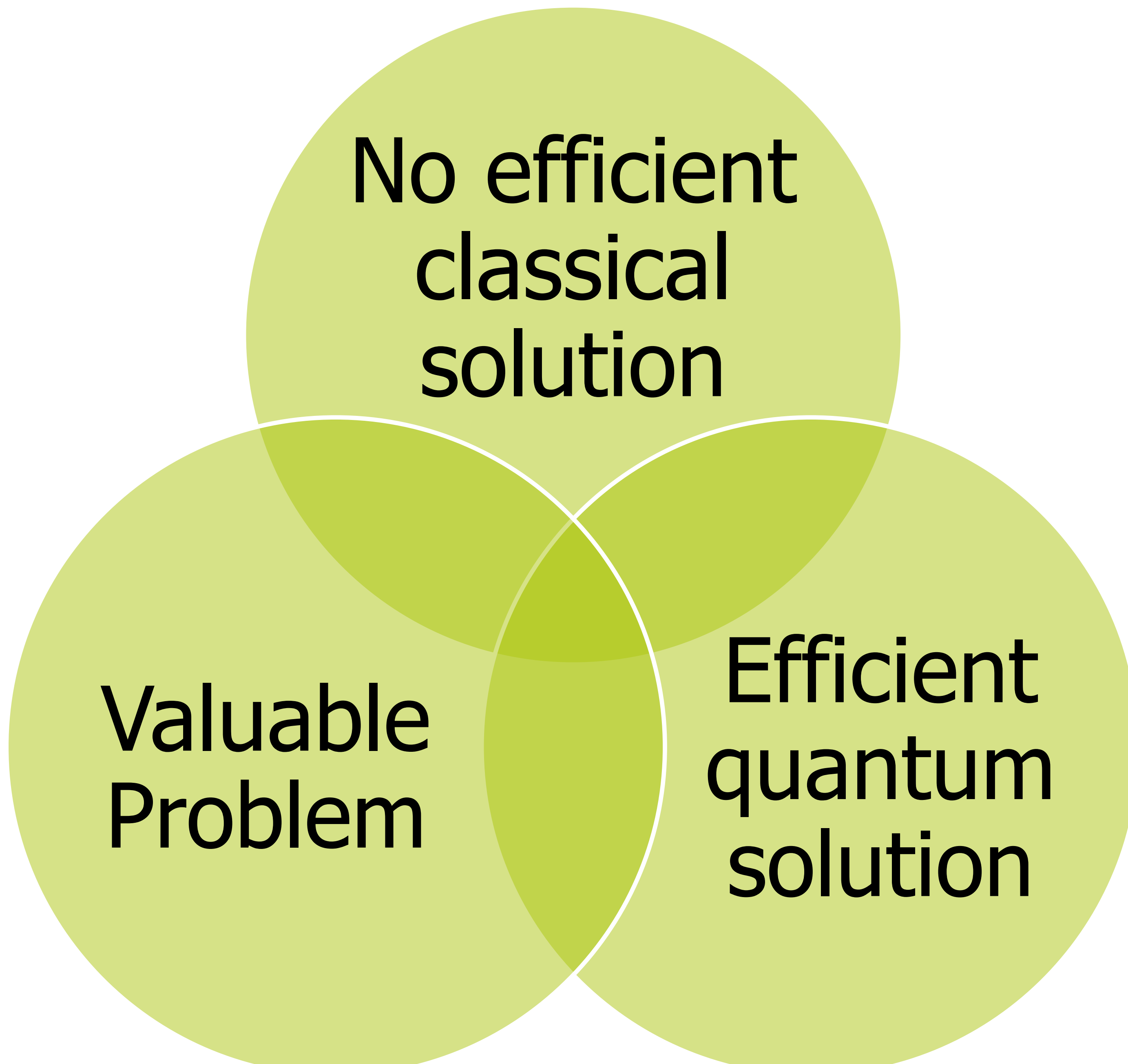






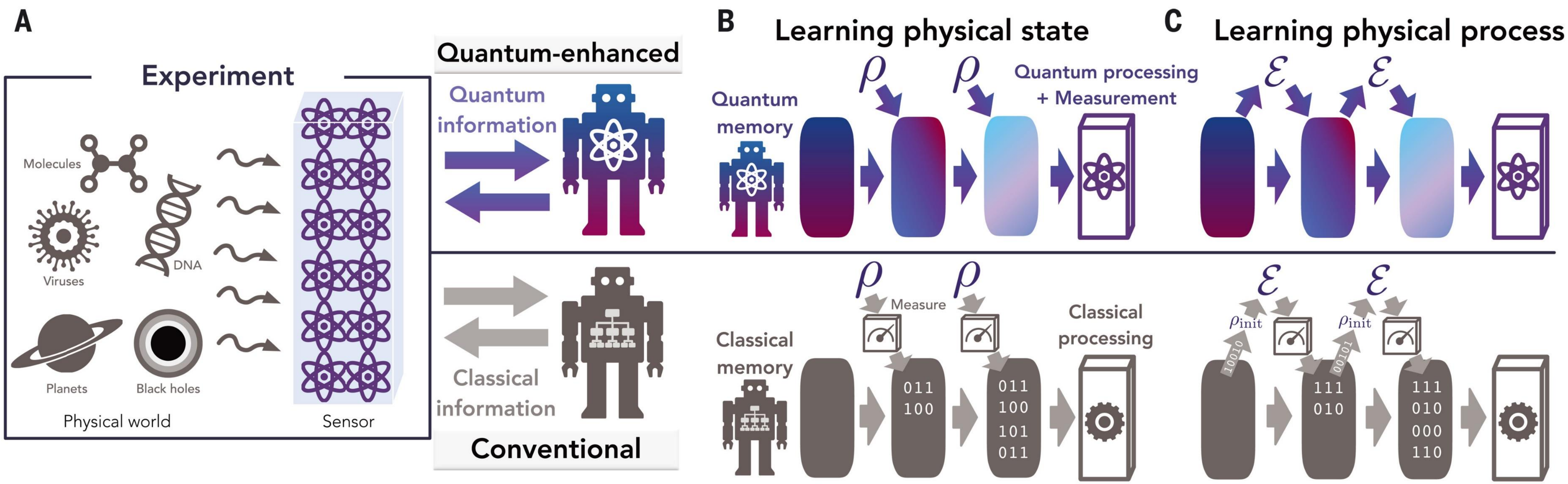
Main problems

- Find the right solution AND the right problem



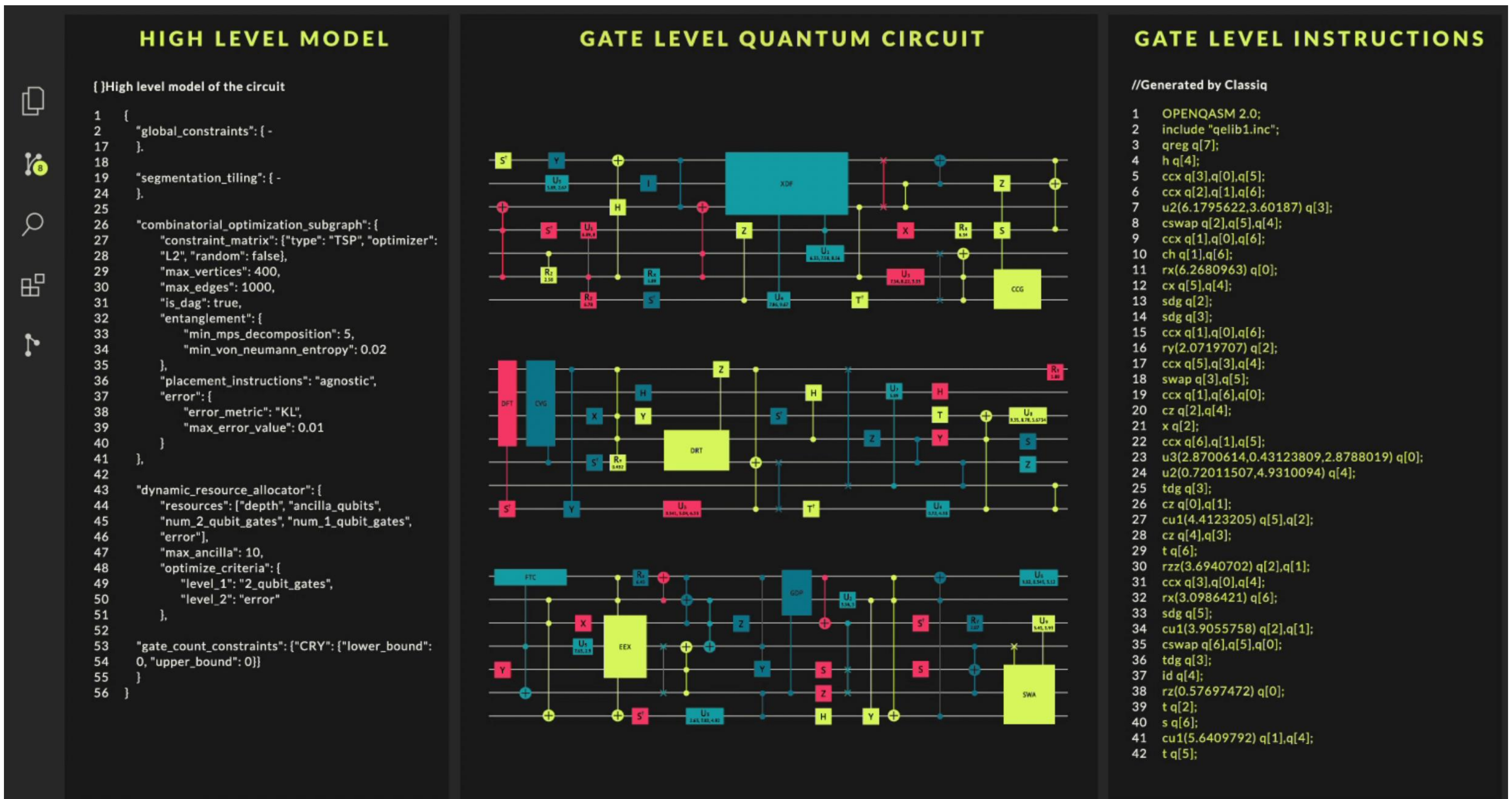
Some highlights

- Proven superiority, see for example „Quantum advantage in learning from experiments“



Some highlights

- Many successful efforts regarding abstraction, SDKs, APIs, …



Fundamentals of quantum computing

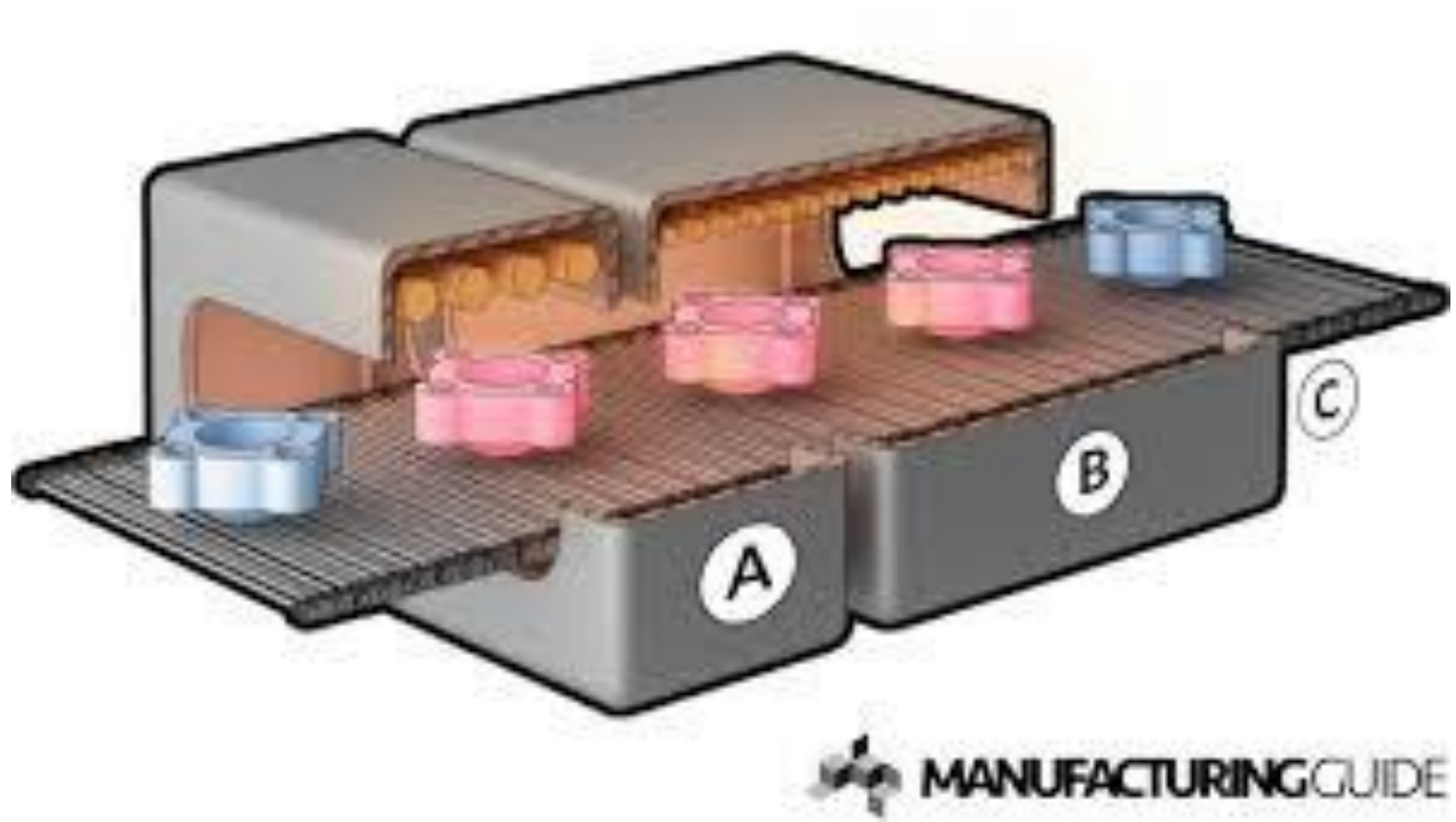
- Overview
- Quantum Gate Model
- **Quantum Annealing**
- Conclusion

A niche product in a niche market

- Basically one vendor of quantum annealing
- But also: quantum-inspired hardware

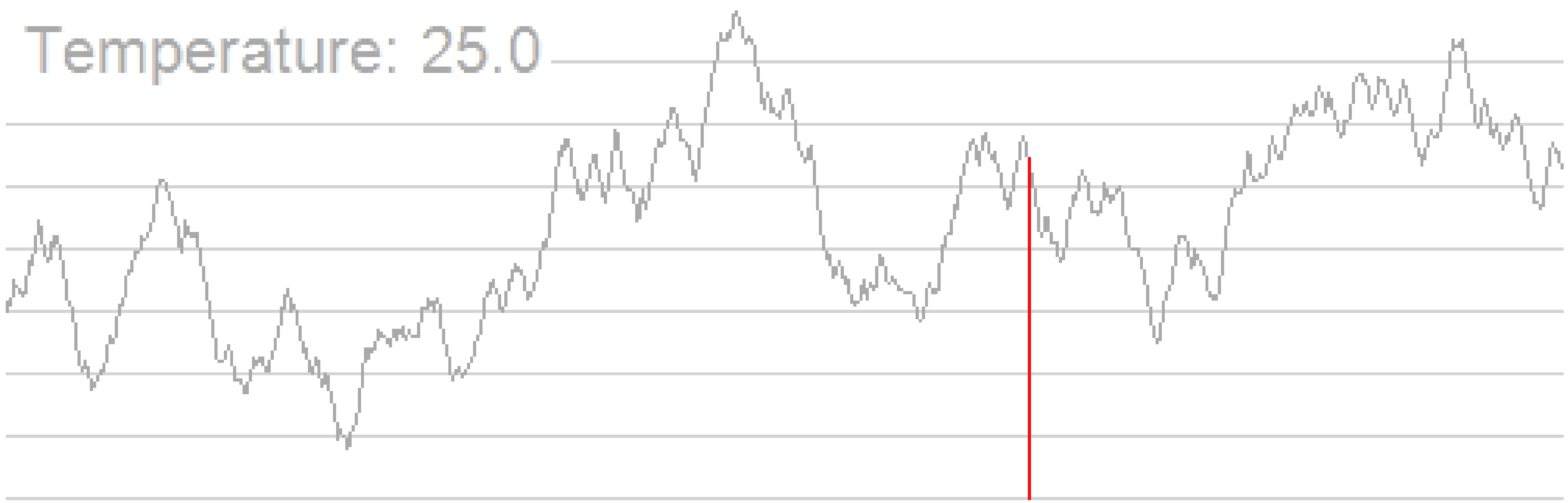
Annealing in metallurgy and materials science

- Heat treatment to alter physical properties of material



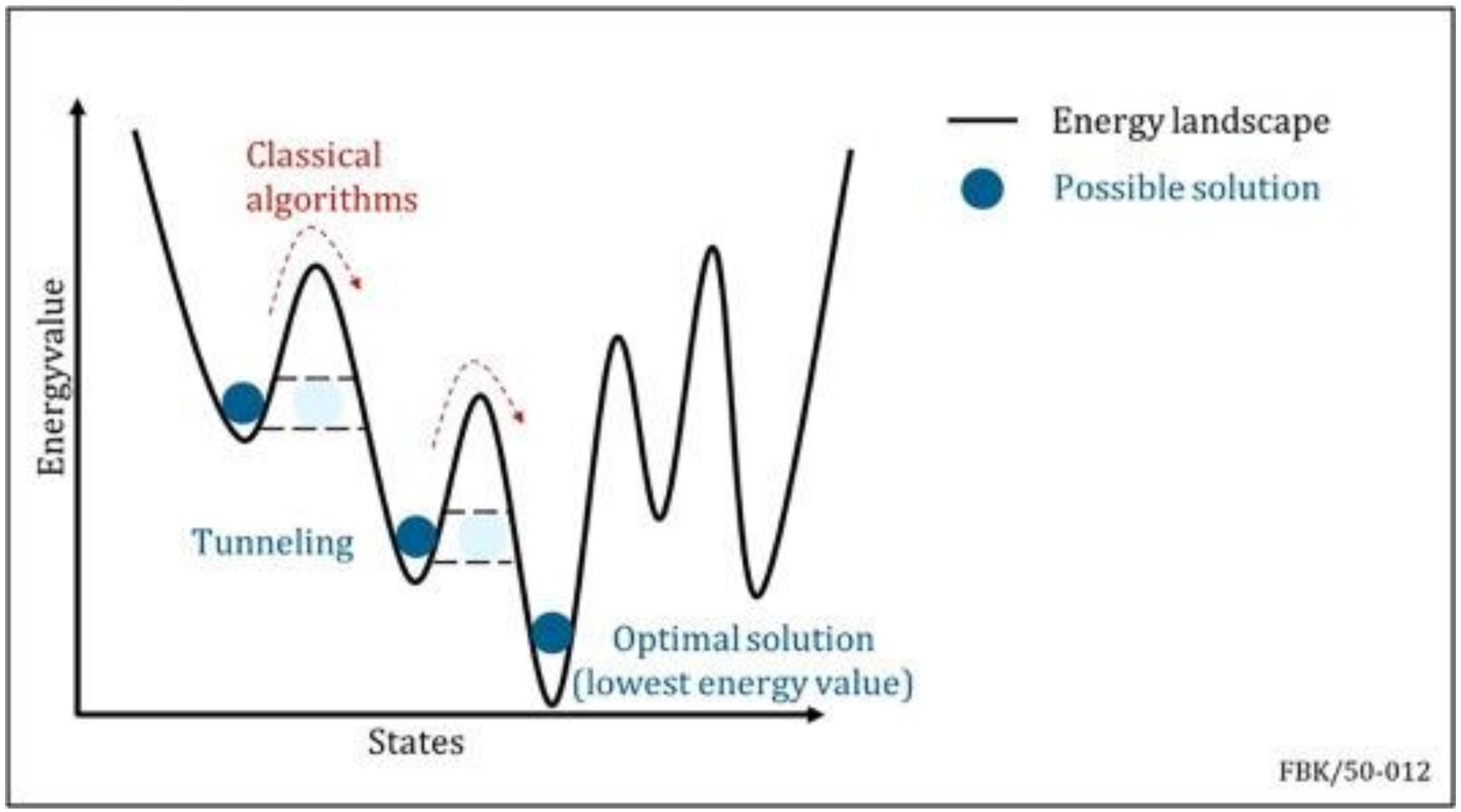
Simulated annealing

- Metaheuristic to find lowest energy state that corresponds to global optimum
- Depending on temperature, worse solution is temporarily accepted



Quantum annealing

- Effect of quantum tunneling is utilized



How to model this solution landscape?

- Hamiltonian function of Ising model from theoretical physics
- Quadratic unconstrained binary optimization (QUBO) from discrete math

- Given a real-valued upper triangular matrix $Q \in \mathbb{R}^{n \times n}$
- Given a binary vector $x \in \mathbb{B}^n$
- This defines function $f_Q: \mathbb{B}^n \to \mathbb{R}$ through

$$f_Q(x) = x^T Q x = \sum_{i=1}^{n} \sum_{j=i}^{n} Q_{ij} x_i x_j$$

- Goal: find binary vector x^* that is minimal w.r.t. f_Q, meaning

$$\forall x \in \mathbb{B}^n: f_Q(x^*) \leq f_Q(x)$$

And now you!

- What binary vector minimizes this QUBO matrix?

1	2	3	4	5	6	7	8
	2	3	4	5	6	7	8
		3	4	5	6	7	8
			4	5	6	7	8
				5	6	7	8
					6	7	8
						7	8
							8

$$f_Q(x) = x^T Q x = \sum_{i=1}^{n} \sum_{j=i}^{n} Q_{ij} x_i x_j$$

And now you!

- What binary vector minimizes this QUBO matrix?
- It‘s $f(00000000) = 0$
- Any other possibility would add some positive number
- For example: $f(00000001) = 8$

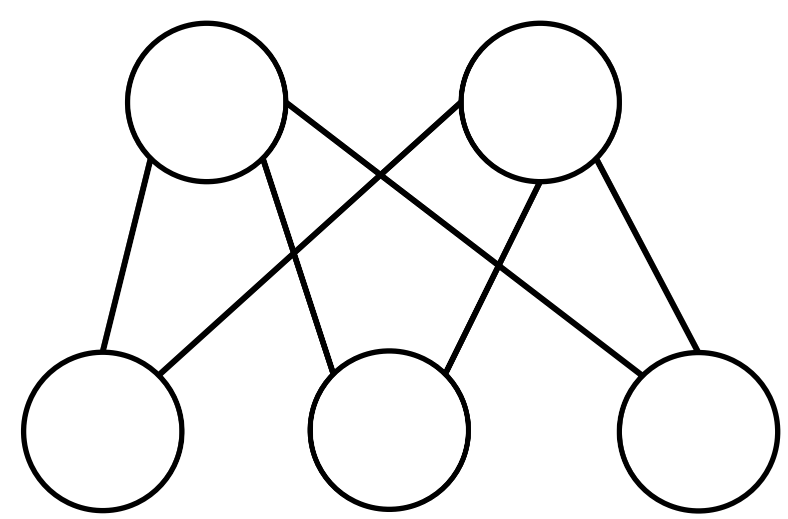
1	2	3	4	5	6	7	8
	2	3	4	5	6	7	8
		3	4	5	6	7	8
			4	5	6	7	8
				5	6	7	8
					6	7	8
						7	8
							8

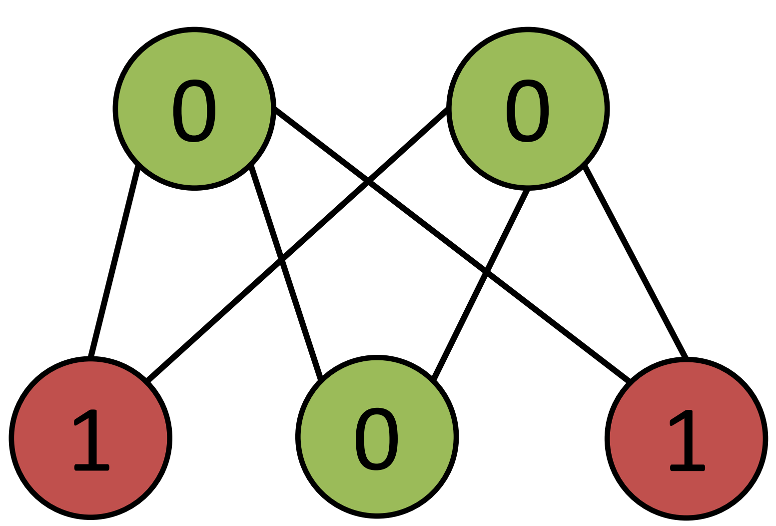
$$f_Q(x) = x^T Q x = \sum_{i=1}^{n} \sum_{j=i}^{n} Q_{ij} x_i x_j$$

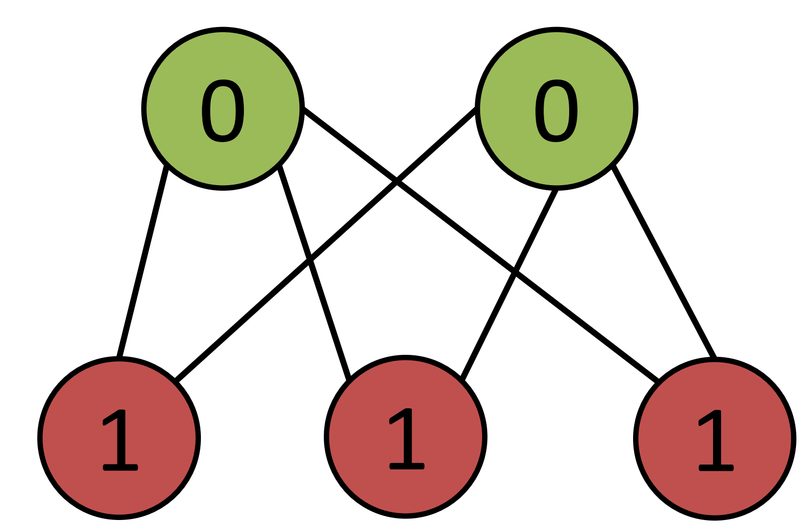
Maximum cut problem (MAXCUT)

- Goal is to split set of graph’s vertices into two disjoint parts …
- … such that number of edges spanning both parts is maximized

$$\min \sum_{(i,j)\in E} \left(-x_i - x_j + 2x_i x_j\right)$$







frontiersin.org/articles/10.3389/fict.2019.00013/full

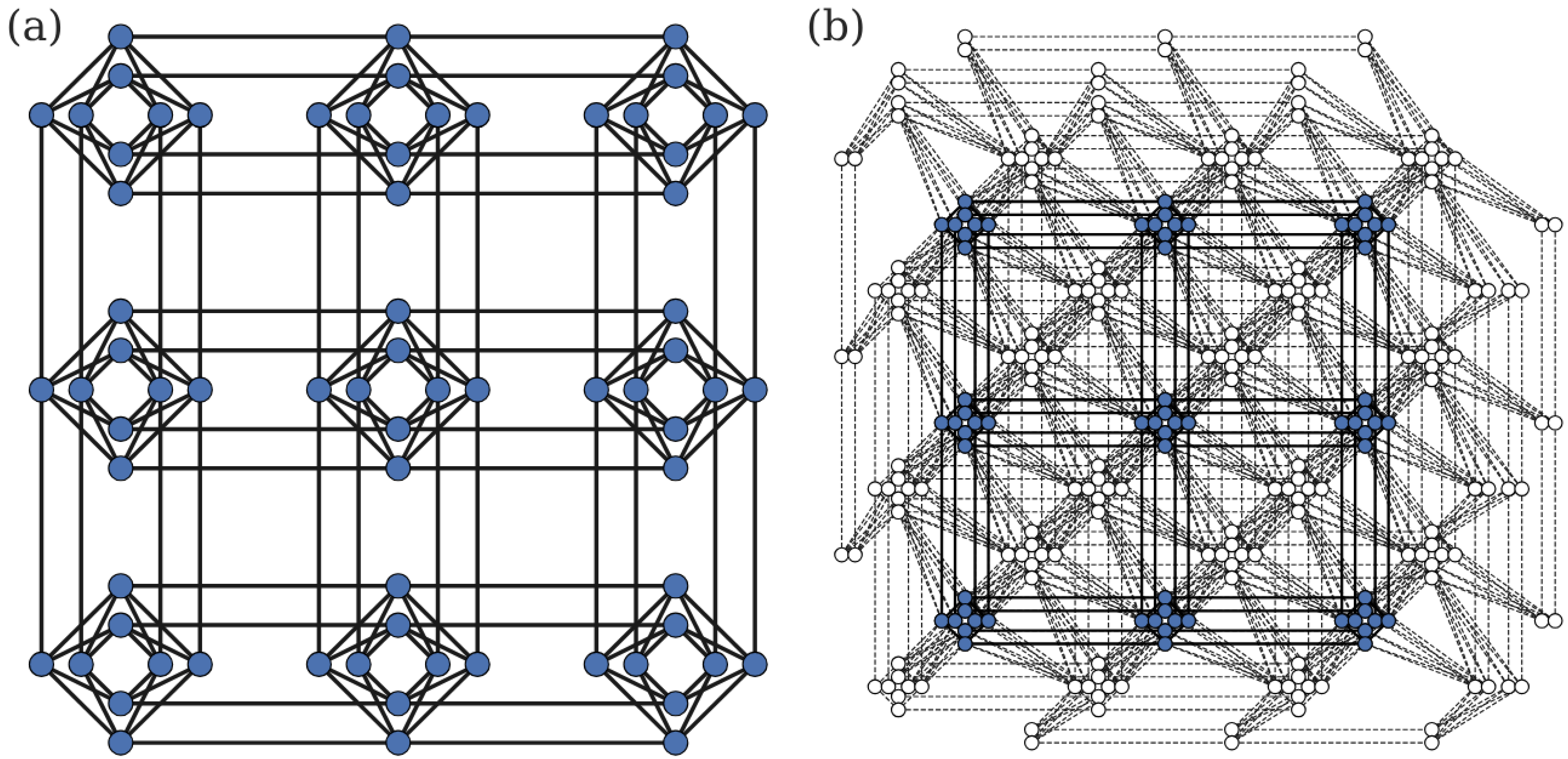
Main problems

- Often inefficient problem transformations

	A1	A2	A3	A4	B1	B2	B3	B4	C1	C2	C3	C4
A1	-1 -1	2	2	2	2	ab		ab	2	ac		ac
A2		-1 -1	2	2	ab	2	ab		ac	2	ac	
A3			-1 -1	2		ab	2	ab		ac	2	ac
A4				-1 -1	ab		ab	2	ac		ac	2
B1					-1 -1	2	2	2	2	bc		bc
B2						-1 -1	2	2	bc	2	bc	
B3							-1 -1	2		bc	2	bc
B4								-1 -1	bc		bc	2
C1									-1 -1	2	2	2
C2										-1 -1	2	2
C3											-1 -1	2
C4												-1 -1

Main problems

- Minor embedding



Some highlights

- As with gate model quantum computing, strongest results in the field of(quantum) physics

Article | Published: 19 April 2023

Quantum critical dynamics in a 5,000-qubit programmable spin glass

Andrew D. King, Jack Raymond, Trevor Lanting, Richard Harris, Alex Zucca, Fabio Altomare, Andrew J. Berkley, Kelly Boothby, Sara Ejtemaee, Colin Enderud, Emile Hoskinson, Shuiyuan Huang, Eric Ladizinsky, Allison J. R. MacDonald, Gaelen Marsden, Reza Molavi, Travis Oh, Gabriel Poulin-Lamarre, Mauricio Reis, Chris Rich, Yuki Sato, Nicholas Tsai, Mark Volkmann, Jed D. Whittaker, ... Mohammad H. Amin

+ Show authors

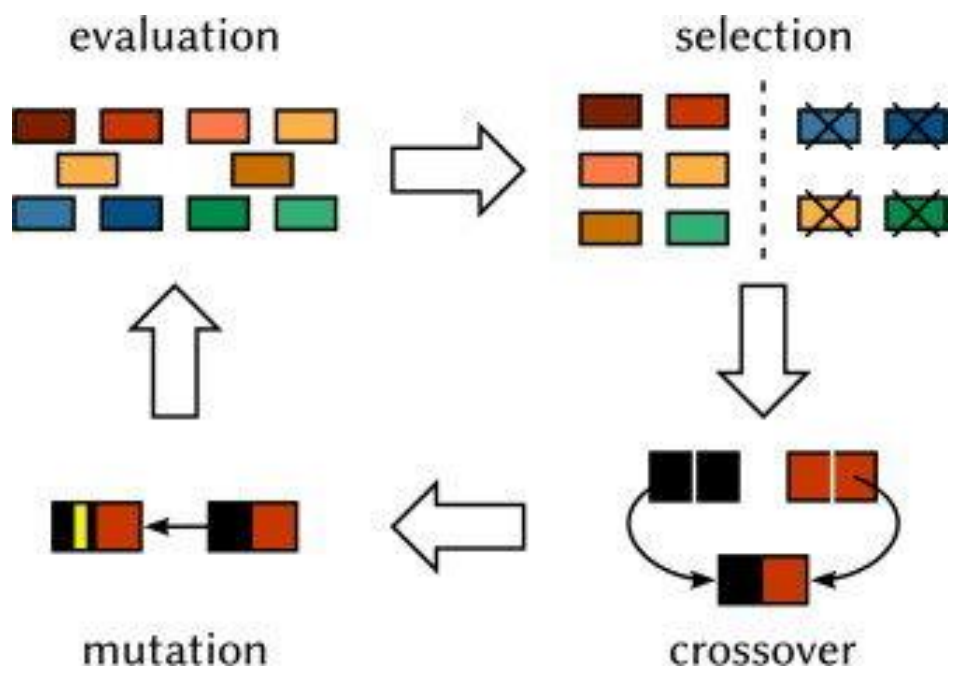
Nature **617**, 61–66 (2023) | Cite this article

Abstract

Experiments on disordered alloys[1,2,3] suggest that spin glasses can be brought into low-energy states faster by annealing quantum fluctuations than by conventional thermal annealing. Owing to the importance of spin glasses as a paradigmatic computational testbed, reproducing this phenomenon in a programmable system has remained a central challenge in quantum optimization[4,5,6,7,8,9,10,11,12,13]. Here we achieve this goal by realizing quantum-critical spin-glass dynamics on thousands of qubits with a superconducting quantum annealer. We first demonstrate quantitative agreement between quantum annealing and time evolution of the Schrödinger equation in small spin glasses. We then measure dynamics in three-dimensional spin glasses on thousands of qubits, for which classical simulation of many-body quantum dynamics is intractable. We extract critical exponents that clearly distinguish quantum annealing from the slower stochastic dynamics of analogous Monte Carlo algorithms, providing both theoretical and experimental support for large-scale quantum simulation and a scaling advantage in energy optimization.

Some highlights

- Research on QUBO benefits not only QA community, but also others
  - Gate model (via QAOA)
  - Meta-Heuristics (Genetic/Evolutionary algorithm, Dynamic programming, …)
  - Operation Research
  - Problem transformation, algorithmics, approximation
  - …

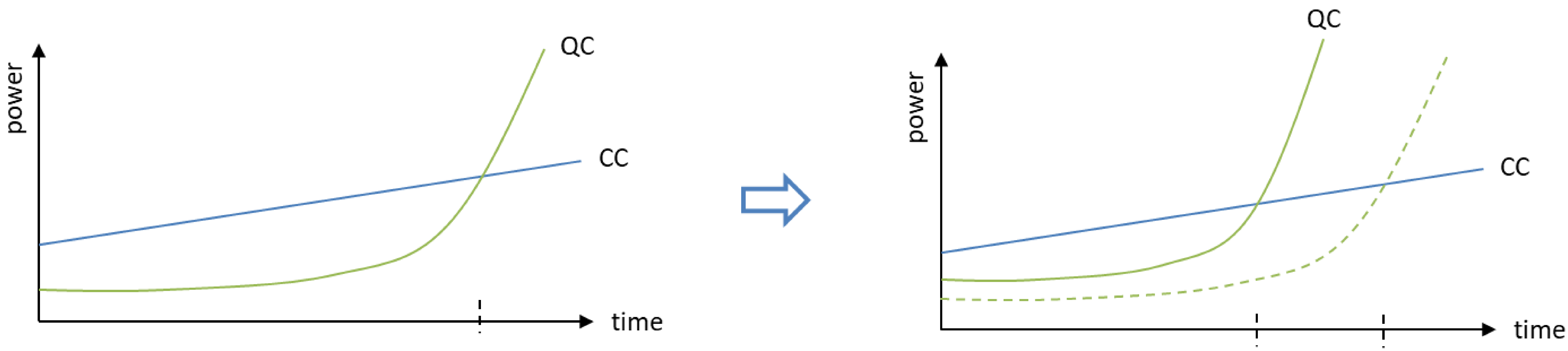


Fundamentals of quantum computing

- Overview
- Quantum Gate Model
- Quantum Annealing
- **Conclusion**

Quantum gate model

- Become quantum ready!



Quantum annealing

- Find the right problem (transformations)

? →

$$\sum_{i=1}^{N} c_i X_i + \sum_{i=1}^{N} \sum_{j=1}^{i} Q_{ij} X_i X_j$$

$$X_i \in \{0,1\}$$

$$c_i, Q_{ij} \in \mathbb{R}$$

Let’s get started!

- The entrance barrier is lower than one thinks…

Qiskit

Cirq

Ocean

… and many more!

Data Management using Quantum Computers
-- the Art of Encoding

LO and Assignment

- Learning Objective
  - Explain how to solve a data management problem on quantum computers

- Assignment
  - Design an approach that solves a data management or data science problem using quantum computers

DB Problems Solved Using QPUs

Reference	DB problem	Subproblem	Formulation	Intermediate quantum algorithm	Quantum computer
I. Trummer et al., VLDB'16	Query optimization	Multiple query optimization	QUBO	–	Annealing-based
T. Fankhauser et al., IEEE Access, 2023				QAOA	Gate-based
M. Schonberger et al., SIGMOD23		Join ordering		QAOA	Gate-based & annealing-based
N. Nayak et al., BiDEDE '23				QAOA, VQE	Gate-based & annealing-based
T. Winker et al., BiDEDE '23			–	VQC	Gate-based
K. Fritsch et al., VLDB'23 Demo	Data integration	Schema matching	QUBO	QAOA	Gate-based & annealing-based
T. Bittner et al., IDEAS'20, OJCC S. Groppe et al., IDEAS'21	Transaction management	Two-phase locking	QUBO	–	Annealing-based

Multiple Query Optimization

T. K. Sellis, Multiple-query optimization, TODS 1988

- Problem: multiple query optimization (MQO) studies how to choose query plans given a set of queries
  - Goal: minimize the total execution cost
  - Key: shared computation between different queries
  - Valid solution
    - A subset of plans selected for query execution
    - 1:1 mapping between the query and the query plan
  - Optimal solution
    - A valid solution with minimal execution cost among all valid solutions

Example

- Two queries

Query	Query plan	Cost
q_1	p_1	2
	p_2	4
q_2	p_3	3
	p_4	1

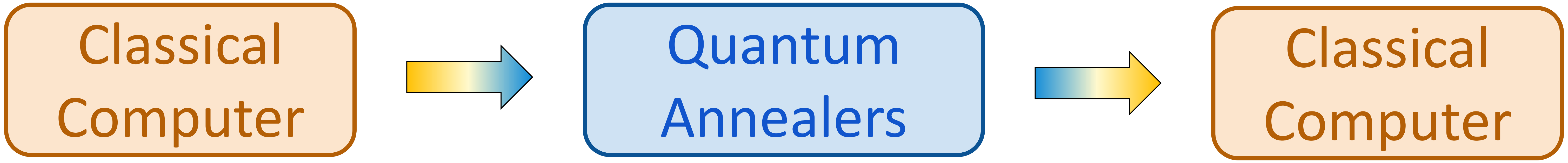
$*$ Shared computation saving between p_2 and p_3 is 5

Which query plans to choose for q_1 and q_2?

Solve MQO Using Quantum Annealing

- Goal: find near-optimal MQO solution
  - Reduce execution cost by sharing computation among queries

- Solution



1. Logical mapping
2. Physical mapping

I. Trummer and C. Koch, Multiple query optimization on the D-Wave 2X adiabatic quantum computer," VLDB'16.

Encoding: Binary Variables

Logical level

$$x_p = \begin{cases} 1, \text{if query plan } p \text{ is selected} \\ 0, \text{otherwise} \end{cases}$$

Physical level

$$b_i = \begin{cases} 1, \text{if the } i - th \text{ qubit has the state of } 1 \\ 0, \text{if the } i - th \text{ qubit has the state of } 0 \end{cases}$$

Example

- Adding binary variable x_p

Query	Query plan	x_p	Cost
q_1	p_1	x_1	2
	p_2	x_2	4
q_2	p_3	x_3	3
	p_4	x_4	1

$*$ Shared computation saving between p_2 and p_3 is 5

Logical Mapping

- Design logical mapping:
  - Transform the MQO problem into a QUBO problem
  - **Minimize** the logical energy formula

Obtain valid solutions

Each query has **at least** one plan

Each query has **at most** one plan

Lower cost

Cost of each plan

Cost saving by intermediate results

Logical Mapping

Query	Query plan	$x_p \in \{\mathbf{0}, \mathbf{1}\}$	Cost
q_1	p_1	x_1	2
	p_2	x_2	4
q_2	p_3	x_3	3
	p_4	x_4	1

∗ Shared computation saving between p_2 and p_3 is 5

1. *Each query has at least one plan*

$$E_L = -\sum_{p \in P} x_p$$

2. *Each query has **at most** one plan*

$x_1 \rightarrow 1$
$x_2 \rightarrow 1$
$x_3 \rightarrow 1$
$x_4 \rightarrow 1$

NOT what we want!

Logical Mapping

Query	Query plan	$x_p \in \{0,1\}$	Cost
q_1	p_1	x_1	2
	p_2	x_2	4
q_2	p_3	x_3	3
	p_4	x_4	1

$*$ Shared computation saving between p_2 and p_3 is 5

1. *Each query has at least one plan*

$$E_L = -\sum_{p \in P} x_p$$

2. *Each query has **at most** one plan*

$$E_M = \sum_{q \in Q} \sum_{\{p1, p2 \subseteq Pq\}} x_{p1} x_{p2}$$

for q_1: $x_1 \; x_2 \rightarrow 0$
for q_2: $x_3 \; x_4 \rightarrow 0$

Logical Mapping

Query	Query plan	$x_p \in \{0, 1\}$	Cost
q_1	p_1	x_1	2
	p_2	x_2	4
q_2	p_3	x_3	3
	p_4	x_4	1

$*$ Shared computation saving between p_2 and p_3 is 5

3. *Sum up the cost of each plan*

$$E_C = \sum_{p \in P} c_p x_p$$

If we choose p_2 for q_1, p_3 for q_2:

$x_1 \rightarrow 0$
$x_2 \rightarrow 1$
$x_3 \rightarrow 1$
$x_4 \rightarrow 0$

$$E_C = 0 * 2 + 1 * 4 + 1*3 + 0*1 = 7$$

Logical Mapping

Query	Query plan	$x_p \in \{0,1\}$	Cost
q_1	p_1	x_1	2
	p_2	x_2	4
q_2	p_3	x_3	3
	p_4	x_4	1

* **Shared computation saving between p_2 and p_3 is 5**

4. *Deduct the cost saved by intermediate results*

$$E_S = -\sum_{\{p1,p2 \subseteq P\}} s_{p1,p2} x_{p1} x_{p2}$$

If we choose p_2 for q_1, p_3 for q_2:

$x_1 \rightarrow 0$
$x_2 \rightarrow 1$
$x_3 \rightarrow 1$
$x_4 \rightarrow 0$

$E_S = -5*1*1 = -5$

Logical Mapping

I. Trummer and C. Koch, Multiple query optimization on the D-Wave 2X adiabatic quantum computer," VLDB'16.

- Final logical energy formula

Obtain valid solutions

Lower cost

Each query has **at least** one plan

Each query has **at most** one plan

Cost of each plan

Cost saving by intermediate results

$$w_L E_L + w_M E_M + E_C + E_S$$

Physical Plan

- Physical mapping
  - Logical energy formula → **physical energy formula** of qubit states
  - Key challenge: mapping variables to qubit

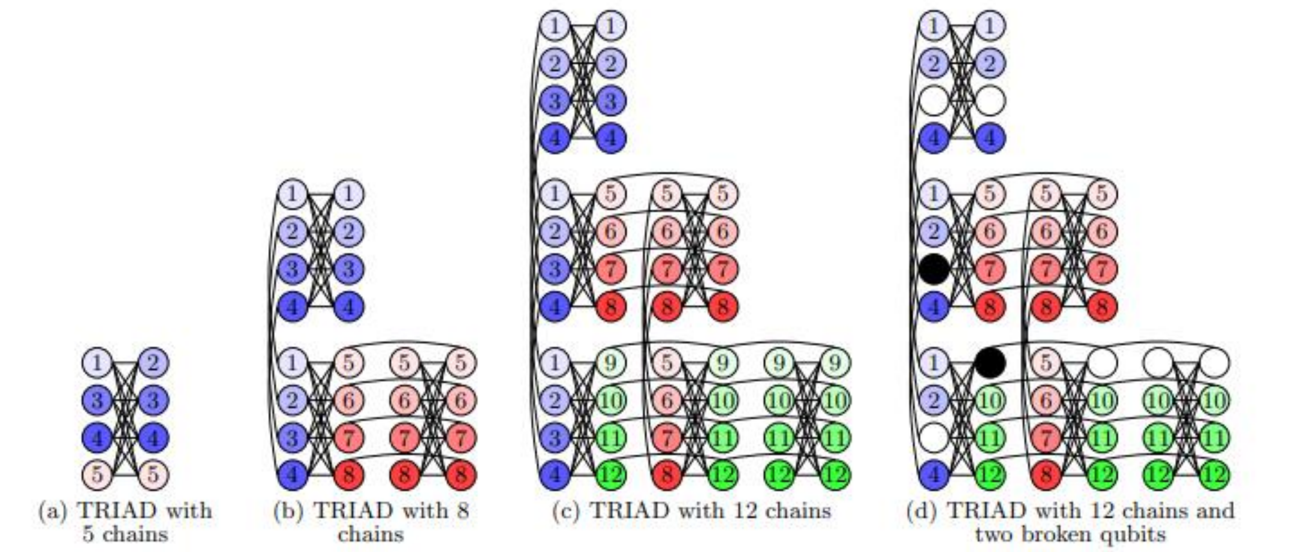


Figure 2: TRIAD pattern in different sizes: we show qubits as circles, annotated by the ID of the logical variable that they represent. The mapping from variables to qubits assures that each variable shares at least one connection (in black) with each of the other variables.

- Hardware-specific constraint:
  - Sparse qubit connectivity
  - All qubits representing the same variable form a chain
  - Logical variables in a quadratic term need to be represented by connected groups of qubits
  - Broken qubits

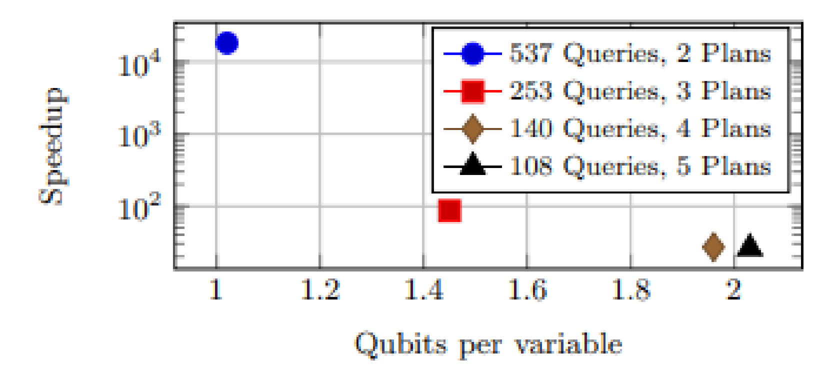


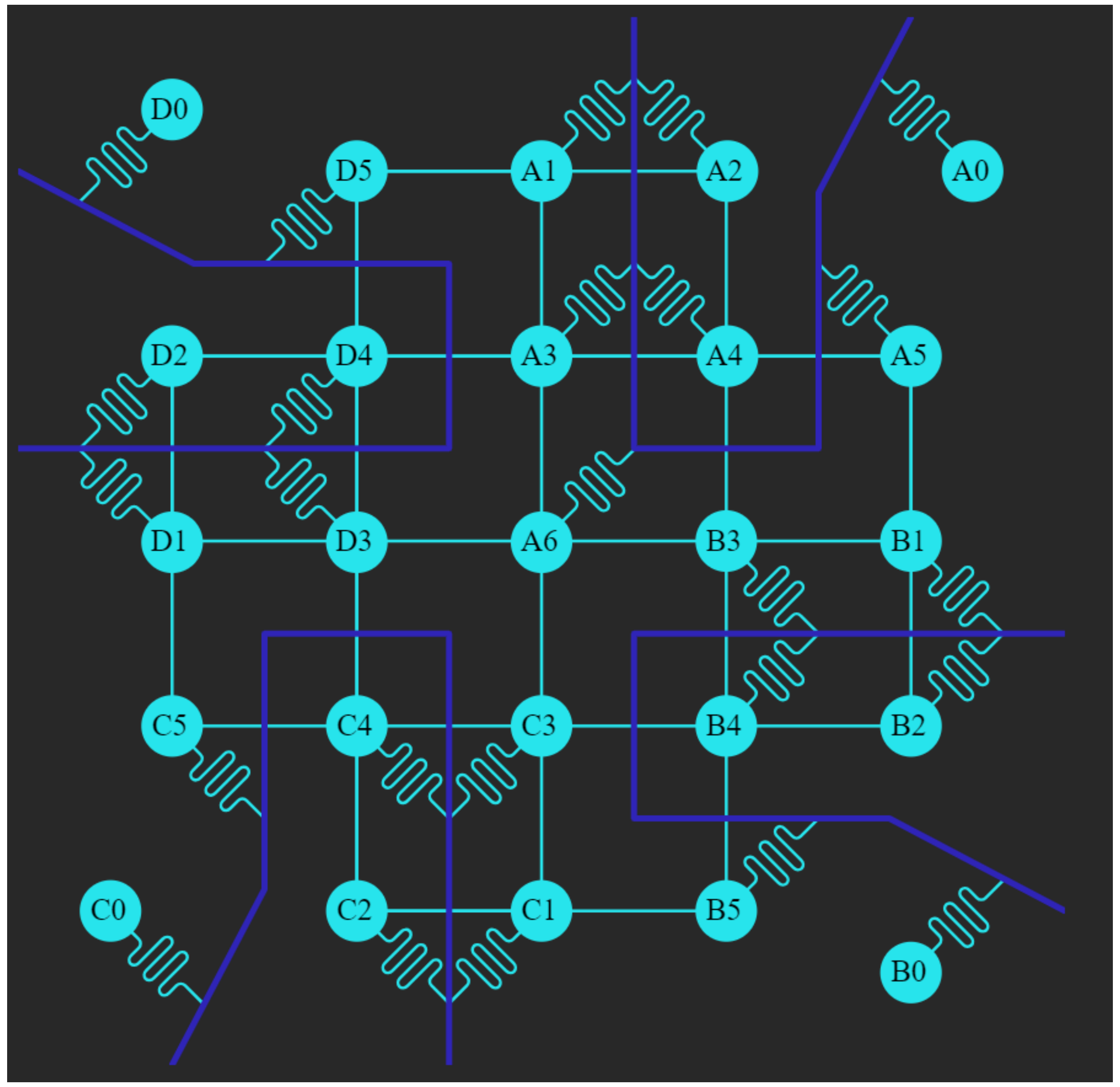
Figure 6: Average quantum speedup for different classes of test cases: having to use more qubits per problem variable decreases the speedup.

Results: 1000x speedup

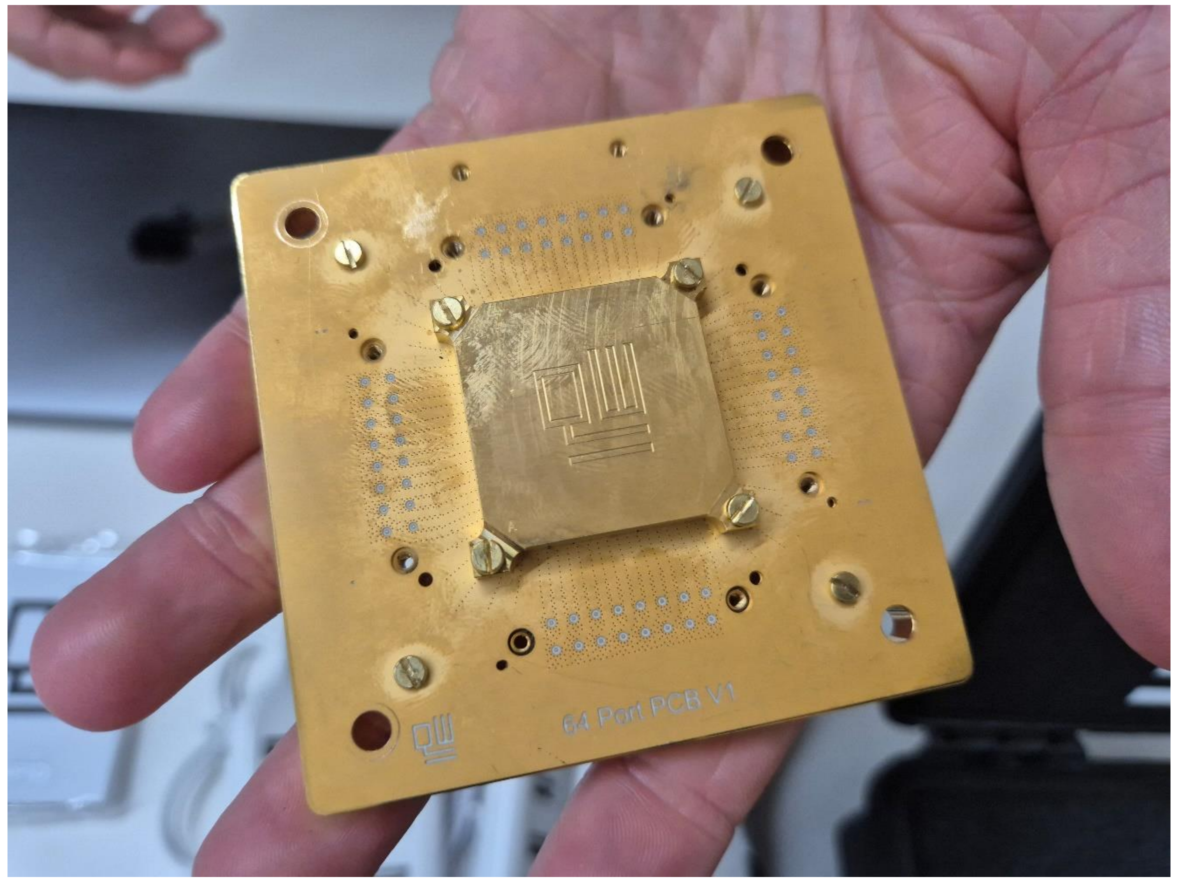
When fewer qubits per variable

Hardware

25 Qubits: 21 fully connected +
4 isolated qubits



QuantWare



Extend MQO to Gate-based QC

- Gate-based quantum computer, more **universal** than quantum annealing
- Hybrid classical-quantum algorithms like the Quantum Approximate Optimization Algorithm (QAOA)
- Similar logical mapping, more effective physical mapping
- Limited scalability due to the number of qubits of gate-based QCs

	Trummer and Koch	**This work**
Max. # of queries	537 queries, 2 plans	7 queries, 2 plans
Qubits used	1074 / 1097 (98%)	14 / 14 (100%)
Max. # of plans	108 queries, 5 plans	2 queries, 7 plans
Qubits used	540 / 1097 (49%)	14 / 14 (100%)

T. Fankhauser, M. E. Soler, R. M. Fuchslin, and K. Stockinger, "Multiple query optimization using a gate-based quantum computer," IEEE Access, 2023.

Join Ordering

- Problem: join ordering (JO) studies on how to identify the optimal ordering of join operations between relations for an efficient query plan

- Solutions
  - Exhaustive search, e.g., dynamic programming
  - Heuristic methods
  - Special-Purpose Solvers, e.g., MILP
  - ML-based methods
  - **SW-HW codesign: methods based on QPU or digital annealer**

Join Ordering on Gate-based QC and QA

- Problem Transformation: from JO to QUBO
- QPU metrics:
  - Overall QPU time
  - **Depth of QAOA circuit (**gate-based QPUs**)**
  - **Number of qubits**

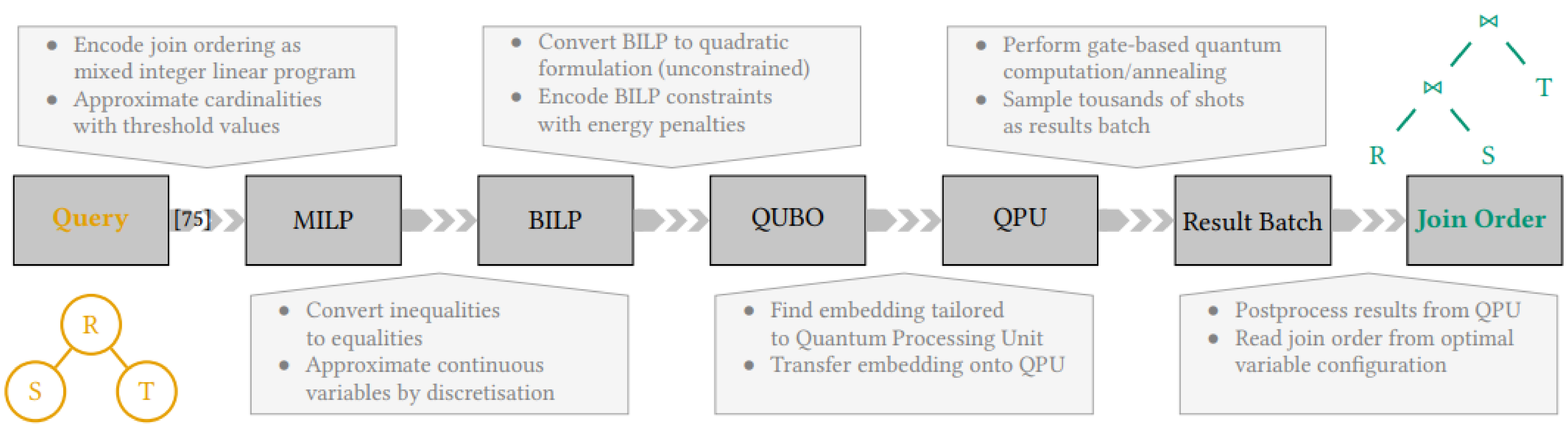


Fig. 1. Overview on all steps required to solve join ordering with quantum computing.

M. Schonberger, et al. "Ready to leap (by co-design)? join order optimisation on quantum hardware. SIGMOD'23.

Join Ordering using Quantum Annealing

M. Schonberger, et al. "Ready to leap (by co-design)? join order optimisation on quantum hardware. SIGMOD'23.

- Gate-based, QA-based
- Encoding: JO→MILP→QUBO
- Without guaranteed result

M. Schonberger, et al. "Quantum-Inspired Digital Annealing for Join Ordering." VLDB'24.

- Fujitsu 2nd Gen digital annealer
  - Fully connected bits
  - 8192 bits/variables
- Encoding: JO→QUBO
  - Include parts not supported by MILP, more native to QA/QUBO
  - Efficient with less encoding size
- With guaranteed result
- Scalability: 50 relations

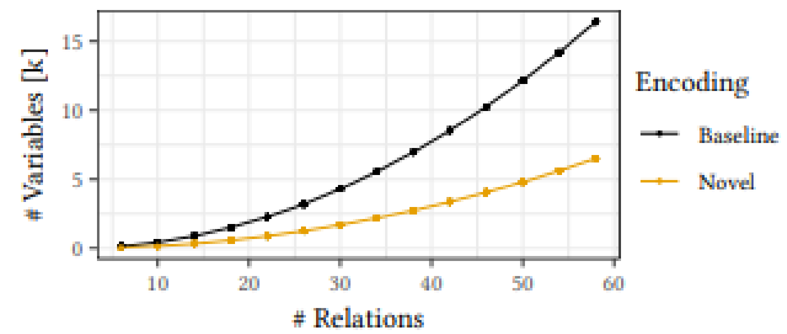


Figure 2: Mandatory variable scaling for cycle queries, comparing the baseline JO-QUBO [41] and our novel encoding.

Encoding

$$H = A(HV + H_p) + HC$$

Terms:

- A: penalty for violating constraints
- HV: validity term to enforce the solution is valid
- H_p: logarithmic cost of a join
- HC: constraint for intermediate result size

M. Schonberger, et al. "Quantum-Inspired Digital Annealing for Join Ordering." VLDB’24.

More References on Join Ordering Using Quantum Computers

1. **M. Schonberger, I. Trummer, and W. Mauerer. "Quantum-Inspired Digital Annealing for Join Ordering." Proceedings of the VLDB Endowment 17.3 (2023): 511-524.**
2. **M. Schonberger, S. Scherzinger, and W. Mauerer, "Ready to leap (by ¨ co-design)? join order optimisation on quantum hardware," Proceedings of the ACM on Management of Data, vol. 1, no. 1, pp. 1–27, 2023.**
3. M. Schonberger, I. Trummer, and W. Mauerer, "Quantum Optimisation ¨ of General Join Trees," in Joint Workshops at 49th International Conference on Very Large Data Bases (VLDBW'23)—International Workshop on Quantum Data Science and Management (QDSM'23), 2023.
4. N. Nayak, J. Rehfeld, T. Winker, B. Warnke, U. C¸ alikyilmaz, and S. Groppe, "Constructing Optimal Bushy Join Trees by Solving QUBO Problems on Quantum Hardware and Simulators," in Proceedings of the International Workshop on Big Data in Emergent Distributed Environments, ser. BiDEDE '23. 2023.
5. T. Winker, U. C¸ alikyilmaz, L. Gruenwald, and S. Groppe, "**Quantum Machine Learning** for Join Order Optimization Using Variational Quantum Circuits," in Proceedings of the International Workshop on Big Data in Emergent Distributed Environments, ser. BiDEDE '23, 2023.

Schema Matching

- Schema matching discovers the correspondences among the given source schemas

	mortality	pname	age	resting HR
0	0	Jack	20	60
1	0	Sam	35	58
2	0	Ruby	22	65
3	1	Jane	37	70

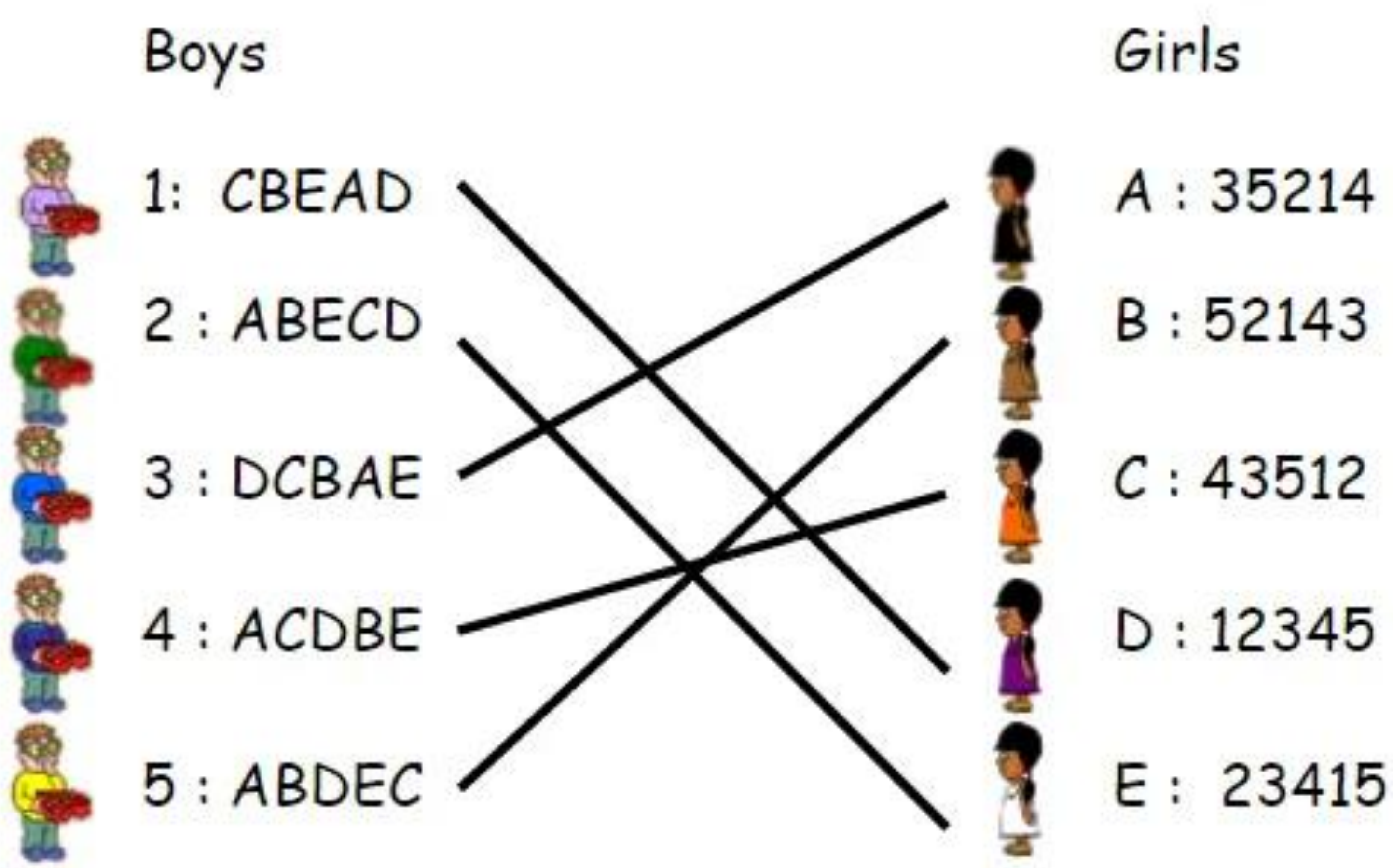
Hospital Amsterdam S_1

	mortality	name	age	oxygen	date diagnosed
0	1	Rose	45	95	1/4/21
1	0	Castiel	20	97	3/8/22
2	1	Jane	37	92	11/5/21

Hospital Delft S_2

Data Integration -- Schema matching

- Approach
  - Transform global matching to the *stable marriage problem*, then into QUBO



  - An **end-to-end** hybrid classical and quantum workflow

K. Fritsch and S. Scherzinger, “Solving hard variants of database schema matching on quantum computers,” VLDB’23 Demo.

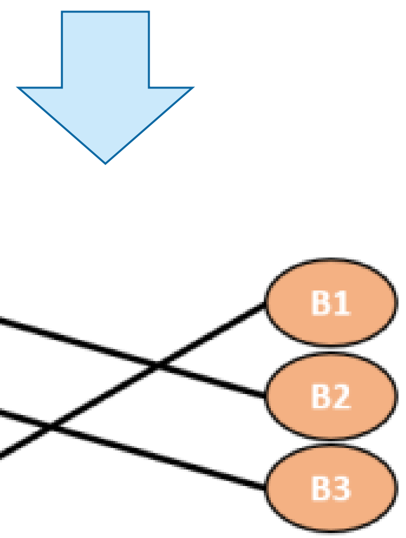
Data Integration -- Schema Matching

- Problem: global matching

  After obtaining similar attribute pairs with similarity scores, how to assign one-to-one correspondences, e.g., if there is a tie in ordering

Table 1: Instance of stable matching problem w/ ranked preferences. Brackets denote a *tie*, i.e., equally preferred partners.

Set A	Preferences of A*i*	Set B	Preferences of B*i*
A1	B2, B1	B1	A2, A1, A3
A2	B3, [B1, B2]	B2	A3, A2, A1
A3	B1, B3, B2	B3	A1, A3, A2



Data Integration -- Schema Matching

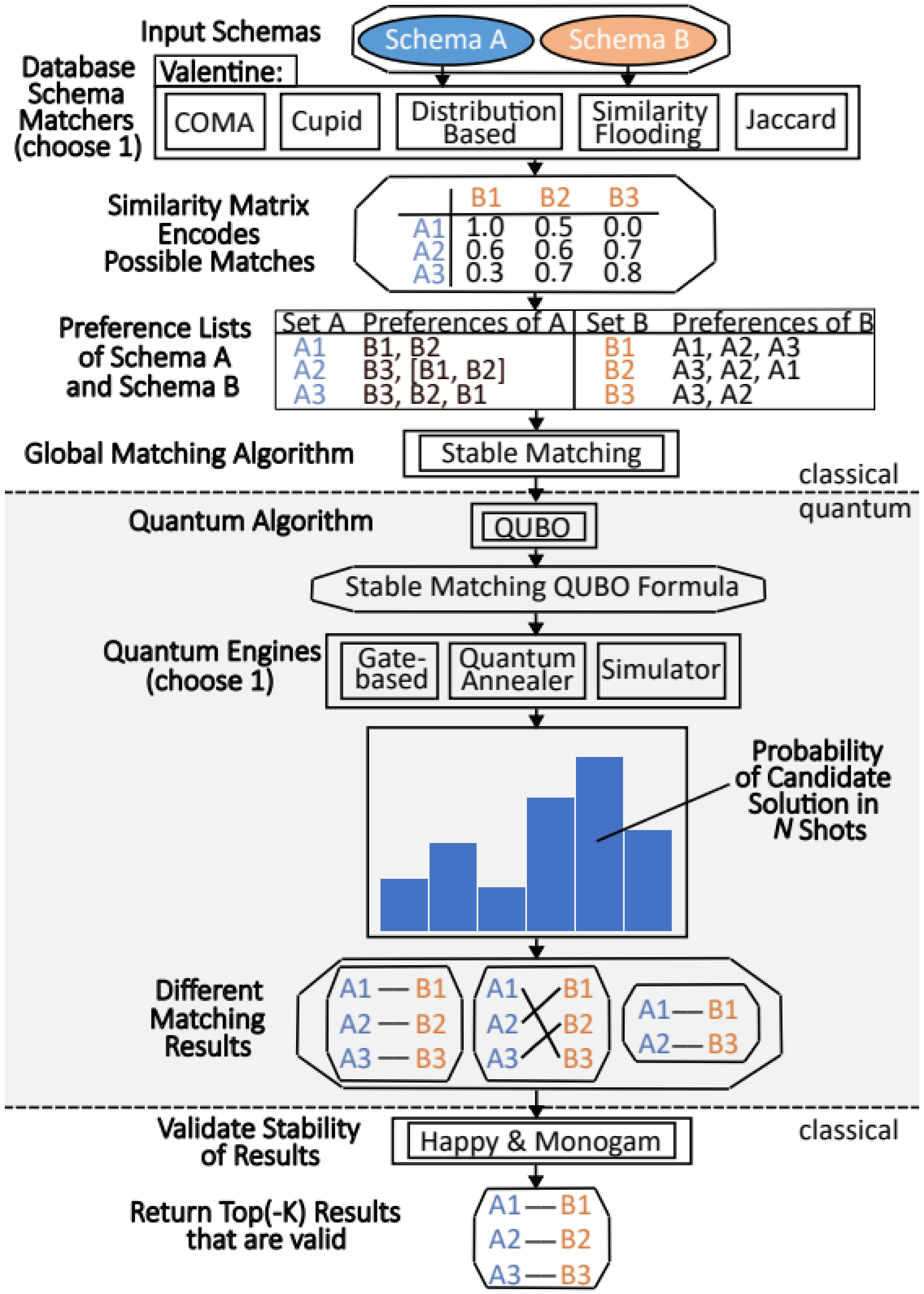
- Solution

$$P = -p_1 X_1 + p_2 X_2 + p_3\, X_3$$

Penalties:

- p_1 : minimize the objective function & to find as many pairs as possible
- p_2 : declare for all candidate attribute pairs
- p_3: prevent nonmonogamous matchings (if candidates are matched more than once)

Hybrid Classical and Quantum Workflow



Transaction Management

- Task: distribute *n* transactions over *k* machines
- Goal: the distribution is valid and optimal
- Binary variable for QUBO formulation：

$$x_{i,\,j,\,s} = \begin{cases} 1, if\ transaction\ t_i\ is\ started\ at\ time\ s\ on\ machine\ m_j, \\ 0, otherwise \end{cases}$$

T. Bittner and S. Groppe. 2020. Avoiding Blocking by Scheduling Transactions using Quantum Annealing, IDEAS 2020

Transaction Management

Minimize logical formula

$$P = A + B + C + D$$

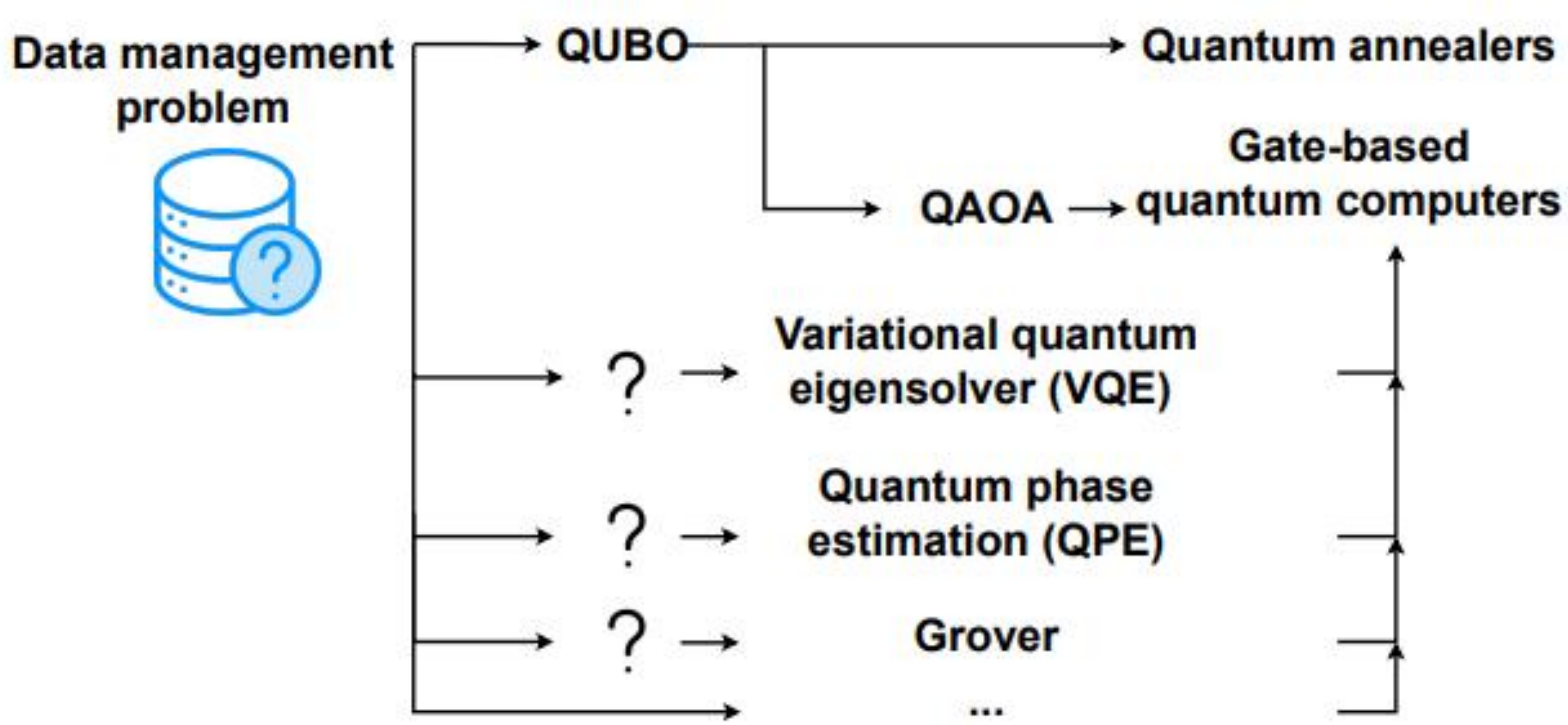
- Constraints for a valid schedule:
  - A: Each transaction starts exactly once
  - B: two or more transactions cannot be executed at the same time on the same machine
  - C: transactions that block each other cannot be executed at the
  - same time
- Optimal solution
  - D: requires the earliest possible start time for each transaction

More References

1. **K. Fritsch and S. Scherzinger, “Solving hard variants of database schema matching on quantum computers,” VLDB’23 Demo.**
2. **T. Bittner and S. Groppe, “Avoiding blocking by scheduling transactions using quantum annealing, IDEAS’20.**
3. Tim Bittner and Sven Groppe, “Hardware accelerating the optimization of transaction schedules via quantum annealing by avoiding blocking, OJCC’20.
4. S. Groppe and J. Groppe, “Optimizing transaction schedules on universal quantum computers via code generation for grover’s search algorithm, IDEAS’21.
5. L. Gruenwald, T. Winker, U. Çalikyilmaz, J. Groppe, and S. Groppe. "Index Tuning with Machine Learning on Quantum Computers for Large-Scale Database Applications." Joint Workshops at 49th International Conference on Very Large Data Bases (VLDBW’23)—International Workshop on Quantum Data Science and Management (QDSM’23). 2023.
6. G. Yuan, et al. "Quantum Computing for Databases: A Short Survey and Vision." Joint Workshops at 49th International Conference on Very Large Data Bases (VLDBW’23)—International Workshop on Quantum Data Science and Management (QDSM’23). 2023.

Roadmap

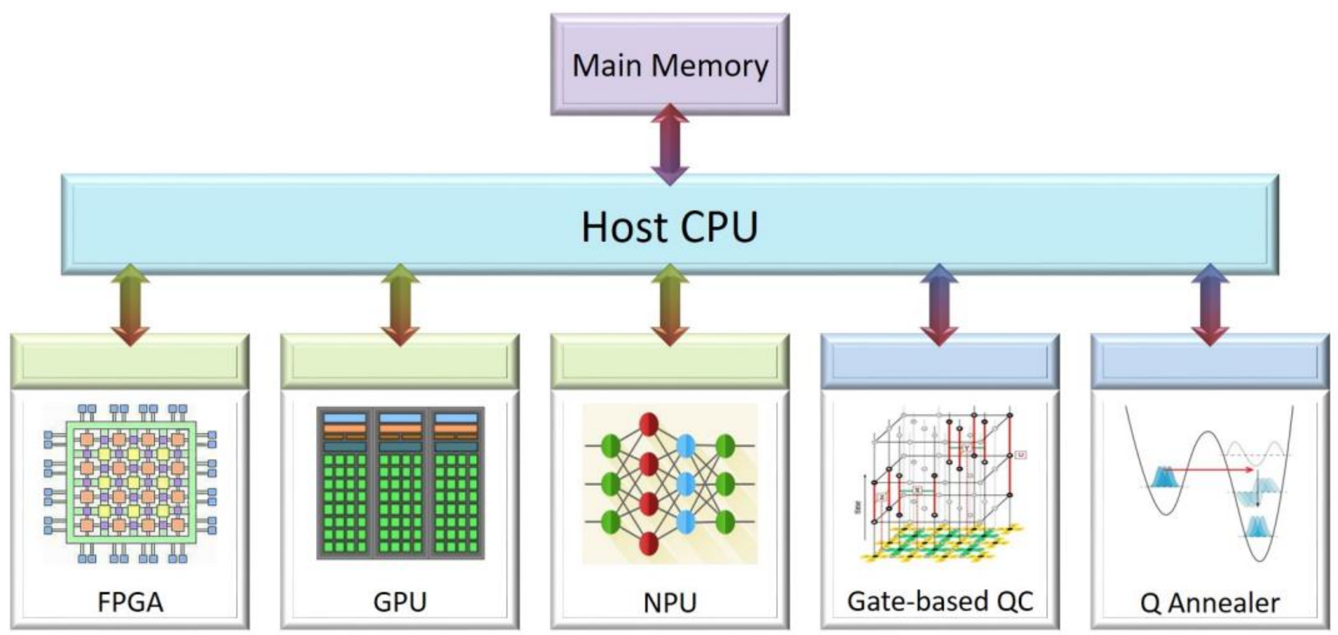
- Solving data management problems on quantum computers
  - Problem benefit from **quantum advantage**, and practically useful
    - Optimization problem
    - Classical approaches have scaling limits
    - Yet it does not require to load a large classical dataset
  - Convert a data management solution to quantum algorithms
  - **Constraints** of current quantum hardware



Research Opportunities

- DB problem **reformulation**
- **Hybrid** approach on classical and quantum computers
- Optimization given quantum computer **constraints**

Quantum computer will enhance, not replace, current HPC systems



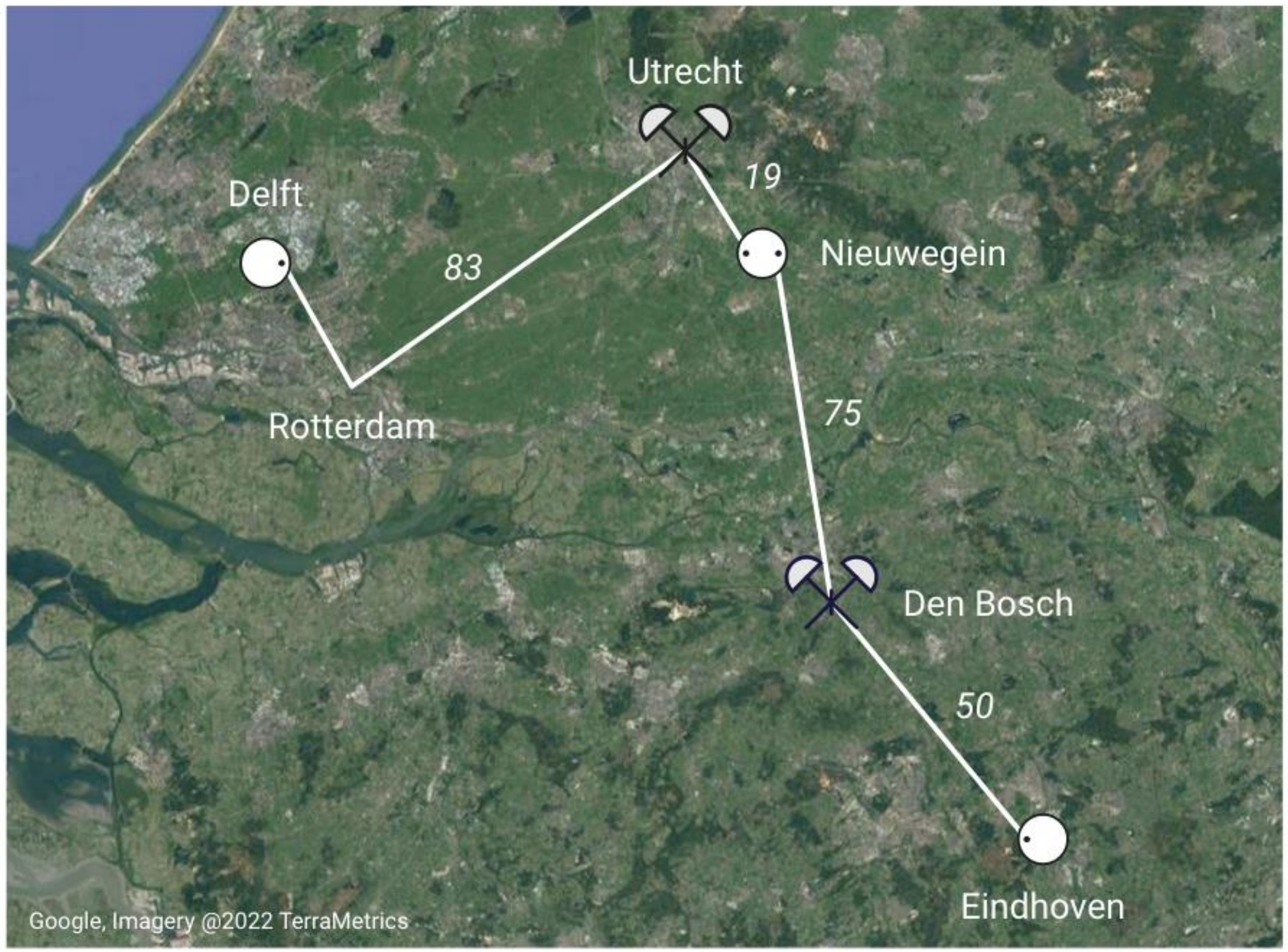
Data Management via Quantum Internet

What is a Quantum Internet?

https://www.youtube.com/watch?v=PCKoT9xcyXI&t=1s

Hypothetical Quantum Internet Connection

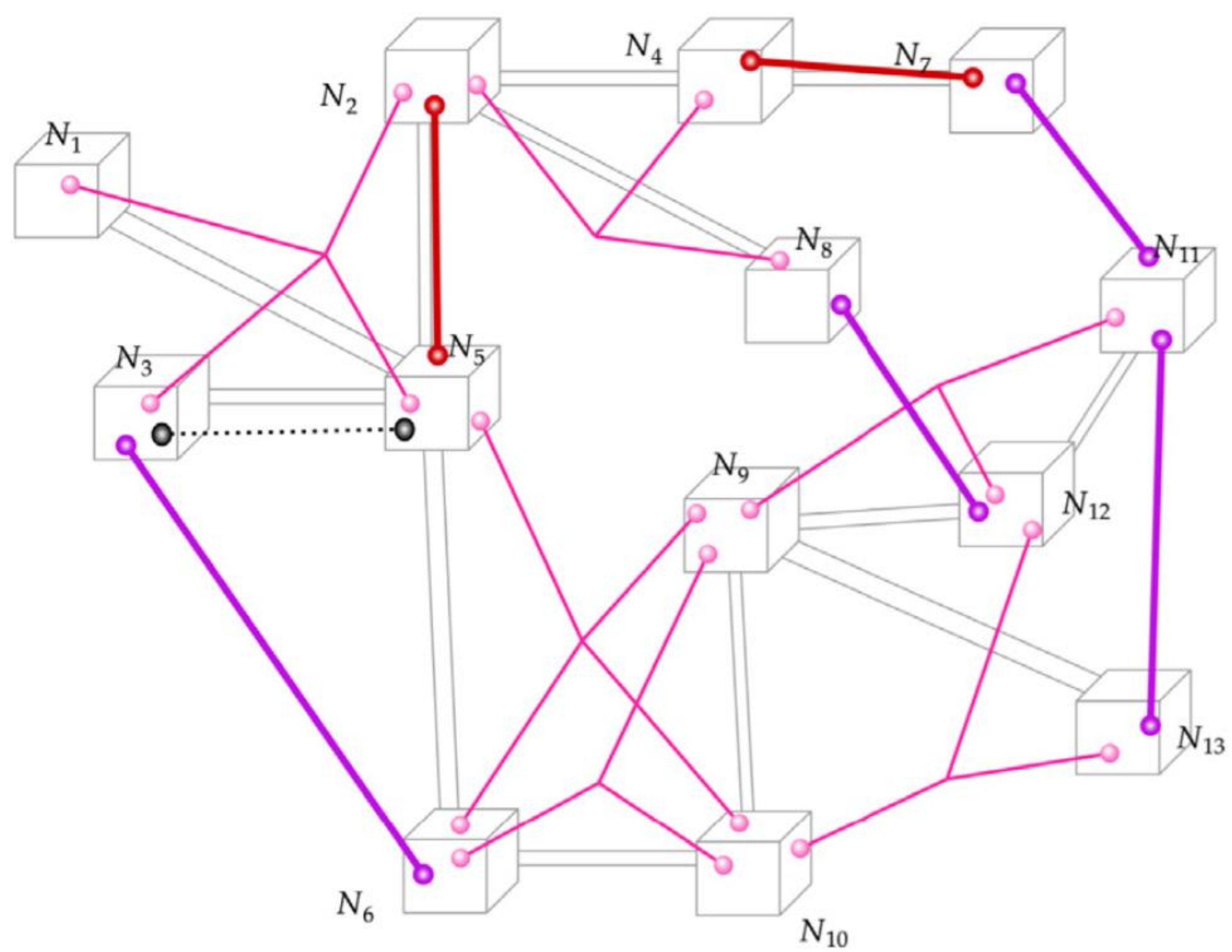
Total fiber distance between Delft and Eindhoven of **226.5** km



Avis, Guus, et al. "Requirements for a processing-node quantum repeater on a real-world fiber grid." NPJ Quantum Information 9.1 (2023): 100.

Quantum Internet

- Like classical internet, quantum internet (QI) allows for information exchange between nodes.
  - QI extends CI to allow joint quantum information processing
  - Physical links: channels established to exchange classical messages
  - **Virtual links**: shared **entangled states** between the nodes



Illiano, Jessica, et al. "Quantum internet protocol stack: A comprehensive survey." Computer Networks 213 (2022): 109092.

Different Architectures to Implement Qubits

A bit of the action

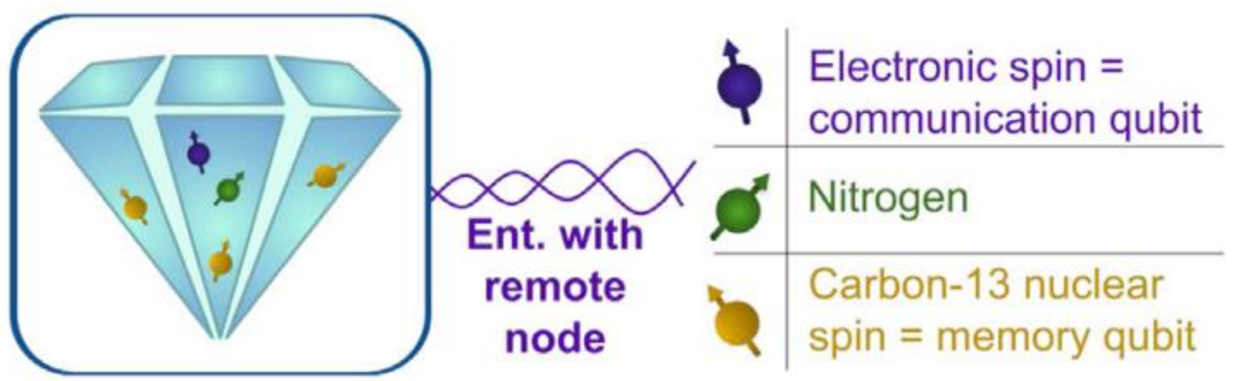
In the race to build a quantum computer, companies are pursuing many types of quantum bits, or qubits, each with its own strengths and weaknesses.

	Superconducting loops	Trapped ions	Silicon quantum dots	Topological qubits	Diamond vacancies
	A resistance-free current oscillates back and forth around a circuit loop. An injected microwave signal excites the current into super-position states.	Electrically charged atoms, or ions, have quantum energies that depend on the location of electrons. Tuned lasers cool and trap the ions, and put them in superposition states.	These "artificial atoms" are made by adding an electron to a small piece of pure silicon. Microwaves control the electron's quantum state.	Quasiparticles can be seen in the behavior of electrons channeled through semi-conductor structures. Their braided paths can encode quantum information.	A nitrogen atom and a vacancy add an electron to a diamond lattice. Its quantum spin state, along with those of nearby carbon nuclei, can be controlled with light.
Longevity (seconds)	0.00005	>1000	0.03	N/A	10
Logic success rate	99.4%	99.9%	~99%	N/A	99.2%
Number entangled	9	14	2	N/A	6
Company support	Google, IBM, Quantum Circuits	ionQ	Intel	Microsoft, Bell Labs	Quantum Diamond Technologies
Pros	Fast working. Build on existing semiconductor industry.	Very stable. Highest achieved gate fidelities.	Stable. Build on existing semiconductor industry.	Greatly reduce errors.	Can operate at room temperature.
Cons	Collapse easily and must be kept cold.	Slow operation. Many lasers are needed.	Only a few entangled. Must be kept cold.	Existence not yet confirmed.	Difficult to entangle.

Note: Longevity is the record coherence time for a single qubit superposition state, logic success rate is the highest reported gate fidelity for logic operations on two qubits, and number entangled is the maximum number of qubits entangled and capable of performing two-qubit operations.

Source: https://www.science.org/doi/10.1126/science.354.6316.1090#F2

Quantum Internet Nodes



(a) the single-NV architecture

Heralded entanglement 1.3km

Mirrors and filters guide the laser beams to the diamond chip

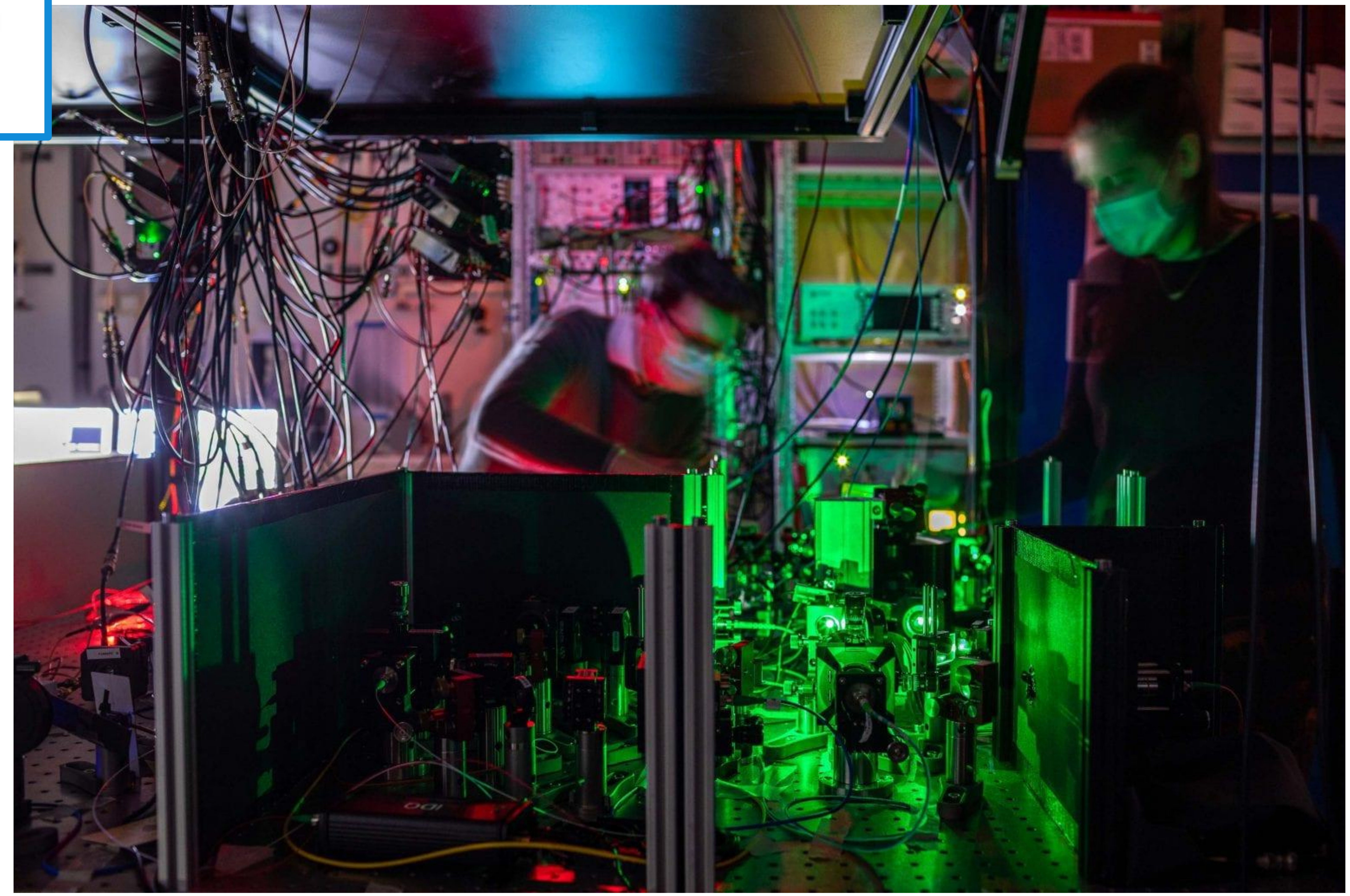
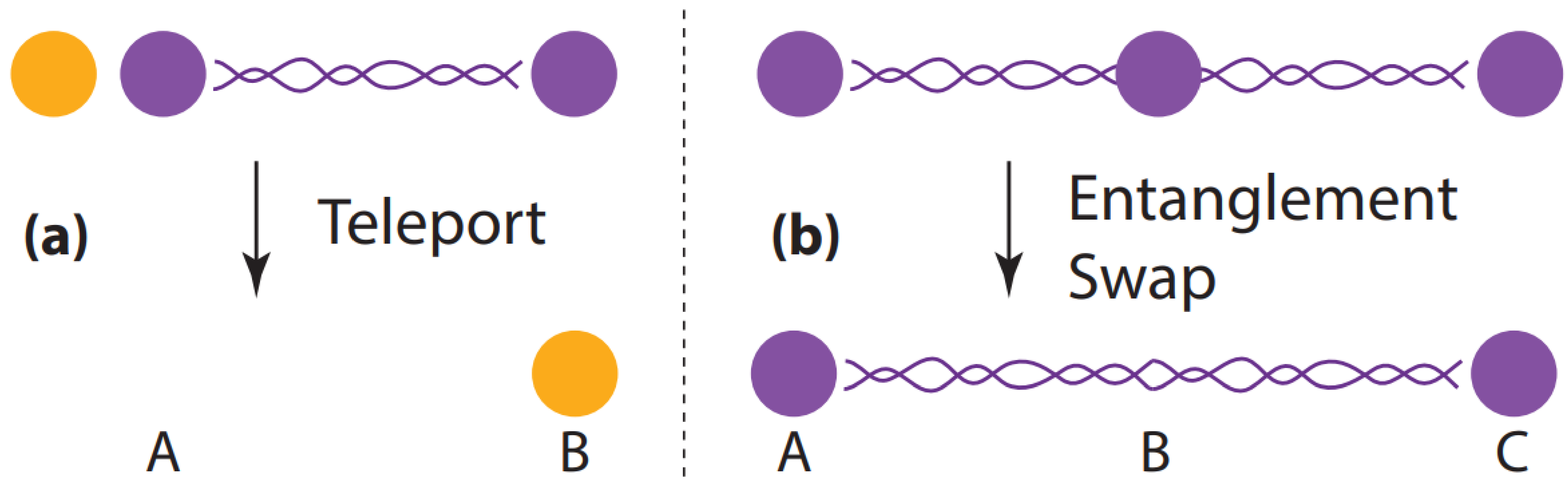


Photo by Marieke de Lorijn for QuTech

How to Distribute Entanglements?

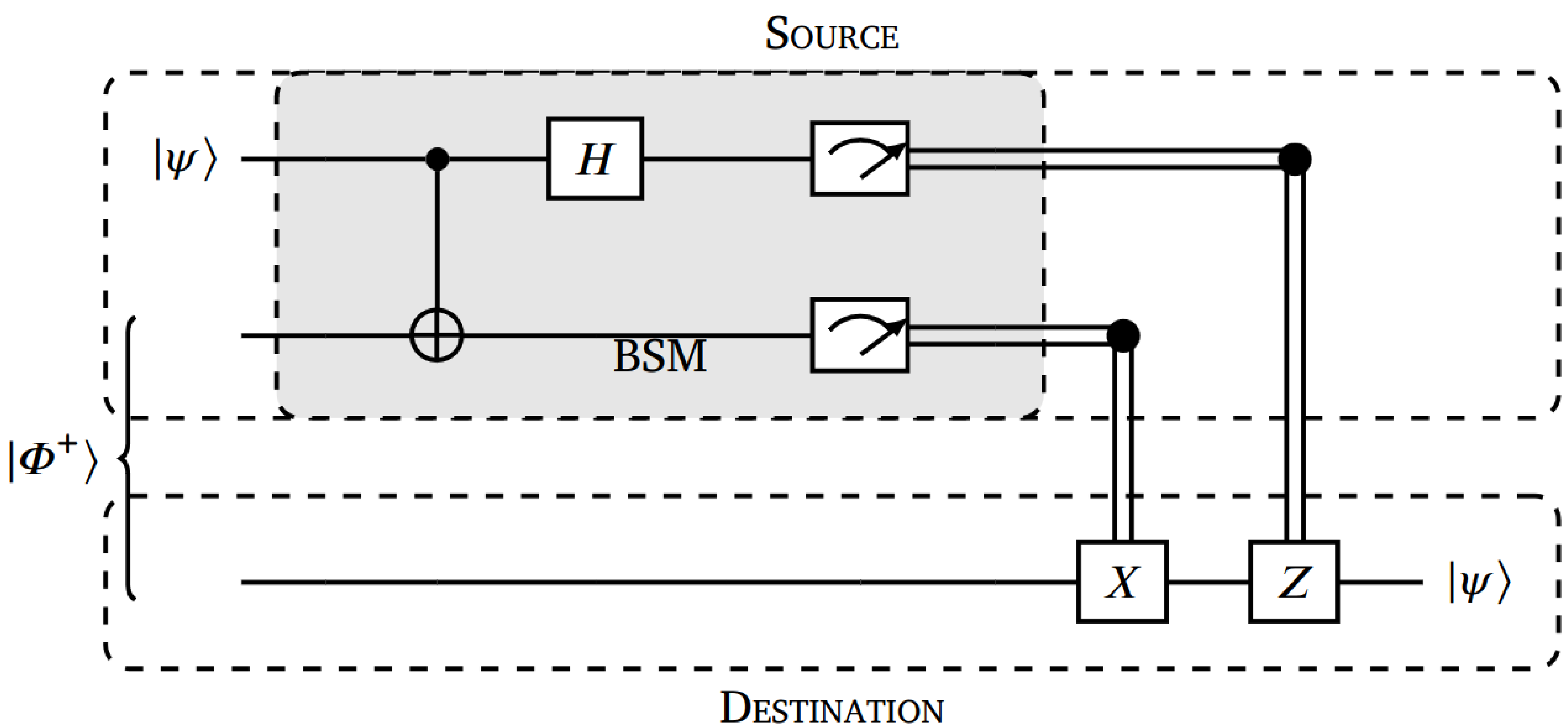


Quantum Teleportation

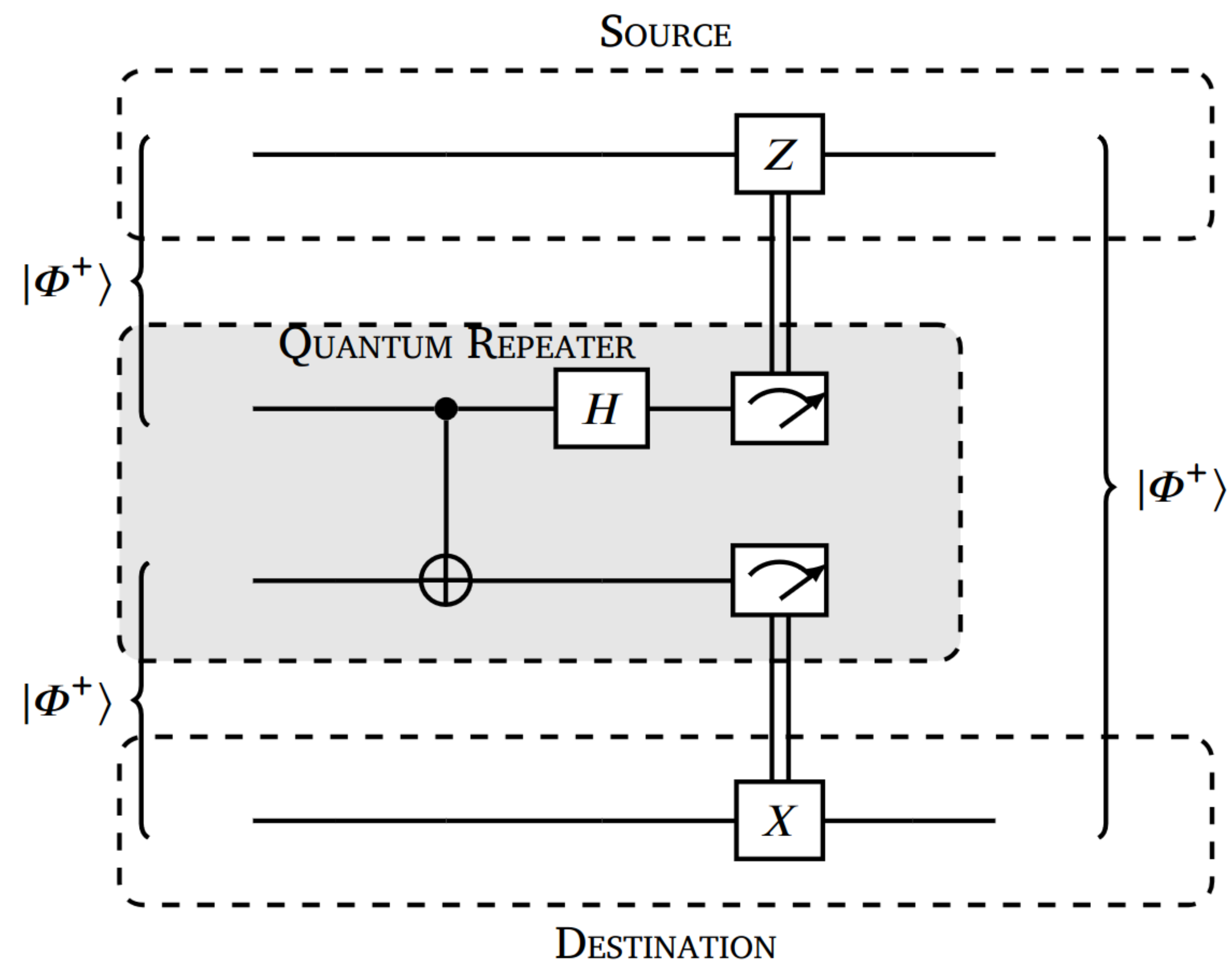
Entanglement Swapping

Dahlberg, Axel, Matthew Skrzypczyk, Tim Coopmans, Leon Wubben, Filip Rozpędek, Matteo Pompili, Arian Stolk et al. "A link layer protocol for quantum networks." In Proceedings of the ACM special interest group on data communication, pp. 159-173. 2019.

How to Distribute Entanglements?



Quantum Teleportation

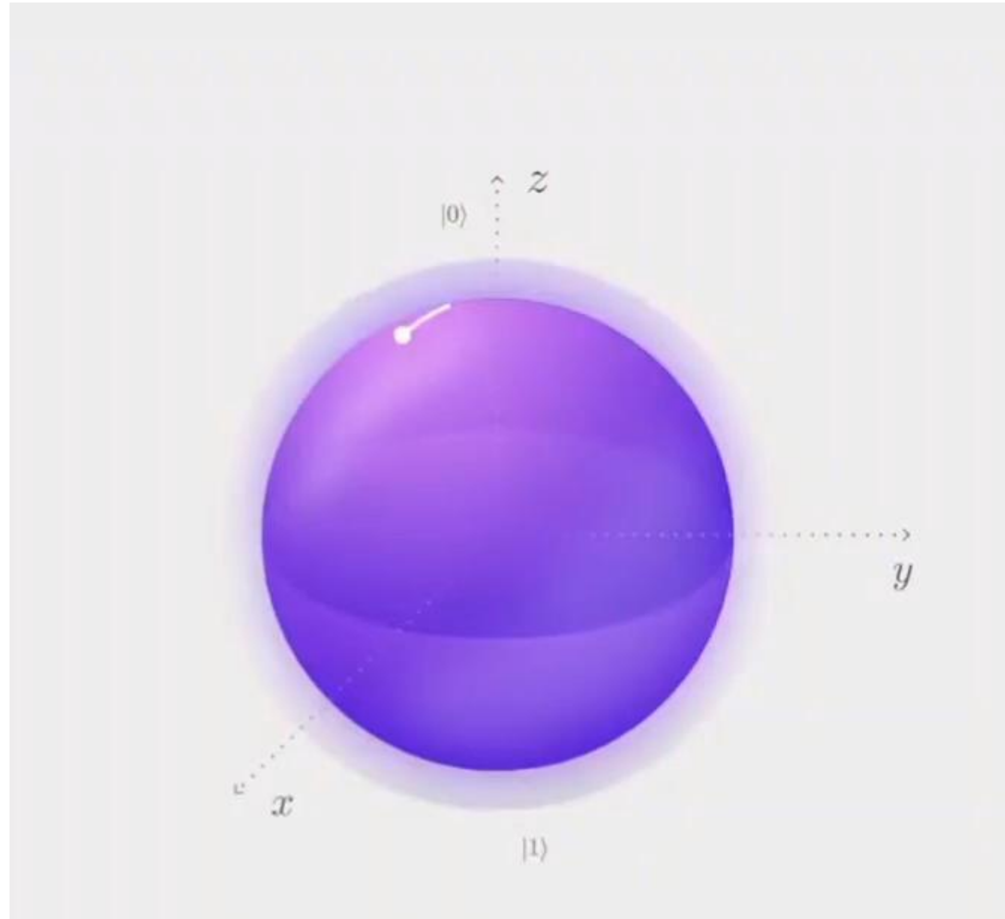


Entanglement Swapping

Illiano, Jessica, et al. "Quantum internet protocol stack: A comprehensive survey." Computer Networks 213 (2022): 109092.

Quantum Noise

- Quantum noise makes it hard to extract information from a quantum computer
- Quantum noise results from unwanted coupling with the environment
  - Depolarizing
  - Bit & phase flipping
  - Amplitude & phase damping



https://physicsworld.com/a/putting-quantum-noise-to-work/
https://www.youtube.com/watch?v=AUAkoEiutOE

Detect Incorrect Quantum Info Processing?

- We expect to only have **noisy intermediate-scale quantum (NISQ)** devices in the near future
  - Handle quantum noise by fixing or removing corruption of quantum data?
  - Detect incorrect quantum operations
  - Generate robust entanglements

Google's Sycamore

USTC's Jiuzhang

What does Quantum Internet Give?

- Full-fledged quantum internet seems necessary for multiparty quantum computation

- Seems a good way to avoid the **synchronous problem**?
  - Einstein–Podolsky–Rosen (EPR) paradox

- Entanglements enhance data integrity
  - Nonlocal games are useful to detect **information leakage** and **incorrect operations,** e.g., quantum key distribution (impossible classically)
  - Enhance security of communication impossible by purely classical means

Quantum Nonlocality

- Correlations that have no interpretations using random coins
- **Bell's state**: a **maximally entangled state** that can be used to achieve non-locality

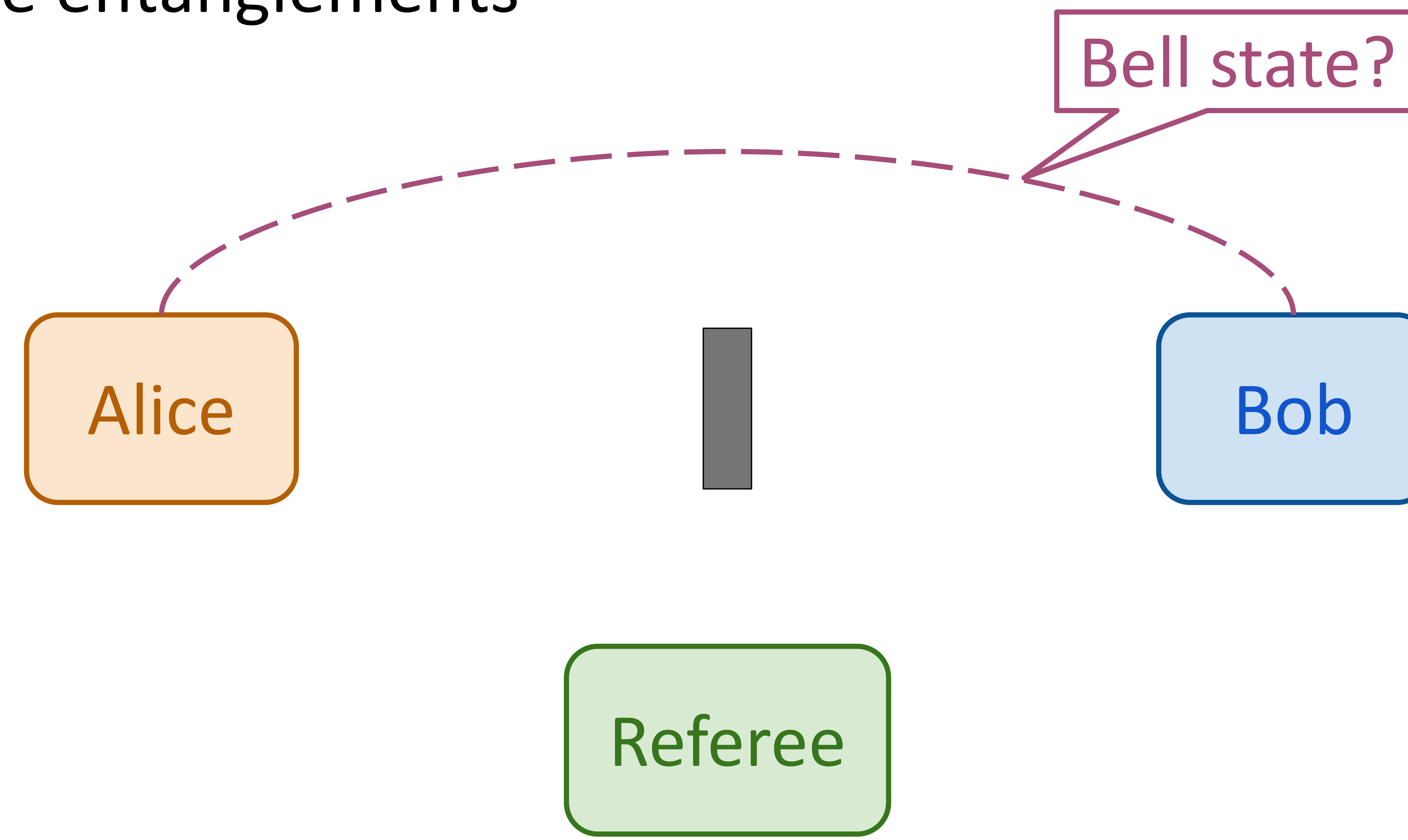
$$|\Psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$$

Measuring any half of the state in any basis, the other "collapses" to the same outcome.

- Nonlocal games: Tests used to test nonlocality, specifically, the shared entanglements!

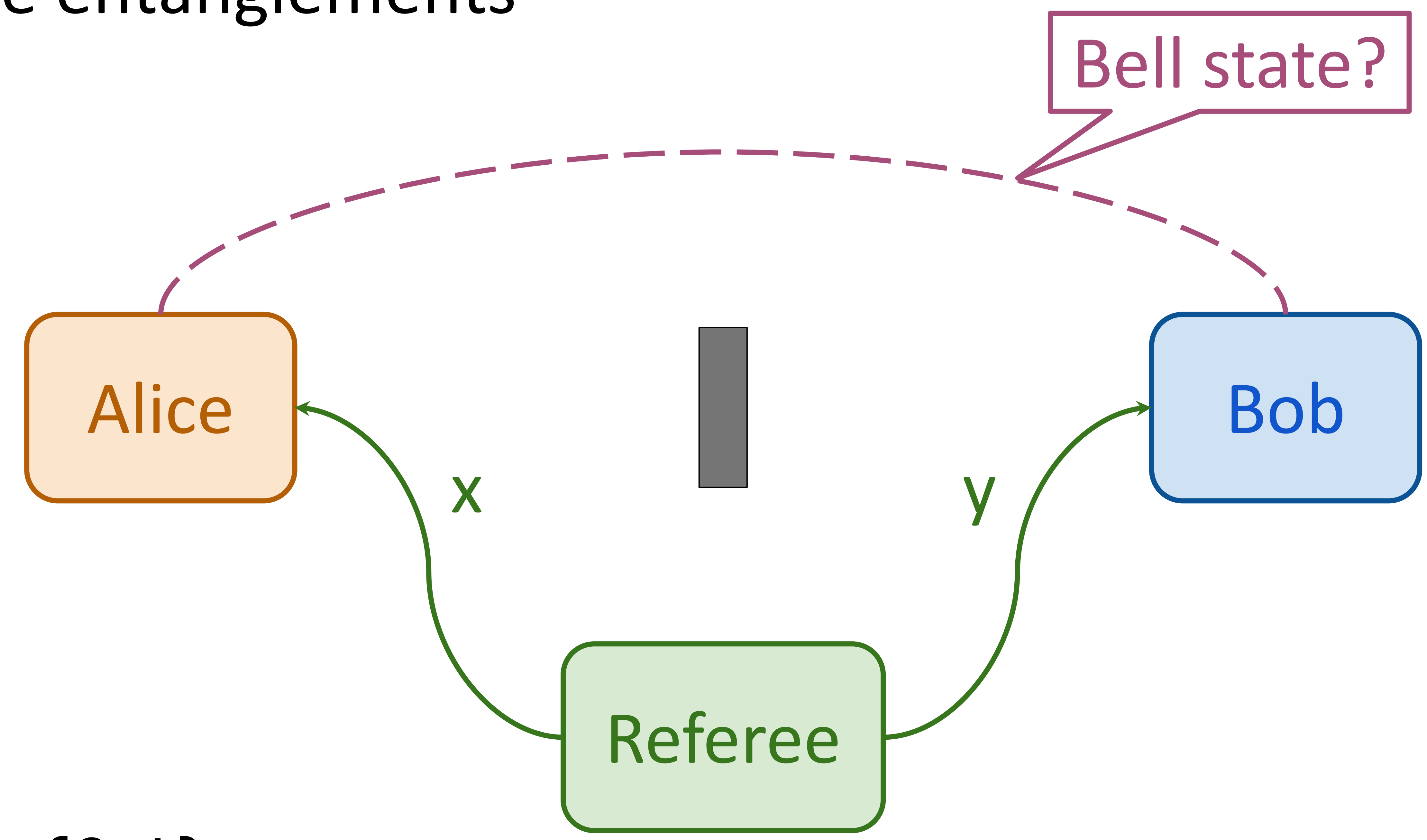
The Clauser-Horne-Shimony-Holt (CHSH) Game

- Two devices that are not allowed to communicate, but can share entanglements



The Clauser-Horne-Shimony-Holt (CHSH) Game

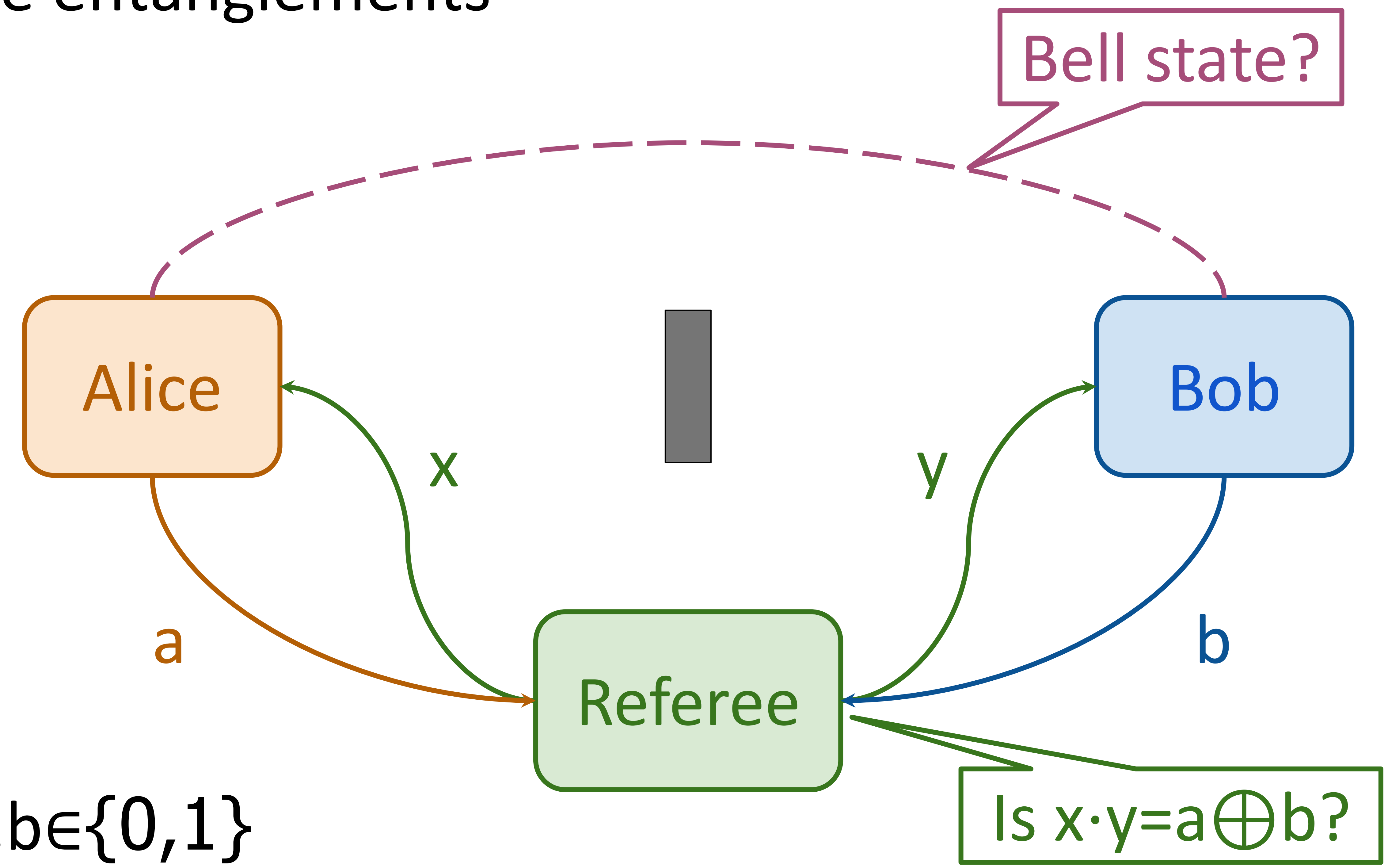
- Two devices that are not allowed to communicate, but can share entanglements



x,y∈{0,1}

The Clauser-Horne-Shimony-Holt (CHSH) Game

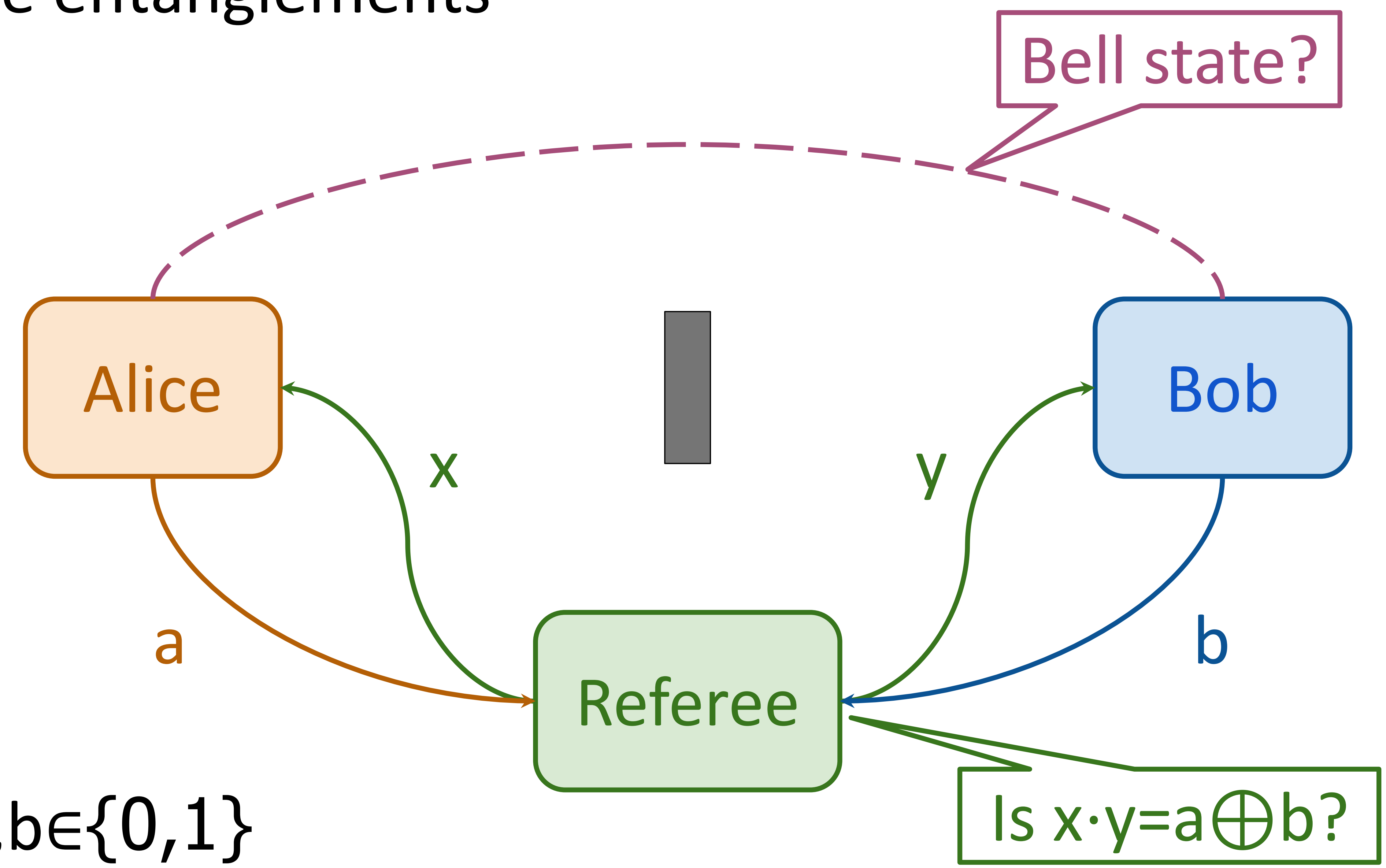
- Two devices that are not allowed to communicate, but can share entanglements



x,y,a,b∈{0,1}

The Clauser-Horne-Shimony-Holt (CHSH) Game

- Two devices that are not allowed to communicate, but can share entanglements

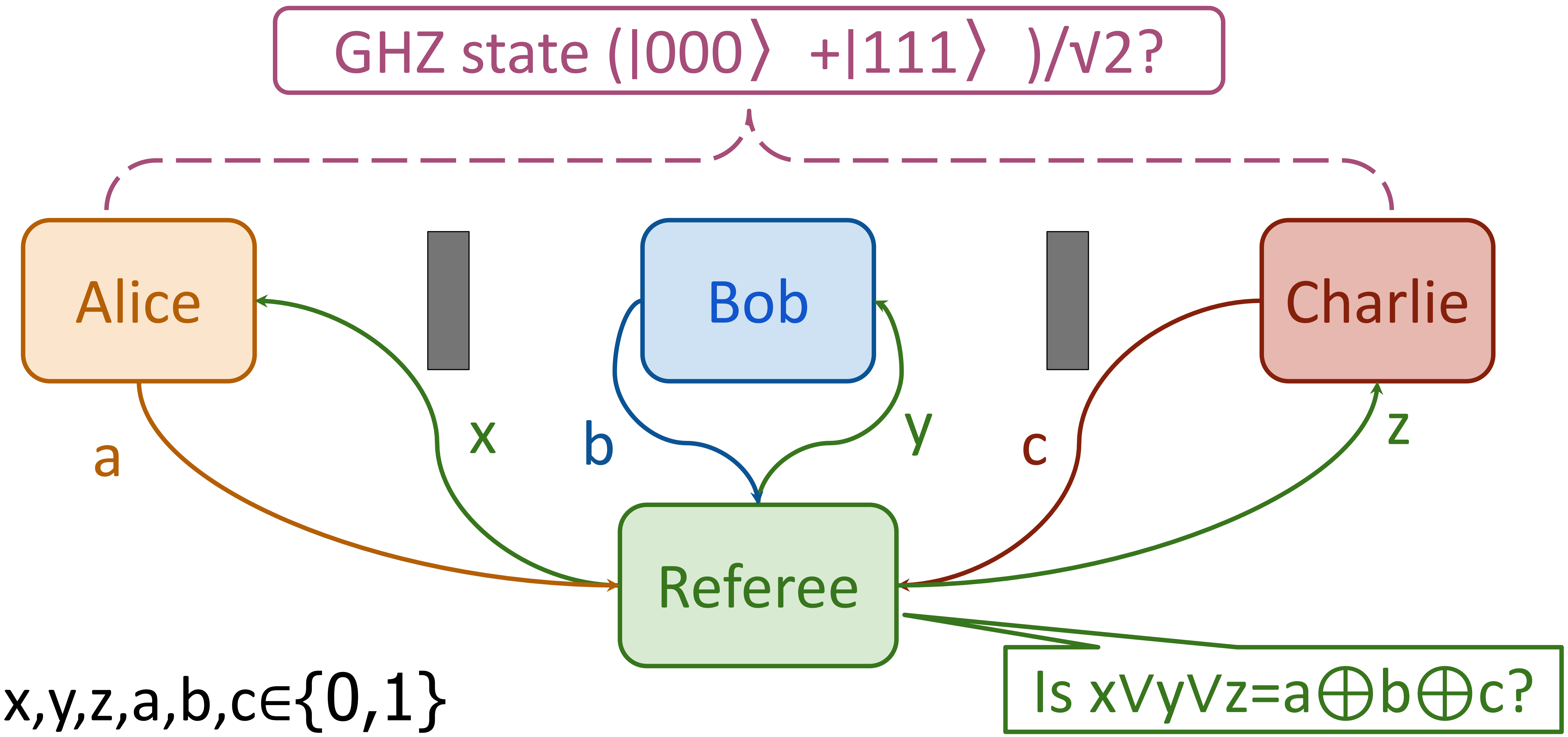


$x, y, a, b \in \{0,1\}$

With a Bell state, they can win the test with probability ≈0.85!

The Greenberger-Horne-Zeilinger (GHZ) Game

- Extension to test multipartite entanglements
- A three-player non-local game for testing GHZ state



x,y,z,a,b,c∈{0,1}

With a GHZ state, they can win the test with probability **1**!

Beyond the Binary Alphabet

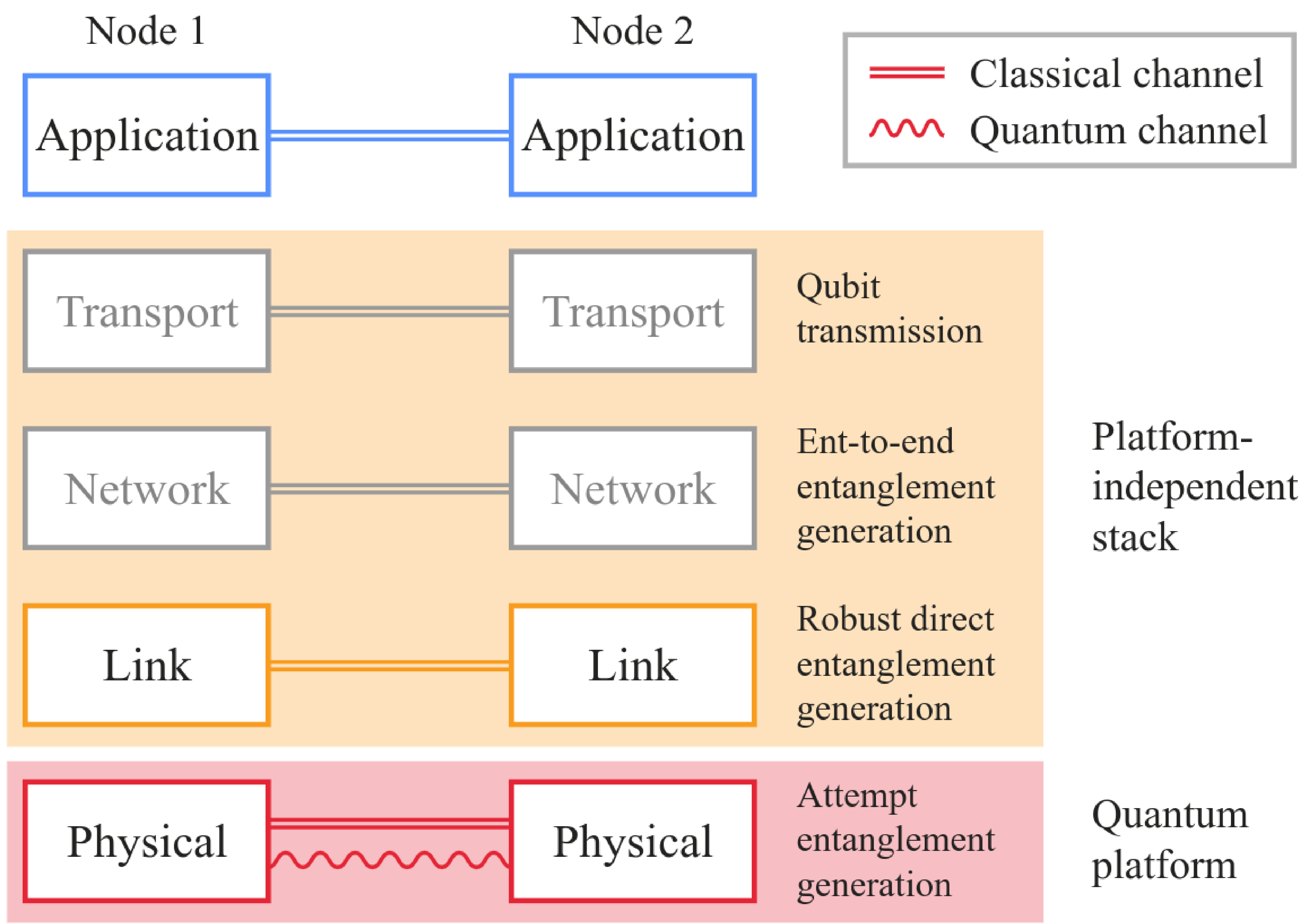
For data management, it is essential to consider the space of data beyond binary strings.

- The tests may be generalized to "qudits," which takes values from an alphabet of size d>2 (as opposed to {0,1}).

pname	gender	age	resting HR	mortality
Alice	0	96	60	0
Bob	1	35	70	1
Charlie	1	22	65	1

Quantum Internet Protocol Stacks

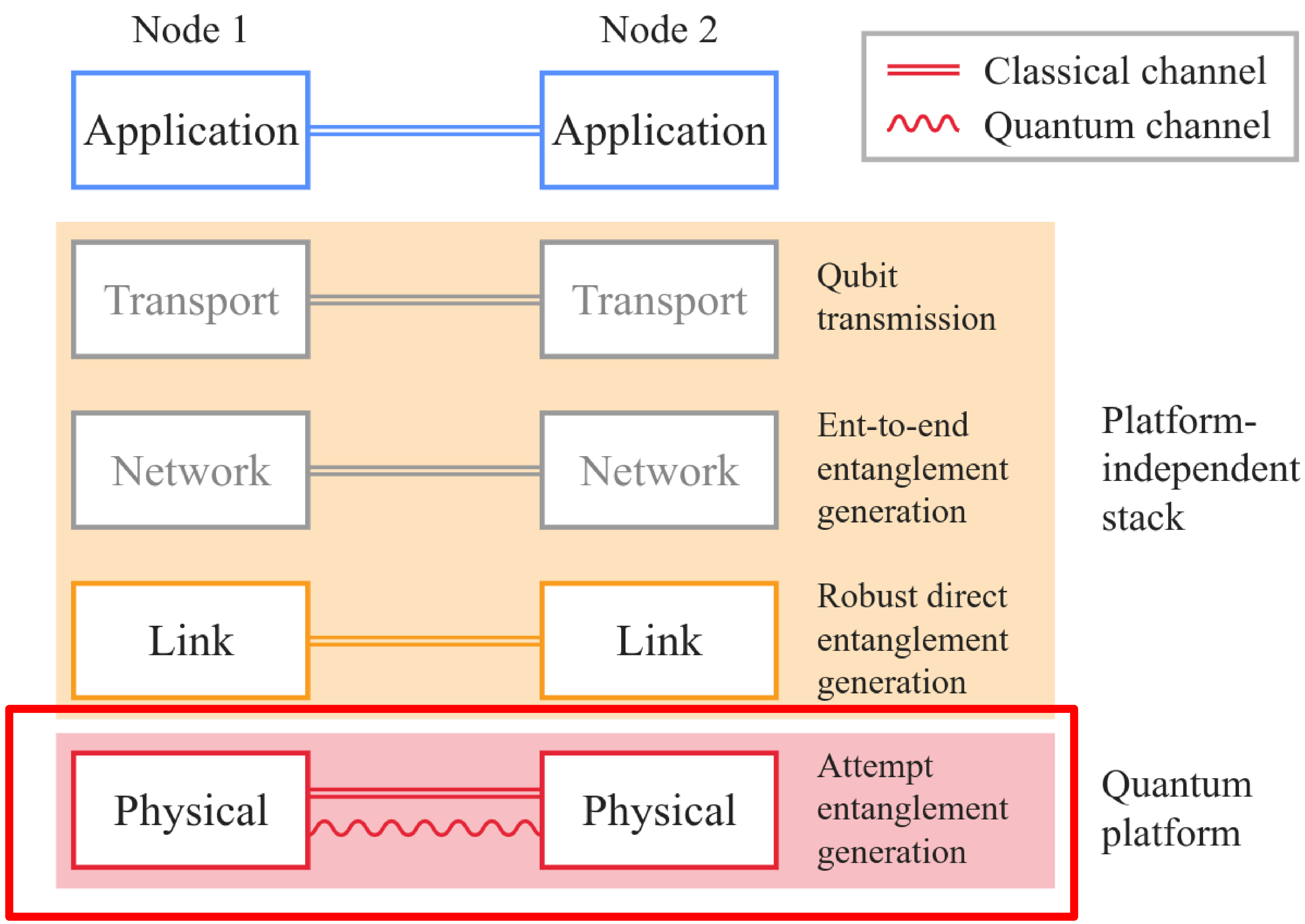
Protocols for scheduling the generation and distribution of quantum entanglements in a quantum network



Pompili, Matteo, et al. "Experimental demonstration of entanglement delivery using a quantum network stack." npj Quantum Information 8.1 (2022): 121.

Quantum Internet Protocol Stacks

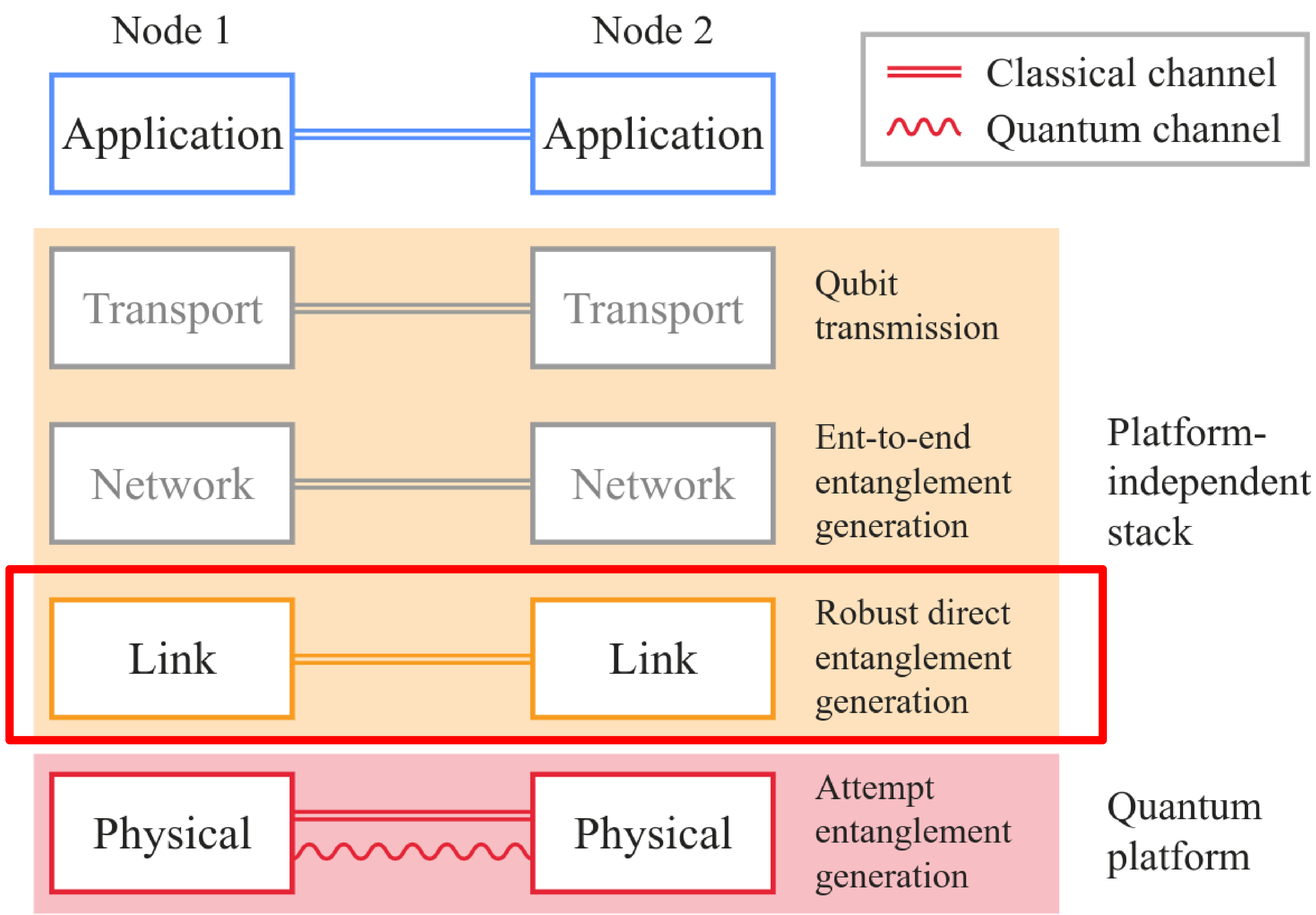
Protocols for scheduling the generation and distribution of quantum entanglements in a quantum network



The **Physical Layer** attempts to generate entanglements between two nodes in a well-defined time slot.

Quantum Internet Protocol Stacks

Protocols for scheduling the generation and distribution of quantum entanglements in a quantum network

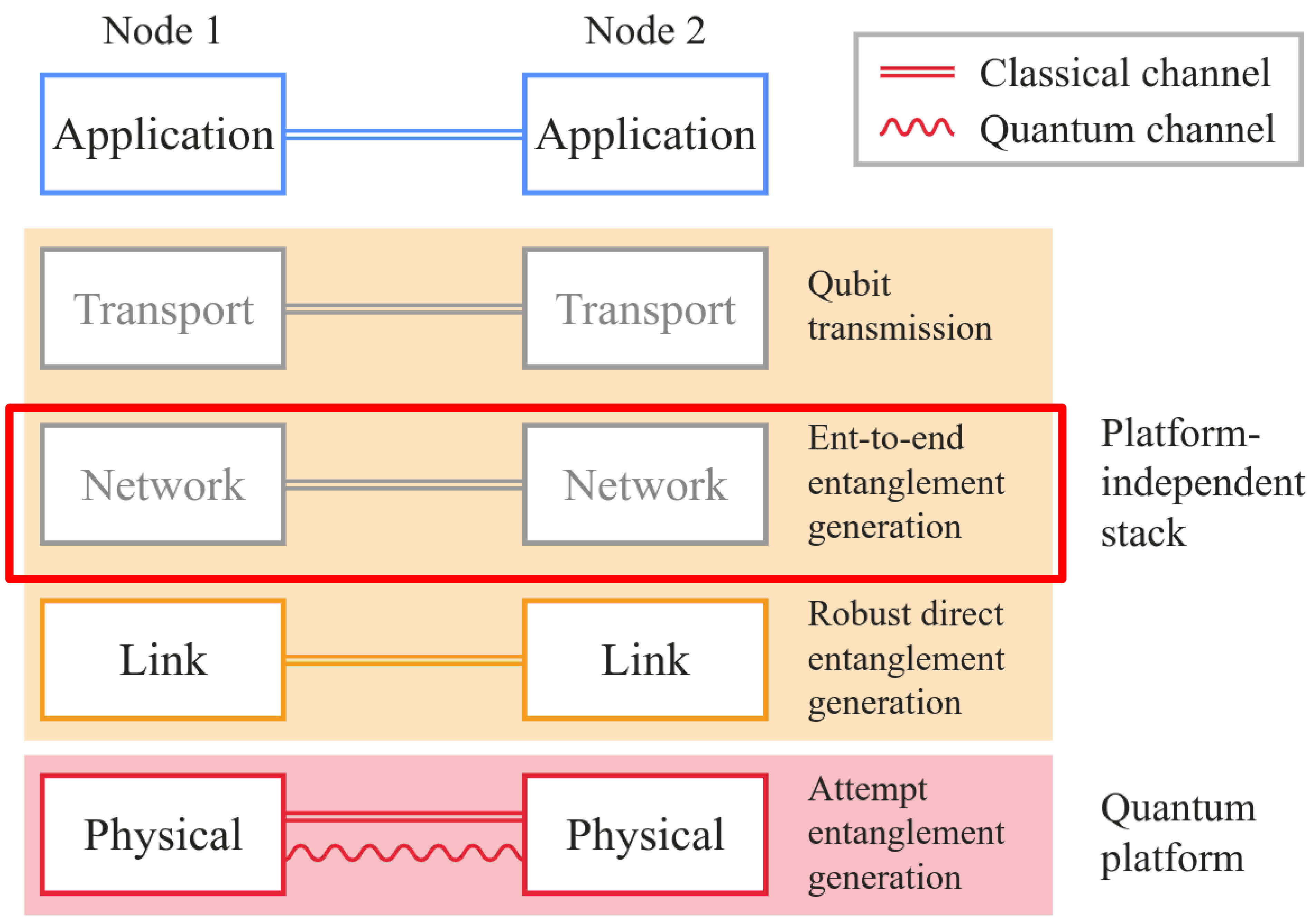


The **Link Layer** manages to generate robust entanglements:

- Receives generation requests
- Perform fidelity evaluation
- Scheduling generation

Quantum Internet Protocol Stacks

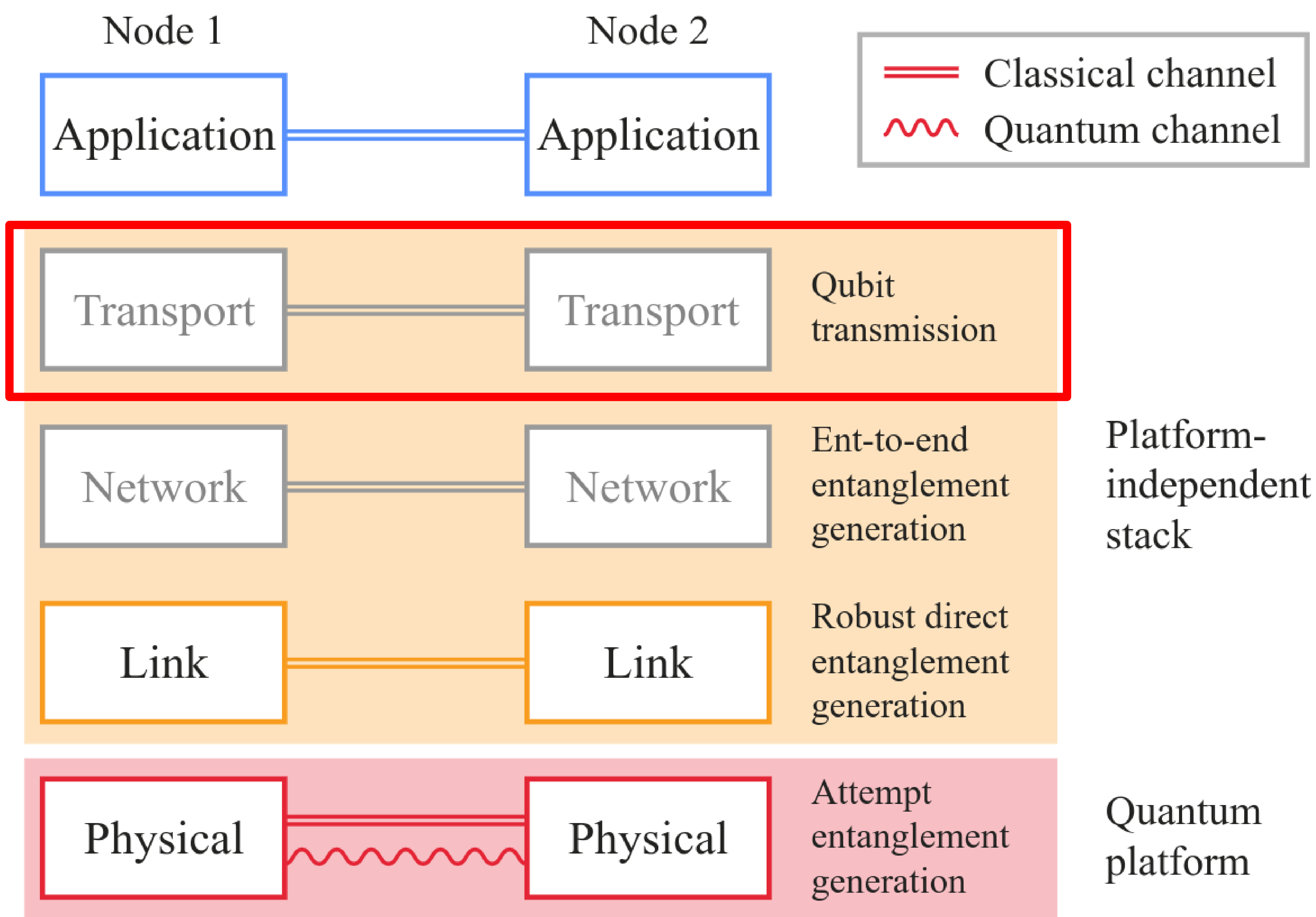
Protocols for scheduling the generation and distribution of quantum entanglements in a quantum network



The **Network Layer** is for producing long-distance entanglements using functionalities provided by the link layer.

Quantum Internet Protocol Stacks

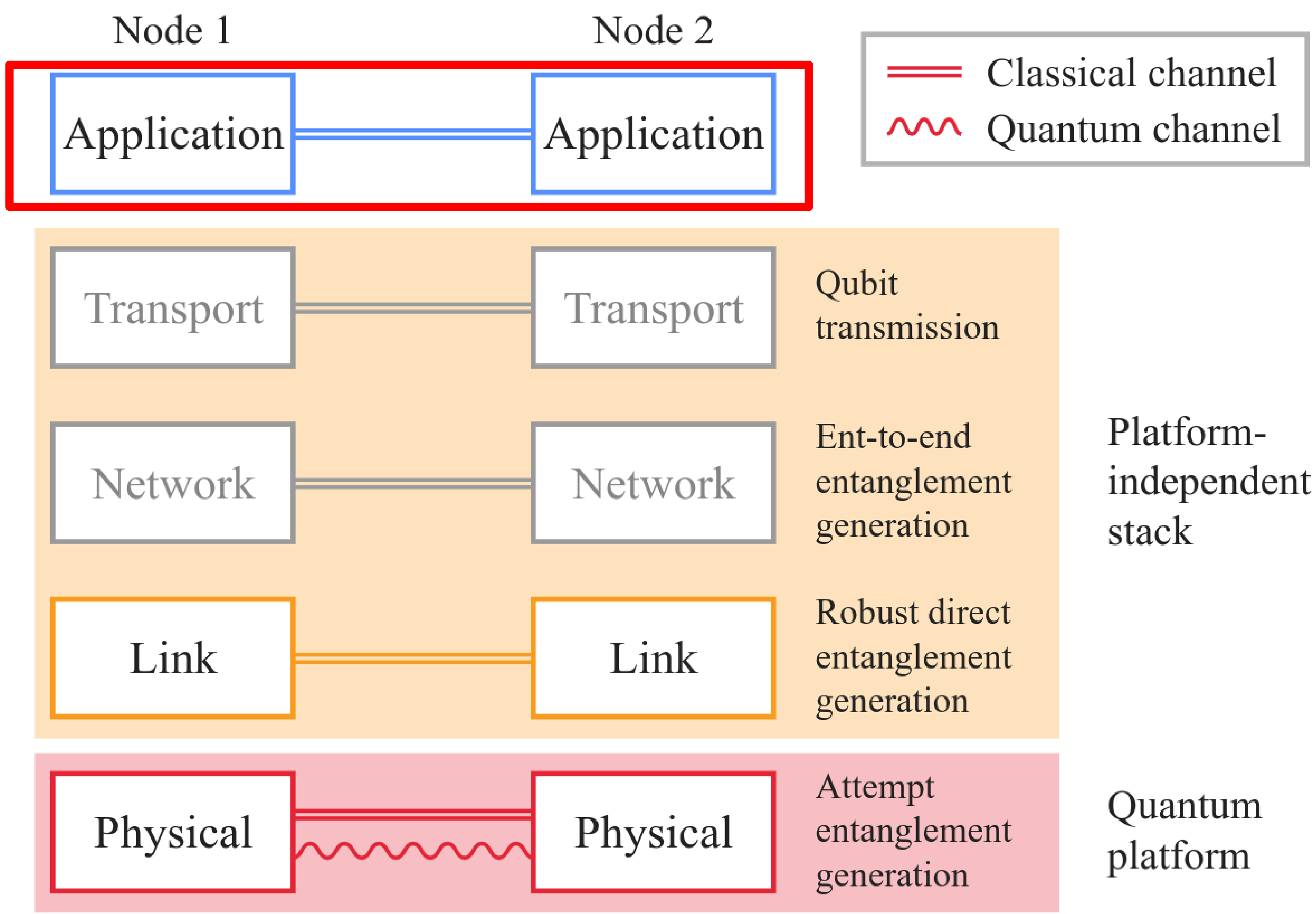
Protocols for scheduling the generation and distribution of quantum entanglements in a quantum network



The **Transport Layer** transmits qubits by using the teleportation process.

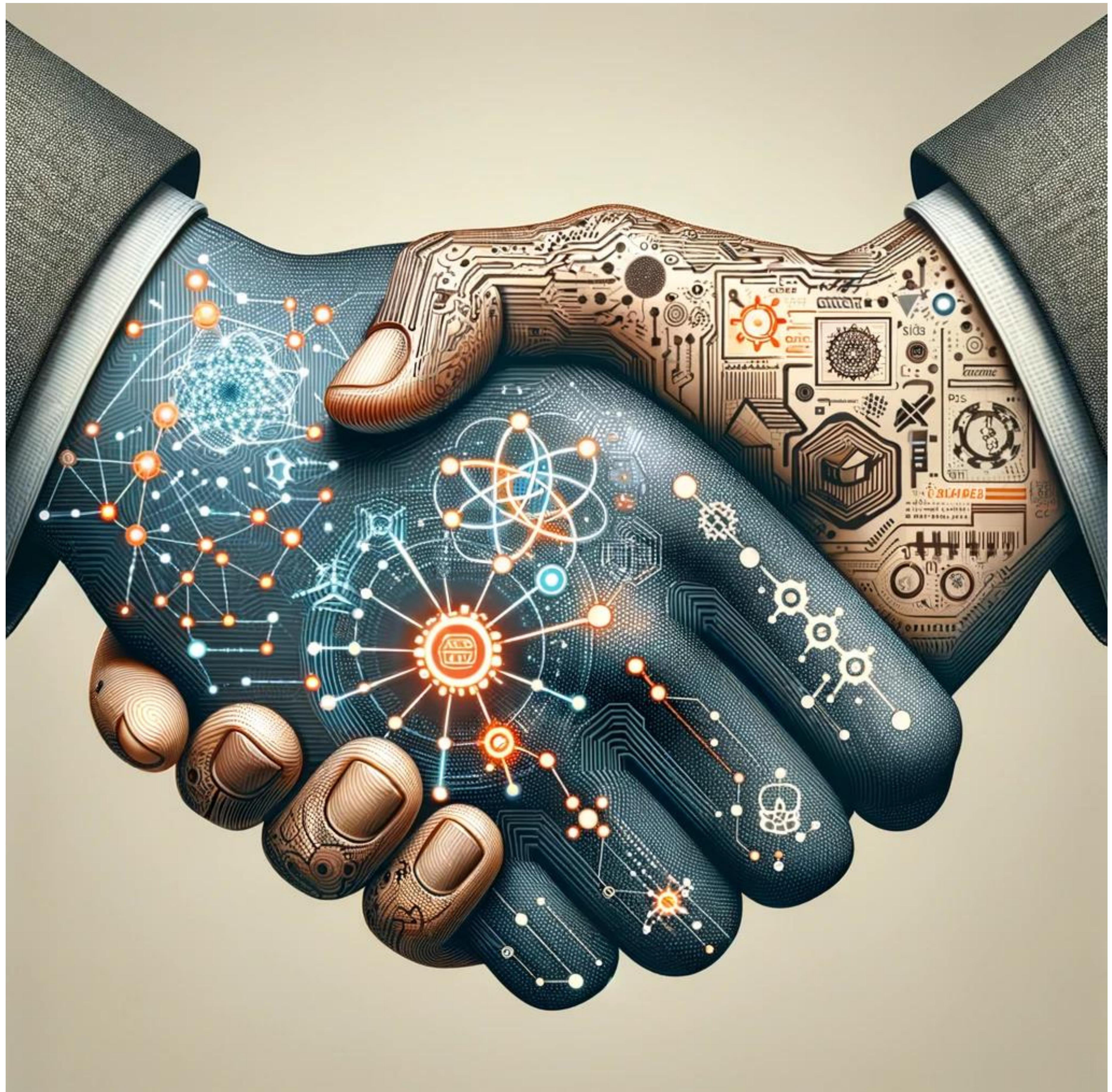
Quantum Internet Protocol Stacks

Protocols for scheduling the generation and distribution of quantum entanglements in a quantum network



The **Application Layer** controls the network abstractly to distribute quantum information and to perform joint quantum computation.

Future Directions



Q&A

- What challenge do you want to tackle in the NISQ era?

Quantum Computer

Quantum Internet