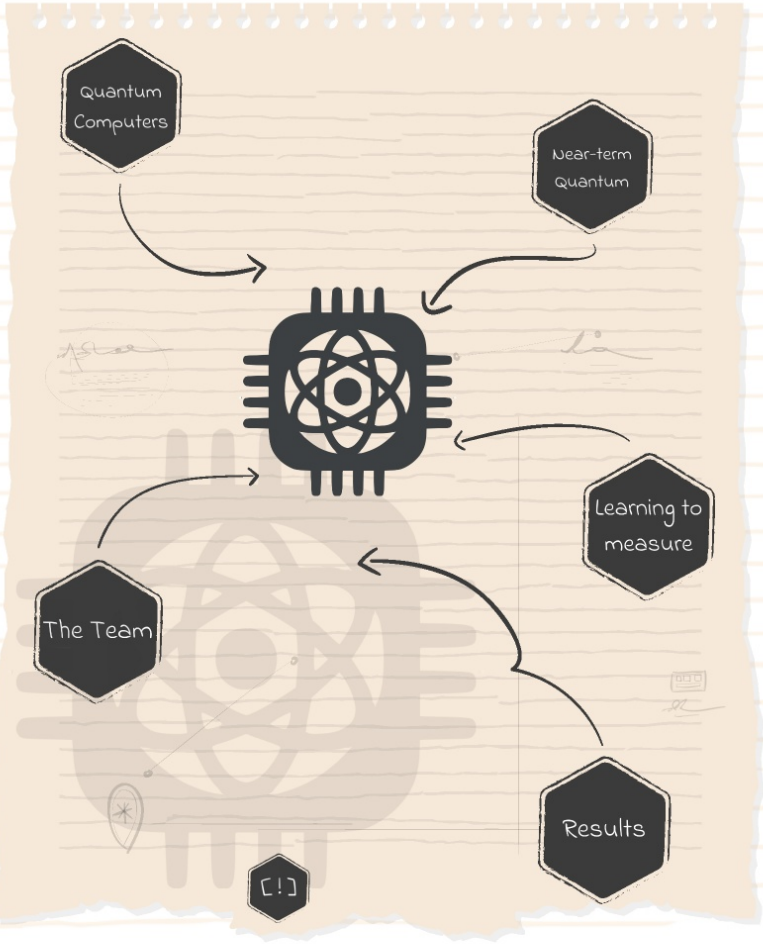


Learning to measure

Adaptive informationally complete POVMs for near-term quantum algorithms

Sabrina Maniscalco, University of Helsinki, Aalto University, Algorithmiq Oy - Finland

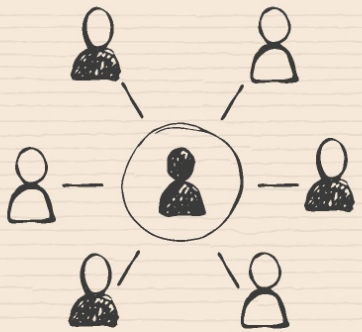
The team

Finnish Team

Guillermo Garcia-Perez, University of Turku, Algorithmiq
Matteo Rossi, University of Turku, Algorithmiq
Boris Sokolov, University of Helsinki, Algorithmiq

IBM Quantum Team

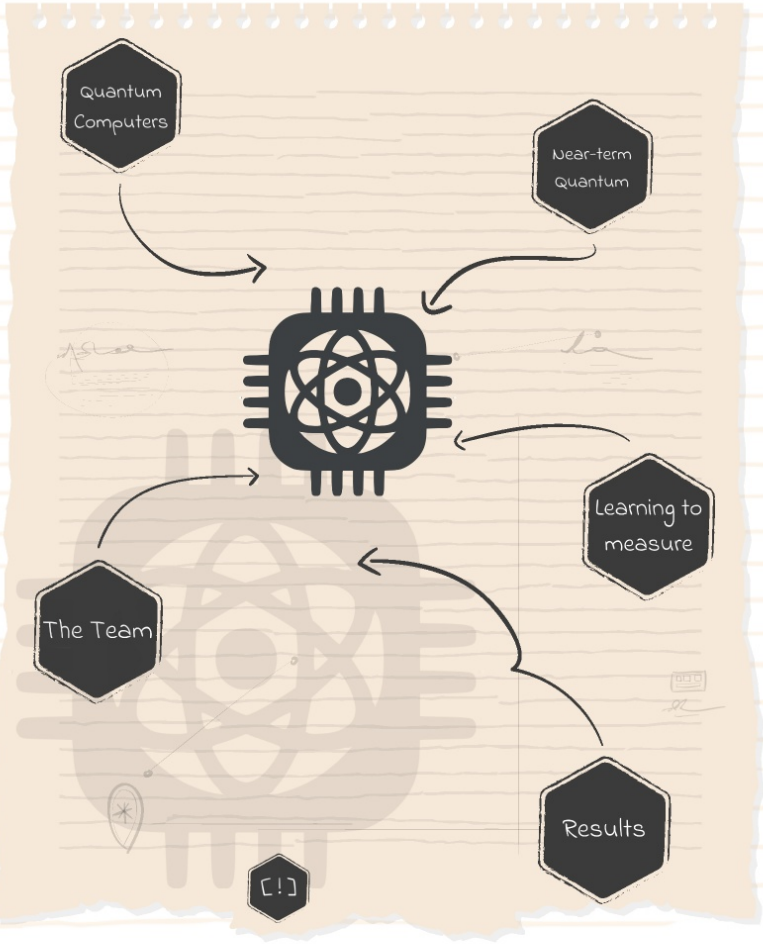
Francesco Tacchino, IBM Quantum, Zurich, Switzerland
Panagiotis Barkoutsos, IBM Quantum, Zurich, Switzerland
Guglielmo Mazzola, IBM Quantum, Zurich, Switzerland
Ivano Tavernelli, IBM Quantum, Zurich, Switzerland



where it all began

Learning to measure

Adaptive informationally complete POVMs for near-term quantum algorithms



Sabrina Maniscalco, University of Helsinki, Aalto University, Algorithmiq Oy - Finland

Quantum Computing

The Goal

Understanding if, and to which extent, it is possible to build computing machines able to perform tasks which are impossible to conventional (classical) computers

Quantum Algorithms

Universal QC

Decoherence

Quantum Algorithms

Quantum Algorithms, exploiting quantum properties of the physical information carriers, outperform all currently existing classical algorithms.

however, they run on ideal fault-tolerant universal quantum computers!

Universal Quantum Computers

However

A quantum Turing machine or universal quantum computer is an abstract machine used to model the effects of a quantum computer. Any quantum algorithm can be expressed formally as a particular quantum Turing machine.

Information is physical

Physical realization of quantum computers

Algorithms as dynamical quantum processes!

Decoherence

The ultimate enemy?

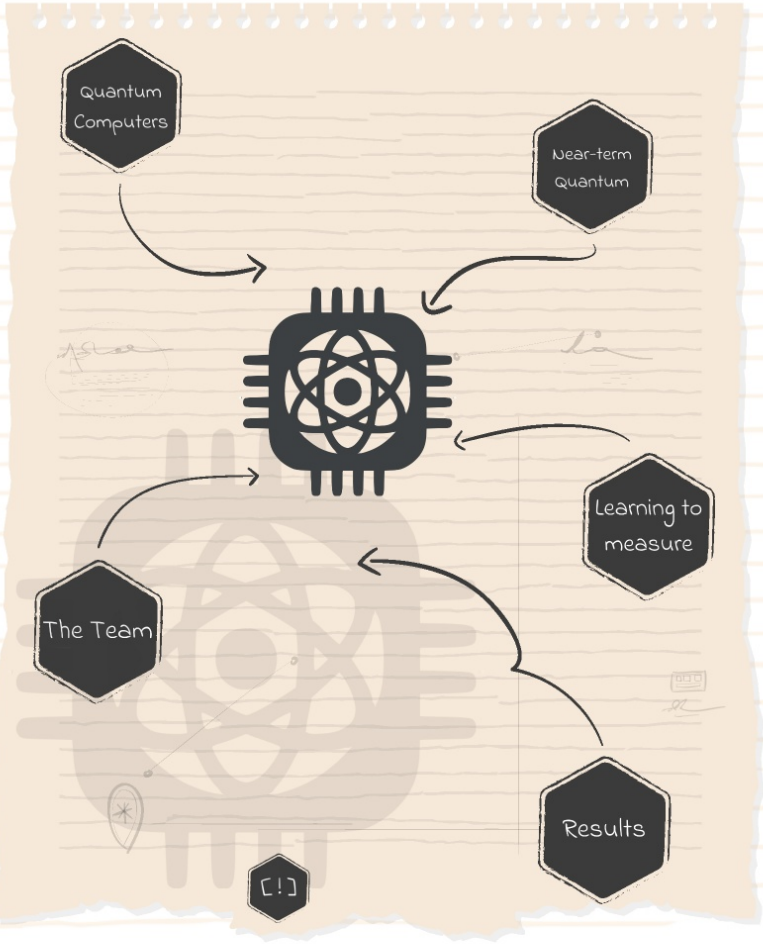
Because of the **extreme fragility of quantum information** storage and processing in presence of environmental noise, **error-correction techniques** required to achieve fault-tolerance are still experimentally in their infancy

The Grand Challenge

It is unclear, at the moment, whether decoherence will eventually be a fundamental unavoidable limiting factor to the manufacturing of universal scalable quantum computers, or not

Learning to measure

Adaptive informationally complete POVMs for near-term quantum algorithms



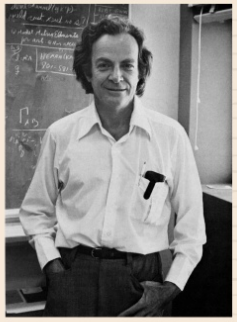
Sabrina Maniscalco, University of Helsinki, Aalto University, Algorithmiq Oy - Finland

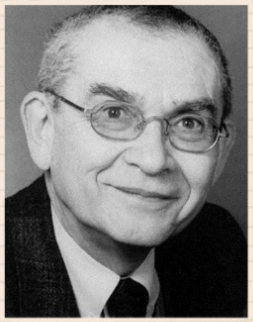
NISQ Devices

What are the **useful problems** which quantum computers can solve **more efficiently** than their classical counterparts?

Which subclass of such problems are **experimentally less demanding**, given the current state-of-the-art quantum hardware?

Quantum Simulations





Feynman ´82 & Manin ’80

Simulating quantum systems is hard on classical computers.
However, one can use quantum systems to simulate other quantum systems efficiently.

Optimization problems

Abandoning the "exact"

Unlikely that quantum computers will be able to solve efficiently worst-case instances of NP-hard problems, like combinatorial optimisation problems.

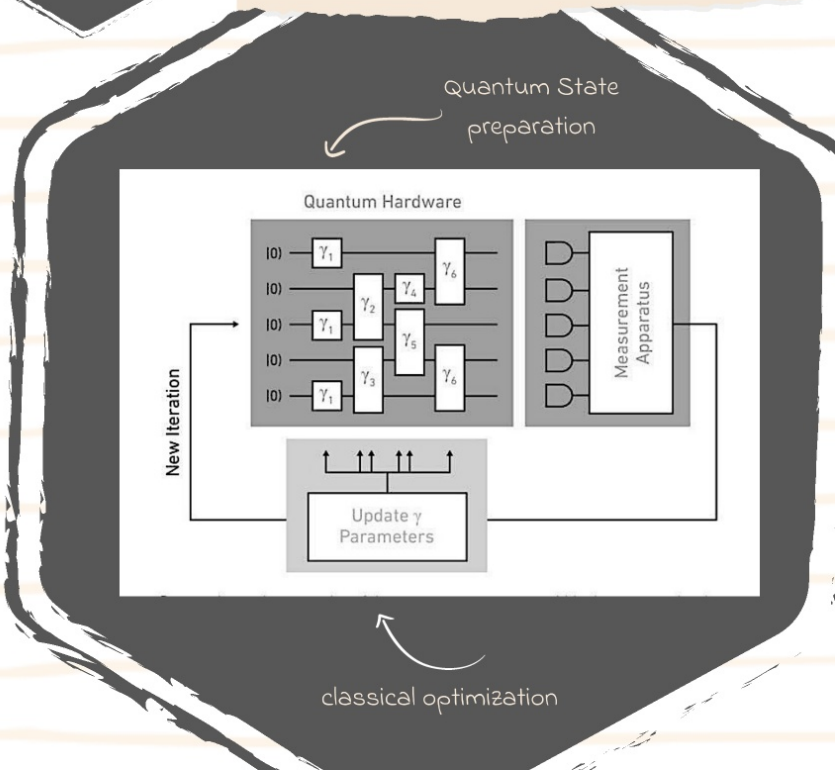
However, it might be possible to build **specific-purpose quantum devices** able to find better **approximate solutions** or **find such approximate solutions faster**

VQE

Impact

Roadblocks

Hybrid quantum-classical algorithm



Variational Quantum Algorithms

M. Cerezo et al, arxiv 2012.09265 (2020)

Variational Quantum Eigensolver

minimizing the energy of a many-body correlated quantum system (chemistry examples)

$$\lambda_{min} \leq \lambda_\theta \equiv \langle \psi(\theta) | H | \psi(\theta) \rangle$$

explore the exponentially-large Hilbert space of the fermionic particles in order to find iteratively the ground state of the Hamiltonian, without solving the full diagonalisation problem.

Quantum Chemistry Hamiltonian

S. McArdle et al, arxiv 1808.10402, (2020)

the knowledge of the ground state of a chemistry compound as a function of the bound length allows one to extract crucial information such as the **equilibrium bond length, bond angle, and dissociation energy**

Electronic Hamiltonian

$$H_e = -\sum_i \frac{\nabla_i^2}{2} - \sum_{i,I} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

Second quantization Hamiltonian

$$H = \sum_{pq} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{pqrs} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s$$

The Hamiltonian contains up to N^4 terms

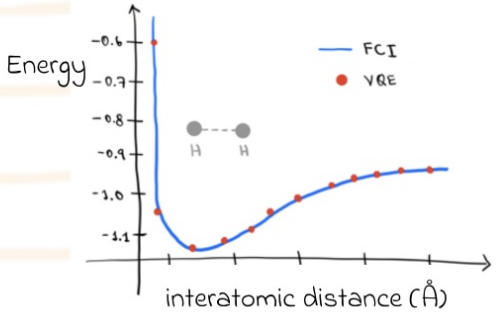
$$h_{pq} = \int \mathrm{d}\mathbf{x} \phi_p^*(\mathbf{x}) \left(-\frac{\nabla^2}{2} - \sum_I \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|} \right) \phi_q(\mathbf{x})$$

kinetic energy terms of the electrons, and their Coulomb interaction with the nuclei

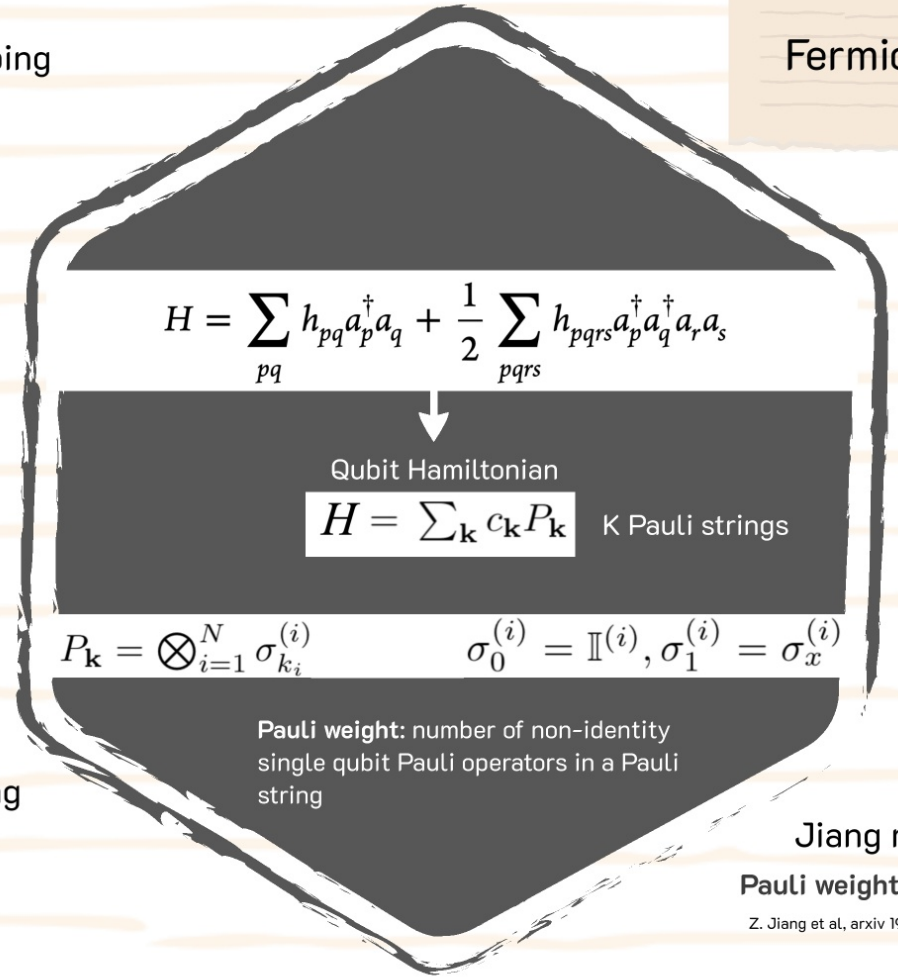
$$h_{pqrs} = \int \mathrm{d}\mathbf{x}_1 \mathrm{d}\mathbf{x}_2 \frac{\phi_p^*(\mathbf{x}_1)\phi_q^*(\mathbf{x}_2)\phi_r(\mathbf{x}_2)\phi_s(\mathbf{x}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|}$$

the electron-electron Coulomb repulsion

excitation of an electron into the single electron spin orbital p



Fermion to qubit mappings



JW

Jordan - Wigner mapping

Pauli weight: O(N)

P

Parity mapping

Pauli weight: O(N)

BK

Bravyi-Kitaev mapping

Pauli weight: O(log(N))

J

Jiang mapping

Pauli weight: O(log(N))

Z. Jiang et al, arxiv 1910.10746 (2020)

Fermion	Jordan-Wigner
$a\|0001\rangle + b\|0010\rangle$	$a\|0001\rangle + b\|0010\rangle$
$+c\|0100\rangle + d\|1000\rangle$	$+c\|0100\rangle + d\|1000\rangle$
a_0	Q_0
a_1	$Q_1 Z_0$
a_2	$Q_2 Z_1 Z_0$
a_3	$Q_3 Z_2 Z_1 Z_0$
$\hat{n}_i = a_i^\dagger a_i$	$\|1\rangle\langle 1\|_i$

Z are parity operators

Jorden-Wigner transformation

stores the occupation number locally and the parity non-locally

Fermion	Jordan-Wigner	Parity
$a\|0001\rangle + b\|0010\rangle$	$a\|0001\rangle + b\|0010\rangle$	$a\|1111\rangle + b\|1110\rangle$
$+c\|0100\rangle + d\|1000\rangle$	$+c\|0100\rangle + d\|1000\rangle$	$+c\|1100\rangle + d\|1000\rangle$
a_0	Q_0	$X_3X_2X_1Q_0$
a_1	Q_1Z_0	$X_3X_2\left(Q_1\|0\rangle\langle 0\|_0 - Q_1^\dagger\|1\rangle\langle 1\|_0\right)$
a_2	$Q_2Z_1Z_0$	$X_3\left(Q_2\|0\rangle\langle 0\|_1 - Q_2^\dagger\|1\rangle\langle 1\|_1\right)$
a_3	$Q_3Z_2Z_1Z_0$	$Q_3\|0\rangle\langle 0\|_2 - Q_3^\dagger\|1\rangle\langle 1\|_2$
$\hat{n}_i = a_i^\dagger a_i$	$\|1\rangle\langle 1\|_i$	$\|1\rangle\langle 1\|_{i=0}$, $\frac{1}{2}\left(\mathbf{1} - Z_iZ_{i-1}\right)_{i=1,2,3}$

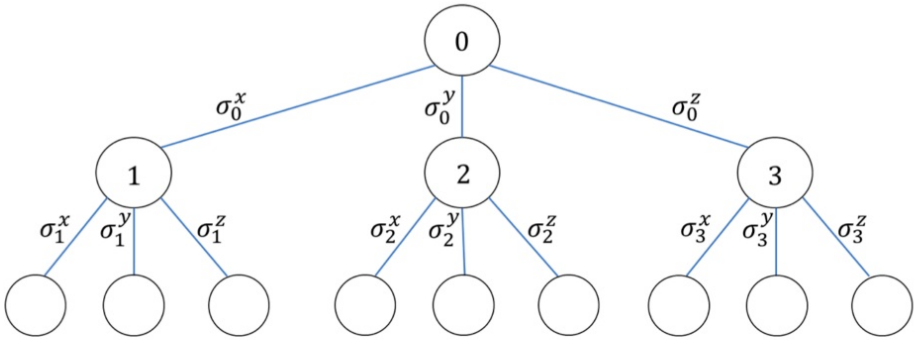
Parity transformation

stores the parity locally, and the occupation number non-locally

Bravyi-Kitaev transformation

Fermion	Jordan-Wigner	Parity	Bravyi-Kitaev
$a\,\|0001\rangle + b\,\|0010\rangle$	$a\,\|0001\rangle + b\,\|0010\rangle$	$a\,\|1111\rangle + b\,\|1110\rangle$	$a\,\|1011\rangle + b\,\|1010\rangle$
$+c\,\|0100\rangle + d\,\|1000\rangle$	$+c\,\|0100\rangle + d\,\|1000\rangle$	$+c\,\|1100\rangle + d\,\|1000\rangle$	$+c\,\|1100\rangle + d\,\|1000\rangle$
a_0	Q_0	$X_3X_2X_1Q_0$	$X_3X_1Q_0$
a_1	Q_1Z_0	$X_3X_2\left(Q_1\,\|0\rangle\langle 0\|_0 - Q_1^\dagger\,\|1\rangle\langle 1\|_0\right)$	$X_3\left(Q_1\,\|0\rangle\langle 0\|_0 - Q_1^\dagger\,\|1\rangle\langle 1\|_0\right)$
a_2	$Q_2Z_1Z_0$	$X_3\left(Q_2\,\|0\rangle\langle 0\|_1 - Q_2^\dagger\,\|1\rangle\langle 1\|_1\right)$	$X_3Q_2Z_1$
a_3	$Q_3Z_2Z_1Z_0$	$Q_3\,\|0\rangle\langle 0\|_2 - Q_3^\dagger\,\|1\rangle\langle 1\|_2$	$\frac{1}{2}\left(Q_3(\mathbf{1}+Z_2Z_1) - Q_3^\dagger(\mathbf{1}-Z_2Z_1)\right)$
$\hat{n}_i = a_i^\dagger a_i$	$\|1\rangle\langle 1\|_i$	$\|1\rangle\langle 1\|_{i=0}\ ,\ \frac{1}{2}\left(\mathbf{1}-Z_iZ_{i-1}\right)_{i=1,2,3}$	$\|1\rangle\langle 1\|_{i=0,2}\ ,\ \frac{1}{2}(\mathbf{1}-Z_1Z_0)_{i=1}\ ,\ \frac{1}{2}(\mathbf{1}-Z_3Z_2Z_1)_{i=3}$

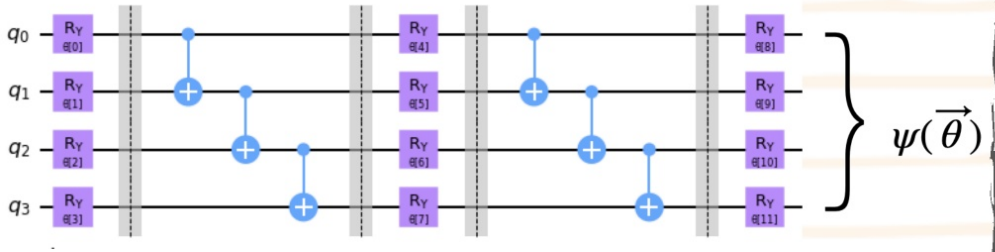
a midway point between the JW and parity methods:
it compromises on the locality of occupation number and parity information



Jiang Mapping

defined on ternary trees, maps any single Majorana operator of an n-mode fermionic system to a multi-qubit Pauli operator acting nontrivially on O(log(N)) qubits

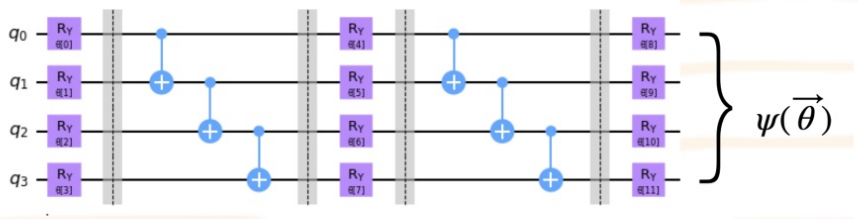
Variational ansatz

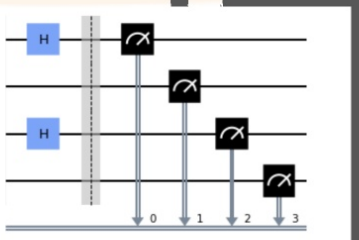


State preparation

a state which has a sufficiently large overlap with the target eigenstate (ground state)

e.g., hardware efficient ansatz and chemically inspired ansatz





$$P_1 = \sigma_x \otimes \sigma_z \otimes \sigma_x \otimes \sigma_z$$

Pauli string measurement

Measurement

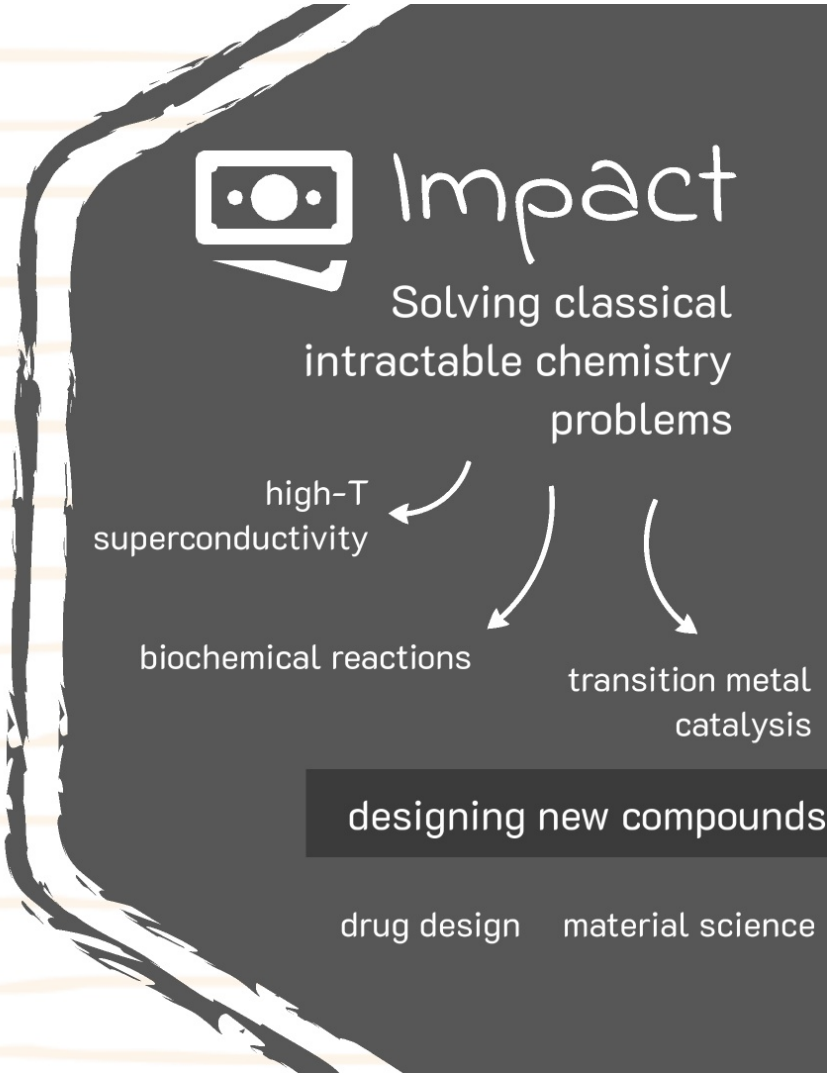
calculate the expectation value

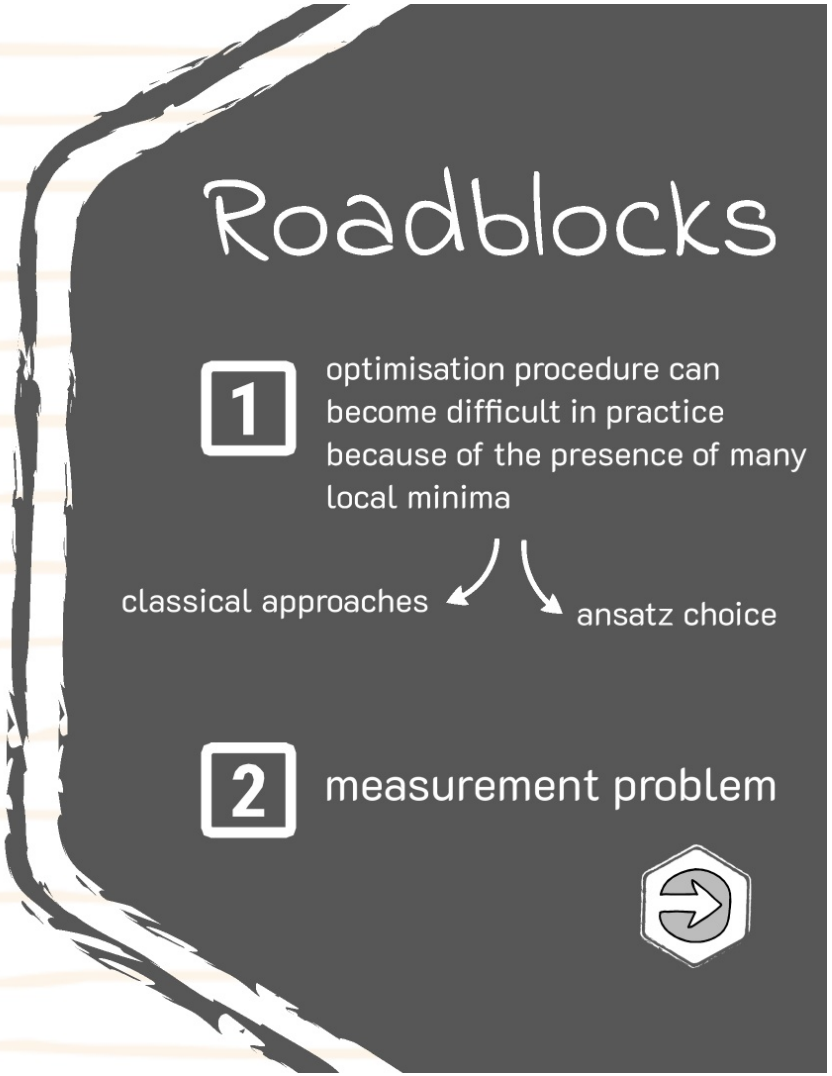
$$\langle H \rangle_{\psi(\theta)} = \sum_k c_k \langle P_k \rangle_{\psi(\theta)}$$

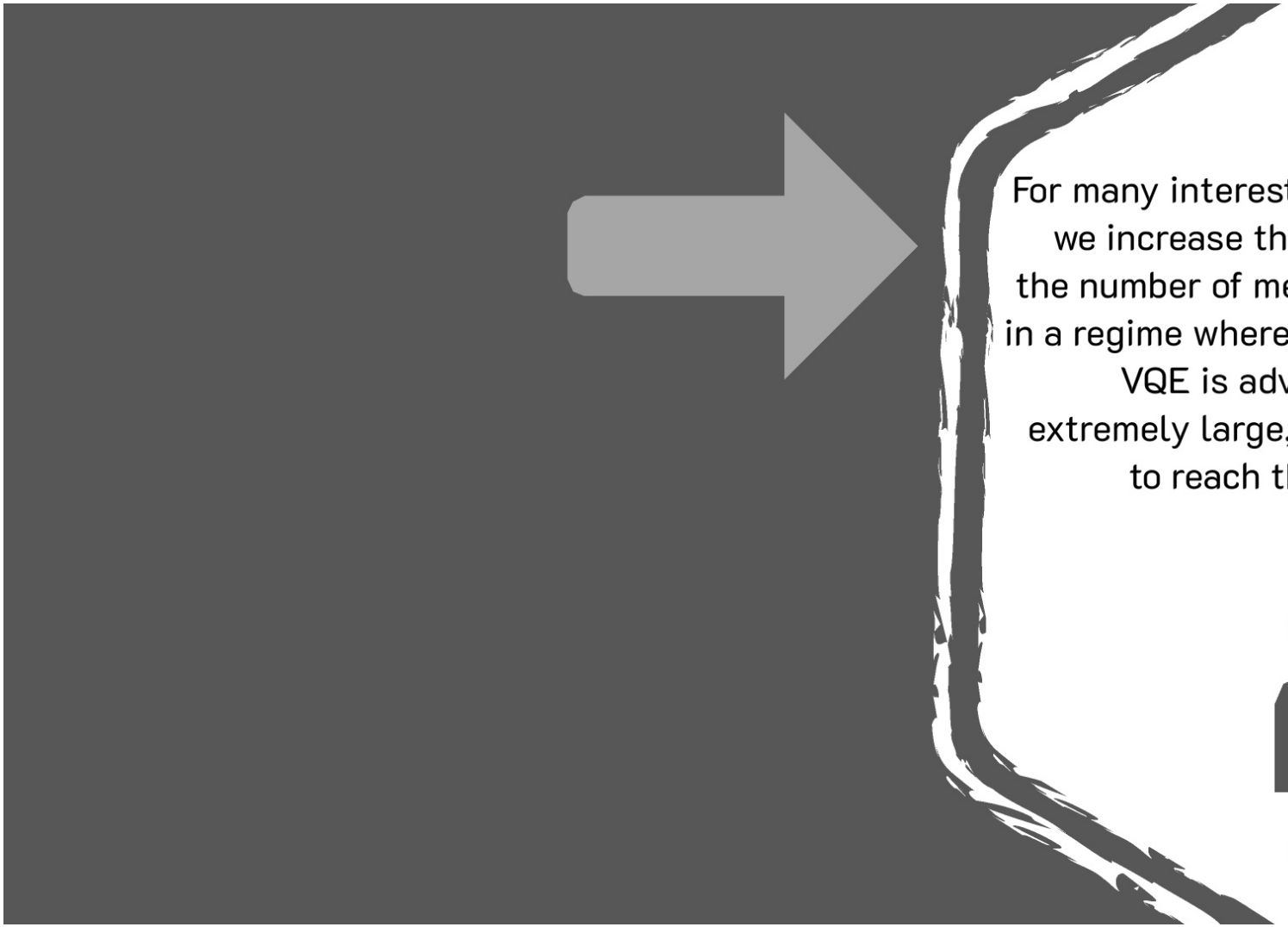
repeat to accumulate statistics

Classical optimization

e.g. gradient descent


Impact
Solving classical intractable chemistry problems
high-T superconductivity
biochemical reactions
transition metal catalysis
designing new compounds
drug design
material science

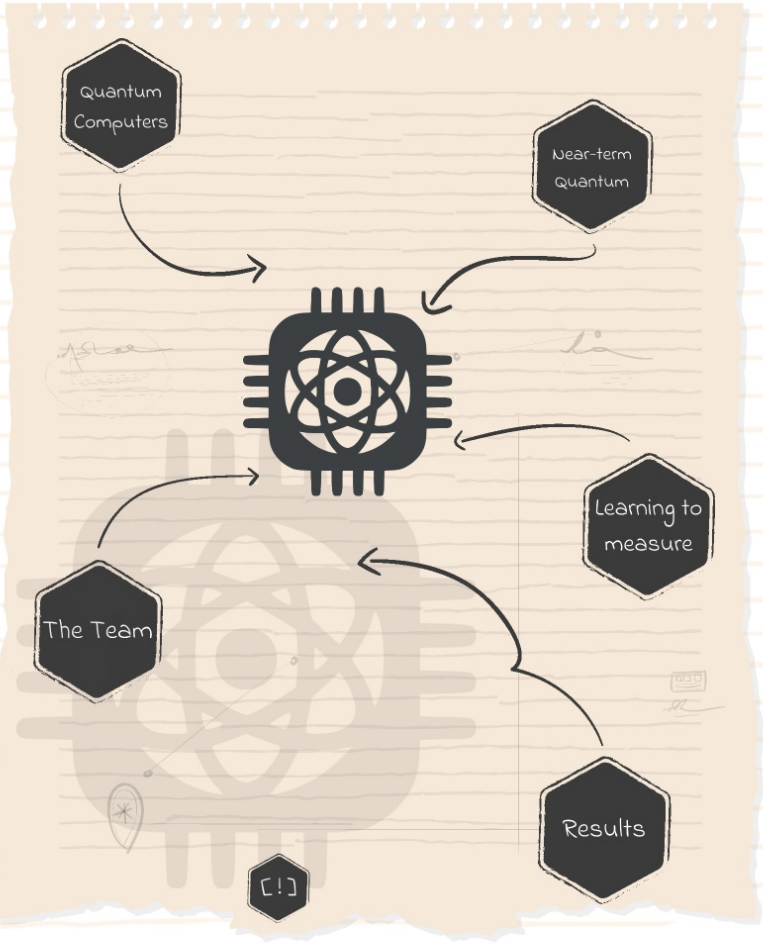

Roadblocks
1
optimisation procedure can become difficult in practice because of the presence of many local minima
classical approaches
ansatz choice
2
measurement problem



For many interesting Hamiltonians, as we increase the size of the problem the number of measurements needed, in a regime where the quantum core of VQE is advantageous, would be extremely large, so it may take years to reach the required precision (chemical accuracy).

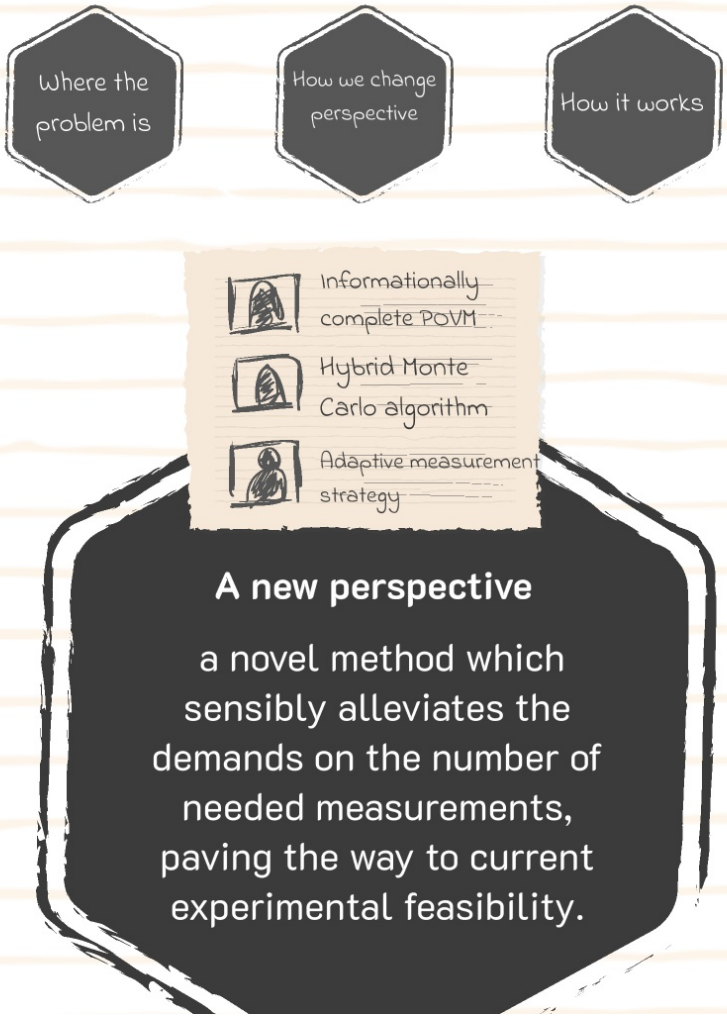
Learning to measure

Adaptive informationally complete POVMs for near-term quantum algorithms



Sabrina Maniscalco, University of Helsinki, Aalto University, Algorithmiq Oy - Finland

Learning to Measure



observable

Pauli strings

$$\mathcal{O} = \sum_{\mathbf{k}} c_{\mathbf{k}} P_{\mathbf{k}}$$

error in the estimation

$$\epsilon = \sqrt{\sum_{\mathbf{k}} |c_{\mathbf{k}}|^2 \mathrm{Var}(P_{\mathbf{k}})/S}$$

number of measurements on each Pauli string

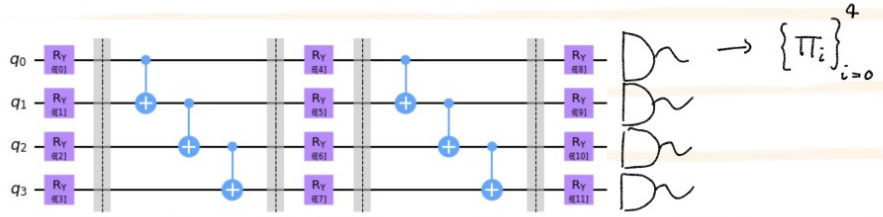
$$\mathrm{Var}(P_{\mathbf{k}}) \simeq \langle P_{\mathbf{k}}^2 \rangle - \langle P_{\mathbf{k}} \rangle^2$$

The measurement problem

sub-optimal measurements scheme, as the variance of O is the sum of the weighted variances of the individual Pauli strings

Grouped Pauli method aims to identify all the Pauli strings that can be measured simultaneously from the same data set

not efficient



$\Pi_m^{(i)}$ $m = 0, 1, 2, 3$ effects

each effect is associated to one of the 4 possible outcomes of 2 qubit measurements $\mathrm{Tr}[\rho\Pi_i]$

quantum-classical hybrid Monte Carlo

POVM optimization

Starting point

informationally complete positive operator-valued measures (IC-POVMs)

can be used to estimate any expectation value of our choice!

quantum-classical hybrid Monte Carlo algorithm

MC sampling

measuring a linear combination of a large number of Pauli strings

$$\mathcal{O} = \sum_{\mathbf{k}} c_{\mathbf{k}} \bigotimes_{i=1}^{N} \sigma_{k_i}^{(i)} = \sum_{\mathbf{k}} c_{\mathbf{k}} \bigotimes_{i=1}^{N} \left(\sum_{m_i} b_{k_i m_i}^{(i)} \Pi_{m_i}^{(i)} \right)$$

$$= \sum_{\mathbf{m}} \left(\sum_{\mathbf{k}} c_{\mathbf{k}} \prod_{i=1}^{N} b_{k_i m_i}^{(i)} \right) \Pi_{\mathbf{m}} = \sum_{\mathbf{m}} w_{\mathbf{m}} \Pi_{\mathbf{m}}.$$

$$\mathbf{m} = (m_1, \ldots, m_N)$$

Monte Carlo integral

$$\langle \mathcal{O} \rangle = \mathrm{Tr}[\rho \mathcal{O}] = \sum_{\mathbf{m}} w_{\mathbf{m}} \mathrm{Tr}[\rho \Pi_{\mathbf{m}}] = \sum_{\mathbf{m}} w_{\mathbf{m}} p_{\mathbf{m}}$$

$$\langle \mathcal{O} \rangle = \langle w_{\mathbf{m}} \rangle_{\{p_{\mathbf{m}}\}}$$

We can estimates the weighted average of all the Pauli strings simultaneously, regardless of whether they commute or not, by exploiting IC data.

1. circumvents costly tomographic reconstruction of quantum states

2. naturally takes into account the covariance between all these parallel measurements

$$\langle \mathcal{O} \rangle = \langle w_{\mathbf{m}} \rangle_{\{p_{\mathbf{m}}\}}$$

repeat the measurement S times using the local POVMs to sample from the probability distribution $\{p_{\mathbf{m}}\}$

estimator

$$\bar{\mathcal{O}} = \frac{1}{S} \sum_{s=1}^{S} w_{\mathbf{m}_s}$$

$$w_{\mathbf{m}_s} = \sum_{\mathbf{k}} c_{\mathbf{k}} \prod_{i=1}^{N} b_{k_i m_i}^{(i)}$$

can be calculated efficiently on a classical computer

estimation error $\sqrt{\mathrm{Var}(w_{\mathbf{m}})/S}$

POVM optimization

via classical gradient estimation

classical postprocessing routine to navigate the space of POVMs towards low-variance ones

one feeds the outcomes of the current POVM to the classical routine, which uses them to evaluate the variance of the new POVMs

local effects $\{\Gamma_r^{(i)}\}$

$$\mathrm{Var}(w_{\mathbf{r}}) = \sum_{\mathbf{r}} w_{\mathbf{r}}^2 \mathrm{Tr}[\rho\Gamma_{\mathbf{r}}] - \underbrace{\left(\sum_{\mathbf{r}} w_{\mathbf{r}} \mathrm{Tr}[\rho\Gamma_{\mathbf{r}}]\right)^2}_{= \langle\mathcal{O}\rangle^2}$$

does not depend on the POVMs

term to minimize

$$\langle w_{\mathbf{r}}^2 \rangle_{\{q_{\mathbf{r}}\}} = \sum_{\mathbf{r}} w_{\mathbf{r}}^2 \mathrm{Tr}\left[\rho \bigotimes_{i=1}^{N} \left(\sum_{m_i} d_{r_i m_i}^{(i)} \Pi_{m_i}^{(i)}\right)\right]$$

$$= \sum_{\mathbf{m}} p_{\mathbf{m}} \sum_{\mathbf{r}} \left(\prod_{i=1}^{N} d_{r_i m_i}^{(i)}\right) w_{\mathbf{r}}^2 .$$

on the fly

calculate on classical computer

gradient descent

on the fly

we do not need to first optimise the POVM until it reaches a small-enough variance to then start estimating the expected value of the observable.

The intermediate POVMs used in the process are also IC, so they can be used for the estimation of the mean value of O as well.

The strategy is to **calculate a weighted average of the estimated means that minimises the resulting variance in the estimation.**

This can be done iteratively as the optimisation algorithm progresses

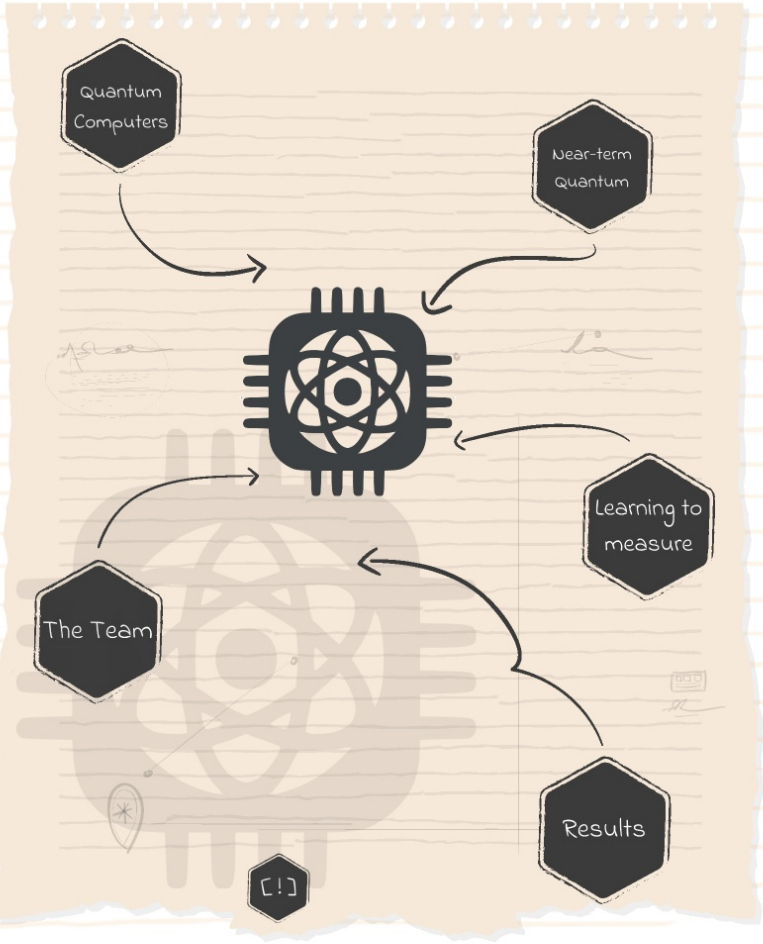
how it works

Key ideas

We use the IC data obtained with the POVM twice

1. to produce an estimation of the mean of the observable.
2. to to find a better POVM in the next experiment

★ the measurement learning procedure improves over the initial POVM with no additional measurement cost

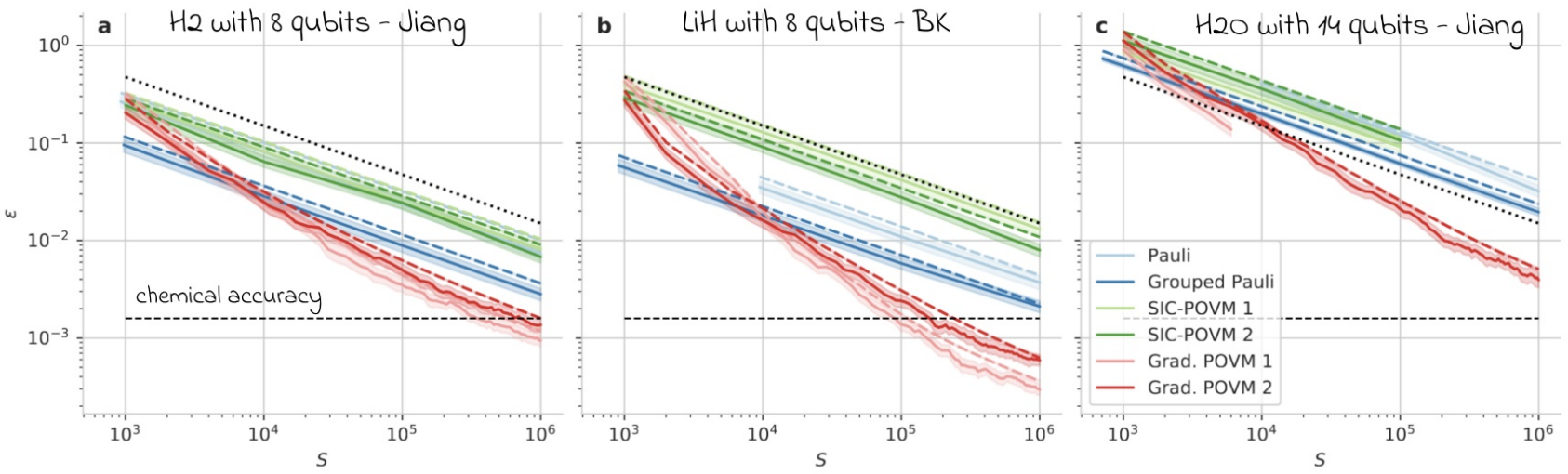


Learning to measure

Adaptive informationally complete POVMs for near-term quantum algorithms

Sabrina Maniscalco, University of Helsinki, Aalto University, Algorithmiq Oy - Finland

estimated (dashed) and actual (solid) error in the estimation of the ground energy



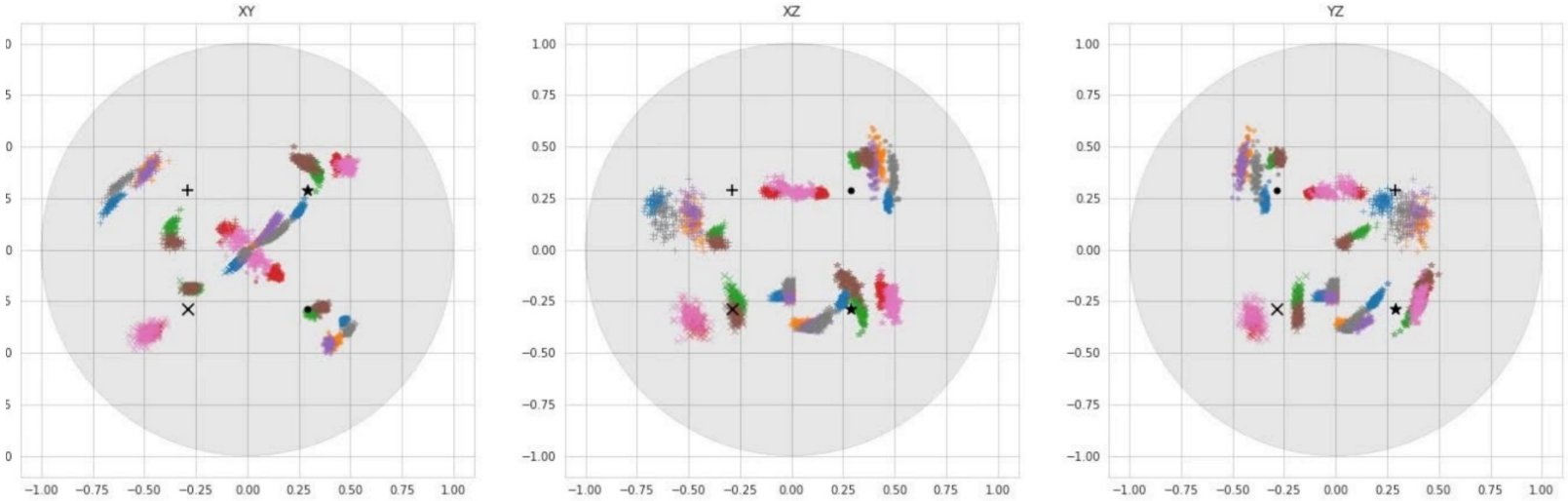
Adaptive measurement for energy estimation

Learning POVM

Scaling

Informational Completeness

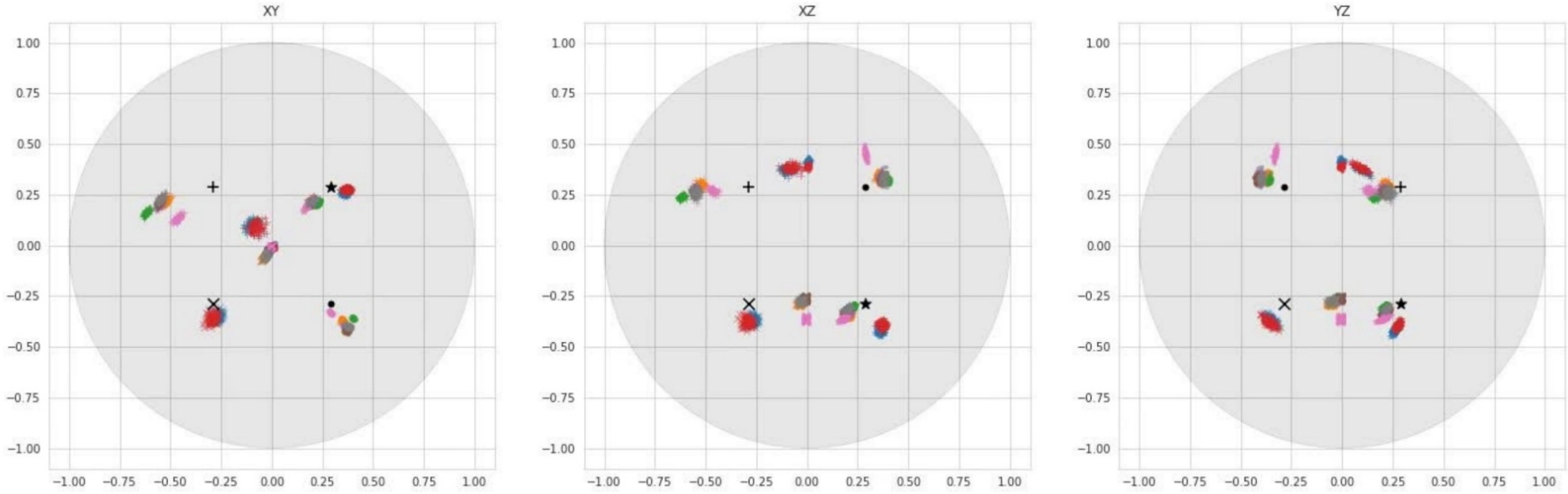
8 qubits H2 Jiang



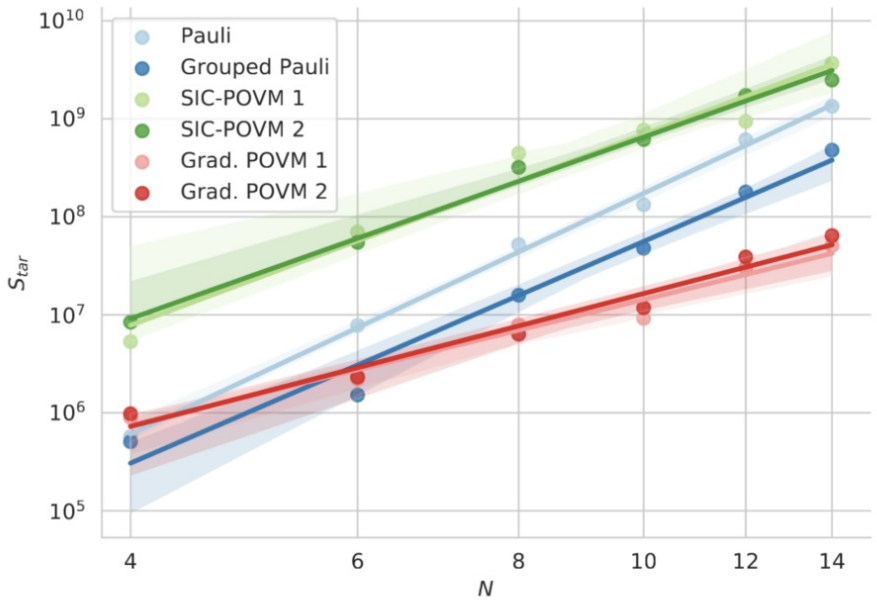
Learning to measure

Example 2

different colors = different qubits
about 100 repetitions


XY
XZ
YZ
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00

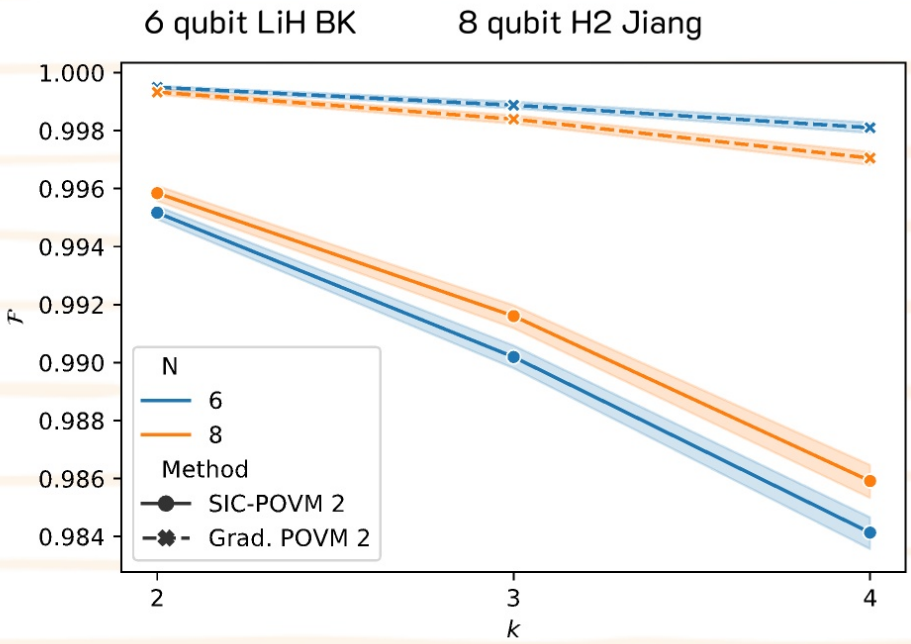
Number of shots S_{tar} required to achieve a target error of 0.5 mHa for H chains as a function of the number of qubits N



Scaling

$$S_{\mathrm{tar}} = aN^b$$

Method	Mapping Parity	BK	JKMN
Pauli	6.0 ± 0.4	6.8 ± 0.5	6.2 ± 0.2
Grouped Pauli	5.5 ± 0.4	6.4 ± 0.5	5.7 ± 0.4
SIC-POVM 1	5.8 ± 0.7	5.7 ± 0.6	4.9 ± 0.5
SIC-POVM 2	5.4 ± 0.5	4.4 ± 0.4	4.7 ± 0.2
Grad. POVM 1	4.3 ± 0.6	4.6 ± 0.5	3.2 ± 0.3
Grad. POVM 2	4.0 ± 0.6	4.4 ± 0.5	3.4 ± 0.3

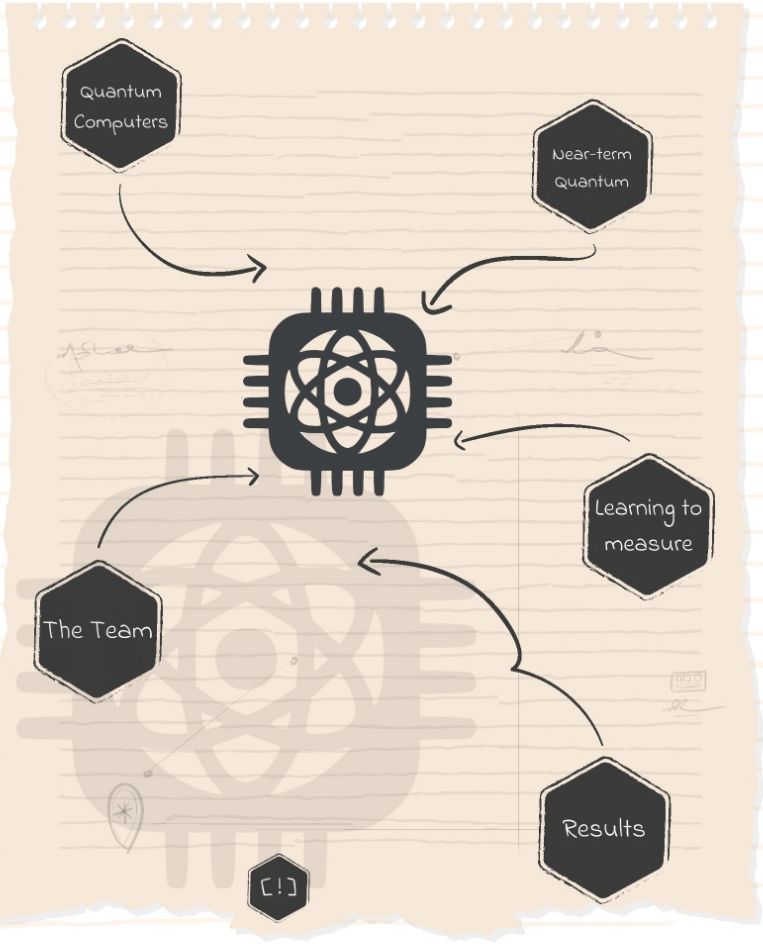


partial tomography

Using the data obtained to calculate the mean energy for partial (2, 3 and 4 qubit) tomography

Learning to measure

Adaptive informationally complete POVMs for near-term quantum algorithms



Sabrina Maniscalco, University of Helsinki, Aalto University, Algorithmiq Oy - Finland

Conclusions

Take home

- New measurement strategy in which the optimal measurement of an operator average is learnt in an adaptive fashion with no measurement overhead
- Does not require exponentially-scaling classical computations
- Our algorithm is completely agnostic to the nature of the qubit Hamiltonian and it is exact.
- The measurement data can be reused to calculate other properties of the state, including its tomographic reconstruction.

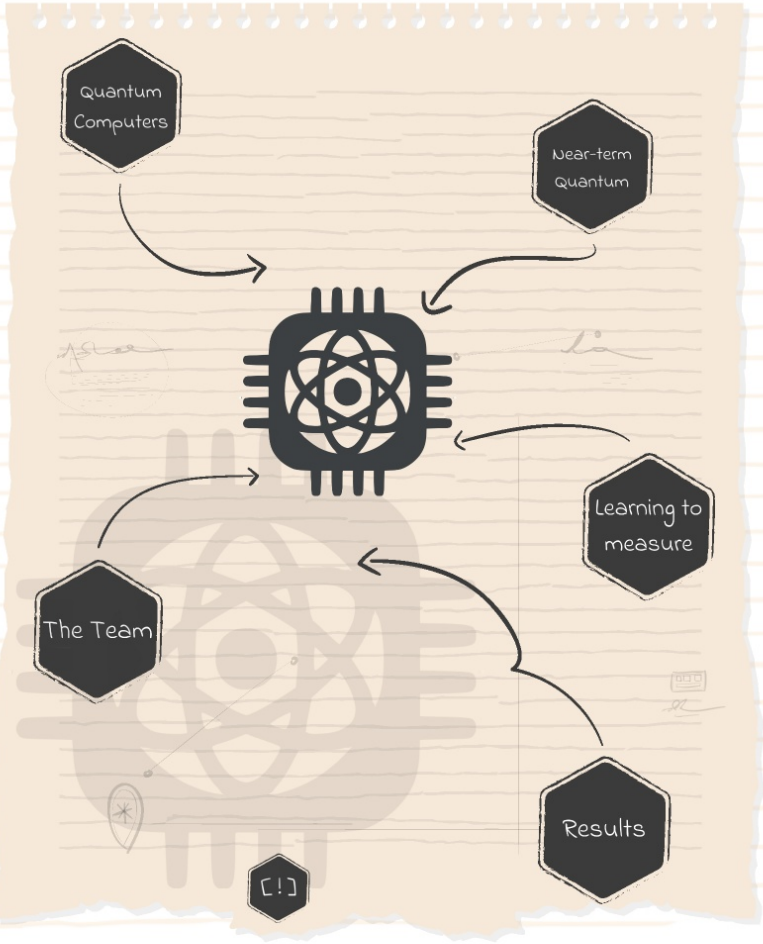
Take home message

The road to useful quantum advantage requires new skillful ways to optimise every aspect of the near-term quantum algorithms

We are one step closer!

Learning to measure

Adaptive informationally complete POVMs for near-term quantum algorithms



Sabrina Maniscalco, University of Helsinki, Aalto University, Algorithmiq Oy - Finland