Comparison of various optimization methods in the quantum QAOA algorithm for the Ising model

Yu. Palii

MLIT, Joint Institute for Nuclear Research

Quantum Approximate Optimization Algorithm (QAOA)

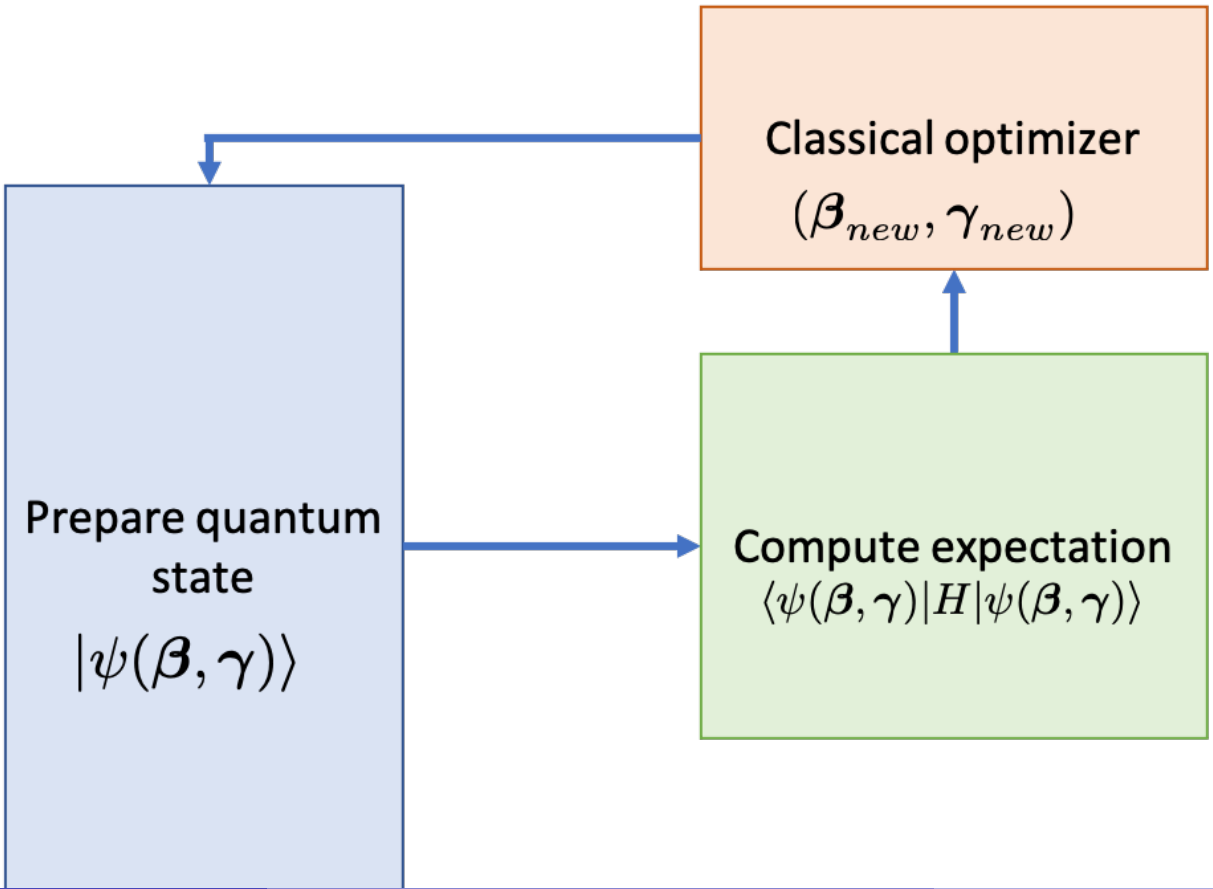
Hybrid quantum-classical algorithm, Edward Farhi and Jeffrey Goldstone. ArXiv:1411.4028

Quantum mechanical variational principle (Rayleigh-Ritz):

for any parametrized trial wave-function (state vector) $|\psi(\alpha)\rangle$, $\alpha = (\alpha_1, \ldots, \alpha_n)^T$

$$\langle\psi(\alpha)|\,\mathcal{H}\,|\psi(\alpha)\rangle \geq E_{Ground}.$$

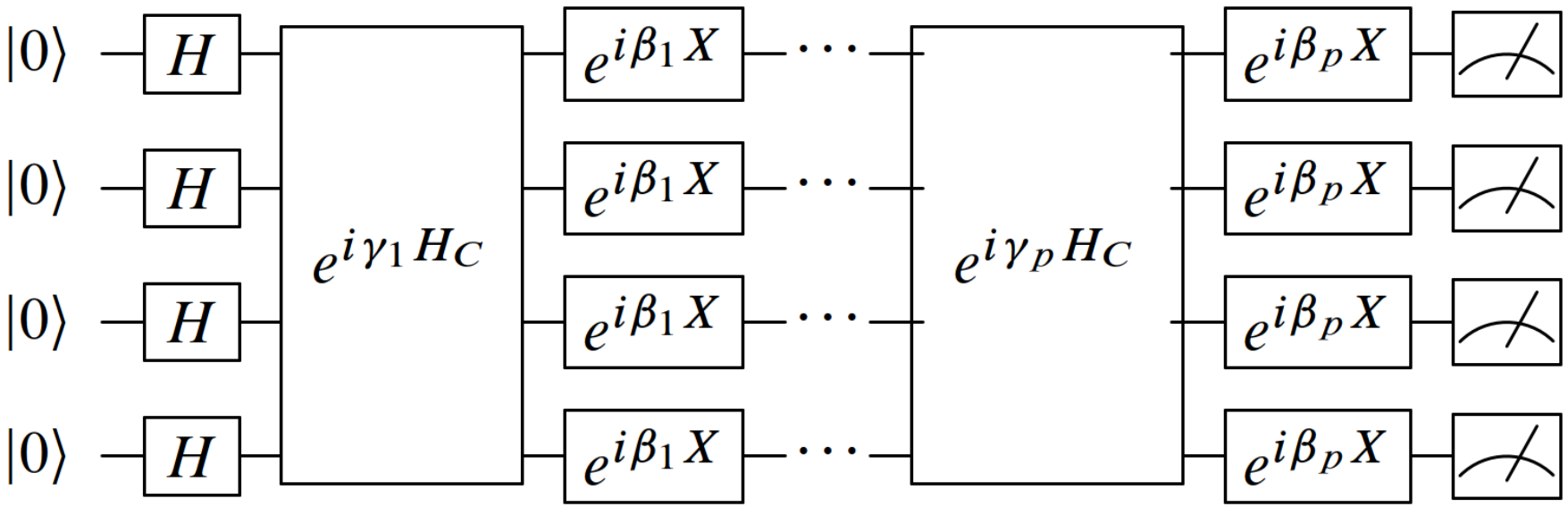
The expectation value of the Hamiltonian plays the role of the cost function.



QAOA Variational Ansatz $|\psi(\gamma,\beta)\rangle$

The driver $U(\gamma_i,\mathcal{H})$ and mixing $U(\beta_i,B), i=1,\ldots,p$, operators are arranged in p layers.

$$|\psi(\gamma,\beta)\rangle = \underbrace{U(\beta_p,B)U(\gamma_p,\mathcal{H})}_{p}\ldots\underbrace{U(\beta_1,B)U(\gamma_1,\mathcal{H})}_{1}H^{\otimes n}|0\rangle^{\otimes n}$$



Main QAOA Theorem

$$\lim_{p\to\infty}\min_{\gamma,\beta} E_p(\gamma,\beta) = E_{Ground}, \qquad E_p(\gamma,\beta) \equiv \langle\psi(\gamma,\beta)|\mathcal{H}|\psi(\gamma,\beta)\rangle$$

Ising model Hamiltonian using Pauli strings

Ising Hamiltonian with nearest-neighbor interaction in an external magnetic field h:

$$\mathcal{H}(Z) = -J\sum_{\langle i,j\rangle} Z^{(i)}Z^{(j)} - h\sum_{i} Z^{(i)}$$

$Z^{(i)} = \mathbb{I}\otimes\cdots\otimes Z\otimes\cdots\otimes\mathbb{I}$ Pauli operator Z on i-th position acts on i-th qubit

Operators for construction of QAOA Variational Ansatz $|\psi(\gamma,\beta)\rangle$

- Driving operator with the Hamiltonian $\mathcal{H}(Z)$ (commutativity of the Hamiltonian terms is used the factorization),

$$\begin{aligned} U(\gamma,\mathcal{H}) =& e^{i\pi\gamma\mathcal{H}(Z)/2} \\ =& \prod_{\langle i,j\rangle} \exp\Big(-i\pi J\gamma Z^{(i)}\cdot Z^{(j)}/2\Big) \text{ (ZZ interaction of the neighbours } i \text{ and } j), \\ & \prod_{i} \exp(-i\pi\gamma h Z^{(i)}/2) \qquad \text{(interaction of } i\text{th site with the magnetic field).} \end{aligned}$$

- Mixing operator with Pauli-X operators,

$$U(\beta,B) = e^{i\pi\beta B(X)/2} = \prod_{j=1}^{n} e^{i\pi\beta X_j/2}, \qquad B(X) = \sum_{j=1}^{n} X_j.$$

Average Energy Value at the state $|\psi(\gamma,\beta)\rangle$.

- Variational QAOA Ansatz on n-qubit register

$$\begin{aligned}|\psi(\gamma,\beta)\rangle =& U(\beta,B)U(\gamma,\mathcal{H})H^{\otimes n}|0\rangle^{\otimes n}\\ =& \alpha_0(\gamma,\beta)|0\ldots00\rangle + \alpha_1(\gamma,\beta)|0\ldots01\rangle + \ldots + \alpha_{2^n-1}(\gamma,\beta)|1\ldots11\rangle\end{aligned}$$

- Energy as expectation value of the Hamiltonian is a function of the variational parameters γ,β,

$$E(\gamma,\beta) = \langle\psi(\gamma,\beta)|\,\mathcal{H}\,|\psi(\gamma,\beta)\rangle,$$

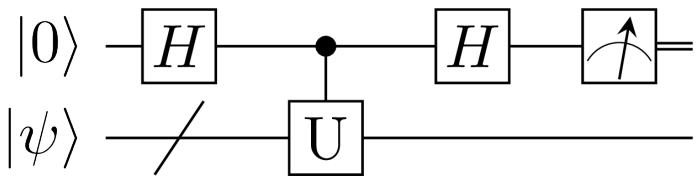
or, if we use the diagonal form of the Hamiltonian,

$$E(\gamma,\beta) = \sum_{i=0}^{2^n-1} |\alpha_i(\gamma,\beta)|^2 E_i.$$

- Sampling (measurement of qubits states) $\mathcal{P}_i(\gamma,\beta) = |\alpha_i(\gamma,\beta)|^2 \approx n_{ith\ state}/n_{meas}$

$$E_p(\gamma,\beta) \approx \sum_{i=0}^{2^n-1} \mathcal{P}_i(\gamma,\beta)E_i, \qquad \sum_{i=0}^{2^n-1} \mathcal{P}_i(\gamma,\beta) = 1$$

Evaluation of an average $\langle\psi(\boldsymbol{\alpha})|U|\psi(\boldsymbol{\alpha})\rangle$ with the Hadamard Test



$$H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad |0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$|0\rangle \otimes |\psi(\boldsymbol{\alpha})\rangle \xrightarrow{\textit{Hadamard Test}} \frac{1}{2}\Big((|0\rangle \otimes (\mathbb{I} + U)|\psi(\boldsymbol{\alpha})\rangle + |1\rangle \otimes (\mathbb{I} - U)\psi(\boldsymbol{\alpha})\rangle\Big)$$

Measurements of the ancilla qubit $|0\rangle$ allow to evaluate the average $\langle\psi|U|\psi\rangle$.

$$P(1) = \frac{1}{2}\langle\psi(\boldsymbol{\alpha})|(\mathbb{I} - Re\, U)|\psi(\boldsymbol{\alpha})\rangle, \qquad P(0) = 1 - P(1) = \frac{1}{2}\langle\psi(\boldsymbol{\alpha})|(\mathbb{I} + Re\, U)|\psi(\boldsymbol{\alpha})\rangle$$

$$Re\,\langle\psi(\boldsymbol{\alpha})|U|\psi(\boldsymbol{\alpha})\rangle = 1 - 2P(1), \qquad 2Re\, U = U + U^{\dagger}$$

Evaluation an average $\langle\psi(\boldsymbol{\alpha})|U|\psi(\boldsymbol{\alpha})\rangle$ needs measuring only the ancilla qubit!

Hard and Soft ware

PC

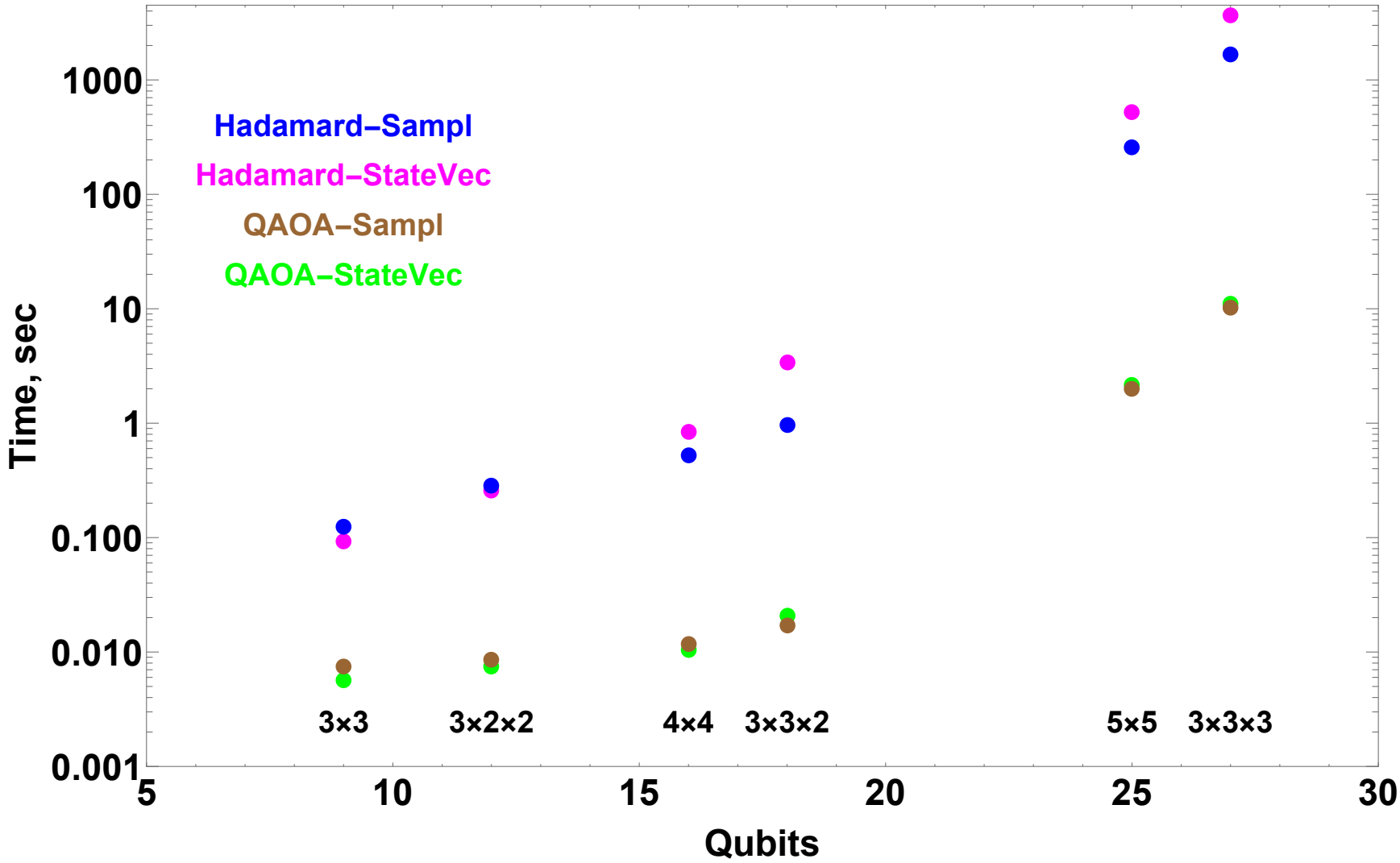
Processor	AMD Ryzen 7 3700X, Socket AM4, 8 x 3.6 GHz, Kernel AMD Matisse.
Motherboard	GIGABYTE B450M DS3H, SSD M.2 Samsung 970 PRO, 512 GB.
Graphics card	NVIDIA® GeForce RTX™ 4070 Ti SUPER 16G VENTUS 2X OC, CUDA CORES® 8448 Units.
RAM	Kingston FURY Beast Black, DDR4, DIMM, total 128 GB, 3200 Mhz.

Ubuntu 24.04, CUDA 12.4,Python, Jupyter Notebook

Simulation of Quantum computing:

Cirq	Framework
Pasqal	
qsim	full wave function simulator written in C++
NVIDIA cuQuantum	SDK of optimized libraries and tools for accelerating quantum computing workflows.

Time for a single energy calculation. CPU



Peak Memory and increment of Memory during Energy evaluation. CPU. MiB

		3D 3×3×2 18 qb	2D 5×5 25 qb	3D 3×3×3 27 qb
State Vector	QAOA	346.45 / 0.04	2637.74 / 512.25	9551.58 / 2048.07
	Hadam	341.46 / 3.90	1867.91 / 1536.41	6477.36 / 6144.55
Sampling	QAOA	348.70 / 0.00	2381.60 / 255.88	8528.04 / 1024.32
	Hadam	341.59 / 00	843.89 / 512.34	2381.30 / 2048.51

Various measurement schemes for Optimization

QAOA+State Vector exact, short time, a large amount of RAM.

QAOA+Sampling less exact, the shortest time, a less amount RAM.

Hadamard+State Vector exact, very long time, a small amount of RAM.

Hadamard+Sampling less exact, long time, the smallest amount RAM.

Gradient-Free Optimization

SciPy Optimize

Local (multivariate) optimization Six variational parameters (three-layer ansatz). Direct search algorithms (gradient-free optimization). Gradient-based failed.

1. COBYLA, the smallest number of energy calculations, high accuracy calculations
2. Powell, larger number of energy calculations, high accuracy
3. Nelder-Mead the largest number of energy calculations, low accuracy

Global optimization with COBYLA and Powell as local minimizers
(Typically, global minimizers efficiently search the parameter space, while using a local minimizer under the hood.)

1. Dual Annealing, smaller number of calculations.
2. SHGO (simplicial homology global optimization) with *sampling_method='sobol'*, outputs a set of minimums.