Introduction to variational quantum algorithms

ACDL 2021 – Tuscany, Italy

Jacob Biamonte and Ernesto Campos

Skolkovo Institute of Science and Technology

Moscow Russia

July, 2021

J.Biamonte@Skoltech.ru

Ernesto.campos@Skoltech.ru

http://Quantum.Skoltech.ru

Contents

Motivating the subject

1. Some physical systems can be described by a Hamiltonian
2. Hamiltonian ground states offer a computational resource. (Ground state quantum computer)
3. Hamiltonian's provide penalty functions for quantum computers to minimize. (Variational training quantum circuits)
4. Effective Hamiltonian's can be realized through repeated measurements of local operators. (Variational model of quantum computation)

Definitions

- Complex Euclidean space $\mathcal{H}$: $[\mathbb{C}^2]^{\otimes n}$
- Linear maps:
  - States: $\psi \in \mathcal{H}$
  - Effects: $\psi^\dagger \in \mathcal{H}^\star$
  - Hamiltonians A in $\mathrm{herm}_{\mathbb{C}}(2^n) \equiv \{A \in \mathcal{L}(\mathcal{H}) \mid A = A^\dagger\}$
  - Propagators U in $\mathbf{U}_{\mathbb{C}}(2^n) \equiv \{U \in \mathcal{L}(\mathcal{H}) \mid U^\dagger U = \mathbb{1}\}$
- Inner product:

$$(\phi, \psi) \to \langle \phi | \, \psi \rangle = \sum_i \bar{\phi}_i \psi^i \in \mathbb{C}$$

  - States/effects are unit ℓ_2 vectors
- Expected value:

$$(A, \psi, \psi^\dagger) \to \langle \psi | \, A \, | \psi \rangle = \sum_{l,m} A_{l,m} \bar{\psi}_m \psi_l \in \mathbb{C}$$

  for $A \in \mathrm{herm}_{\mathbb{C}}(2^n)$

Qubits

Definition (Computational basis)

Single qubit basis: $|0\rangle = [1 \;\; 0]^T, |1\rangle = [0 \;\; 1]^T$

n-qubit basis: $\{|0\rangle, |1\rangle\}^{\otimes n}$

Properties: 2^n *vectors,* $\langle l|m\rangle = \delta_{lm}$

Single qubit state: $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ *;* $\alpha, \beta \in \mathbb{C}$ *and* $|\alpha|^2 + |\beta|^2 = 1$

Measurement: $M_j(|\psi\rangle) = |\langle\psi|j\rangle|^2$ *for* $j \in \{0, 1\}$

Operators and propagators

Definition (Pauli/Sigma operators)

$$\sigma^x = |0\rangle\langle 1| + |1\rangle\langle 0| = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

$$\sigma^y = -\imath|0\rangle\langle 1| + \imath|1\rangle\langle 0| = \begin{bmatrix} 0 & -\imath \\ \imath & 0 \end{bmatrix},$$

$$\sigma^z = |0\rangle\langle 0| - |1\rangle\langle 1| = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

$$\sigma^0 = |0\rangle\langle 0| + |1\rangle\langle 1| = \mathbb{1}$$

A general single qubit rotation by angle θ about an axis defined by unit vector $\boldsymbol{a}$ is $R_{\boldsymbol{a}}(\theta) = e^{-\imath\theta(\boldsymbol{a}.\boldsymbol{\sigma})} = \mathbb{1}.\cos\theta - \imath(\boldsymbol{a}\cdot\boldsymbol{\sigma})\sin\theta$, where $\boldsymbol{a}.\boldsymbol{\sigma} \equiv n_x\sigma^x + n_y\sigma^y + n_z\sigma^z$

Traditional quantum algorithm: Grover's search algorithm

It is a quantum algorithm for unstructured search with time complexity $O(\sqrt{N})$, as compared to the classical complexity $O(N)$. It consists of alternating application of two operators

- $U_s = 2|s\rangle\langle s| - \mathbb{1}$, where $|s\rangle = \frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} |i\rangle$
- $U_\omega = \mathbb{1} - 2|\omega\rangle\langle\omega|$, where $|\omega\rangle$ is the target state in the computational basis

Applying the operators $\sim \frac{\pi}{4}\sqrt{2^n}$ times will maximize the probability of finding the state $|\omega\rangle$

Variational vs Traditional Quantum Algorithms

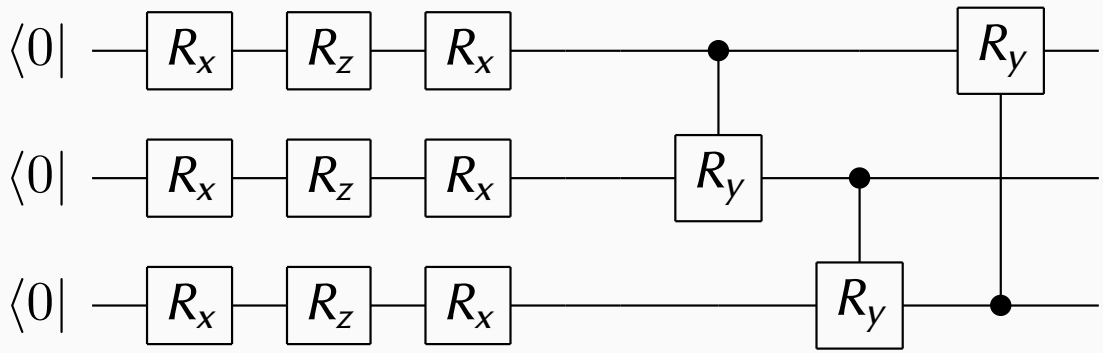
Variational	Traditional
1. Improved robustness to systematic errors 2. Tightly connects hardware with software to overcome hardware constraints 3. Optimizes short depth circuits for optimal use 4. Emulates Hamiltonians by local measurements 5. Outer loop optimization can require significant classical computing resources 6. Coherence time and error rates limit circuit depth	1. Intuitive and familiar, textbook quantum algorithms adhere to the circuit model 2. Theoretical analysis, including complexity, has largely been proven possible 3. Impossible to execute all but the shortest circuits (smallest examples) with current hardware 4. Ignores hardware constraints and susceptible to both systematic and random errors

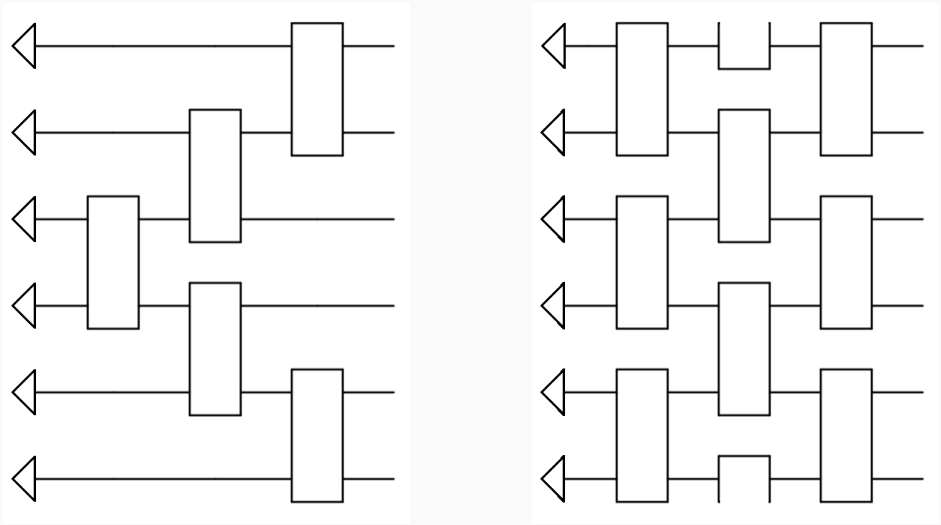
Variational algorithms: algorithms for NISQ devices

1. Errors limit quantum circuit depth
2. State-of-the-art quantum platforms can reach physical error rates near 10^{-3}
3. Noisy intermediate scale quantum (NISQ) era variational algorithms consider a fixed error tolerance and tune a short quantum circuit to minimize an objective function
4. Typically minimising the objective function in this setting is equivalent to finding the ground state (lowest eigenvalue) of a Hamiltonian (hermitian matrix)
5. Circuits with dozens of gates can now be realized with negligible accumulated total error

Short Quantum Circuits (to Minimize Hamiltonians) a.k.a. Quantum circuit as a machine learning model

Ansatz states[1]



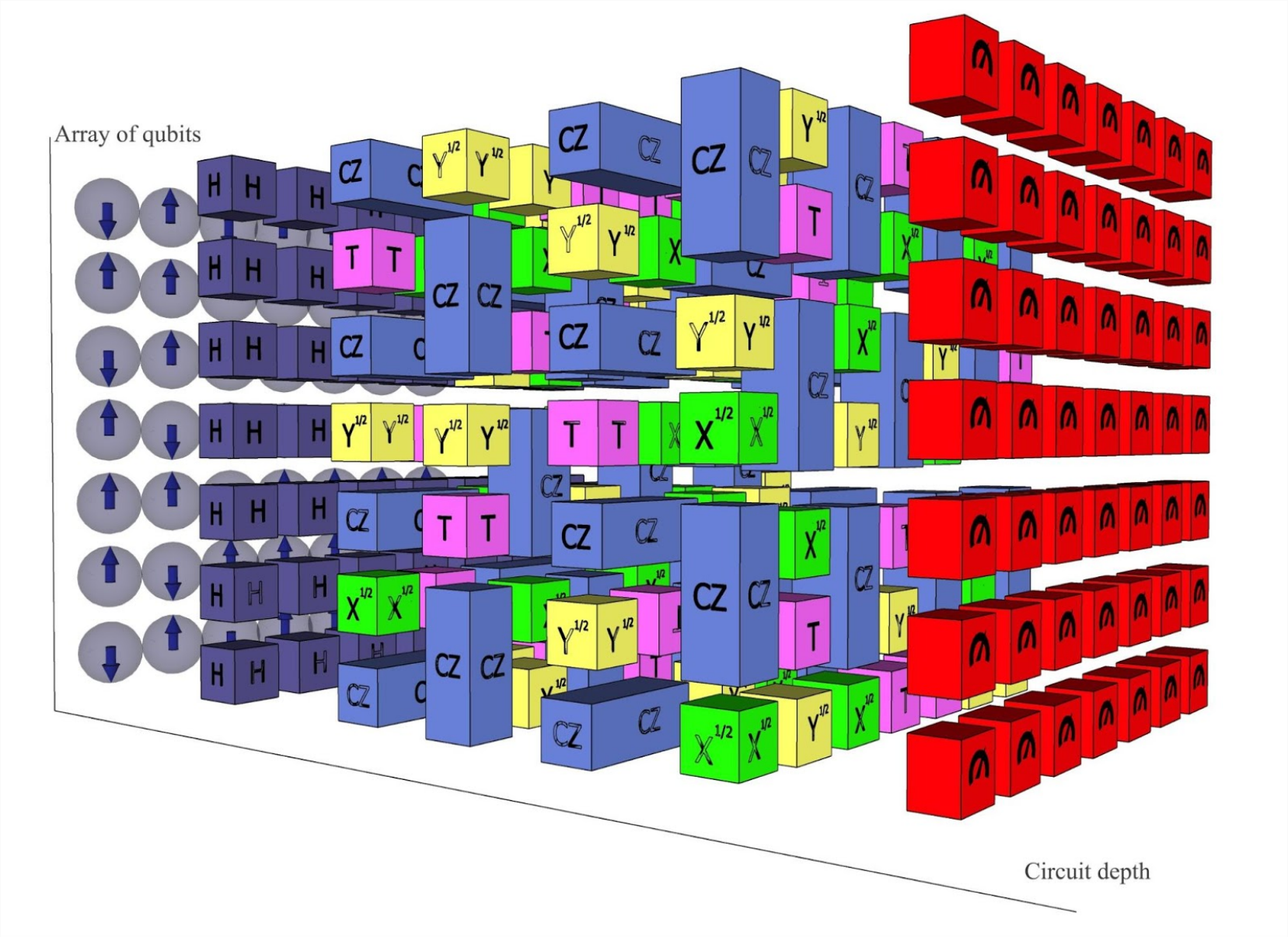


Top: Single layer of the hardware efficient ansatz; Bottom left: Tree tensor network; Bottom right: Checkerboard or brick-layer ansatz.

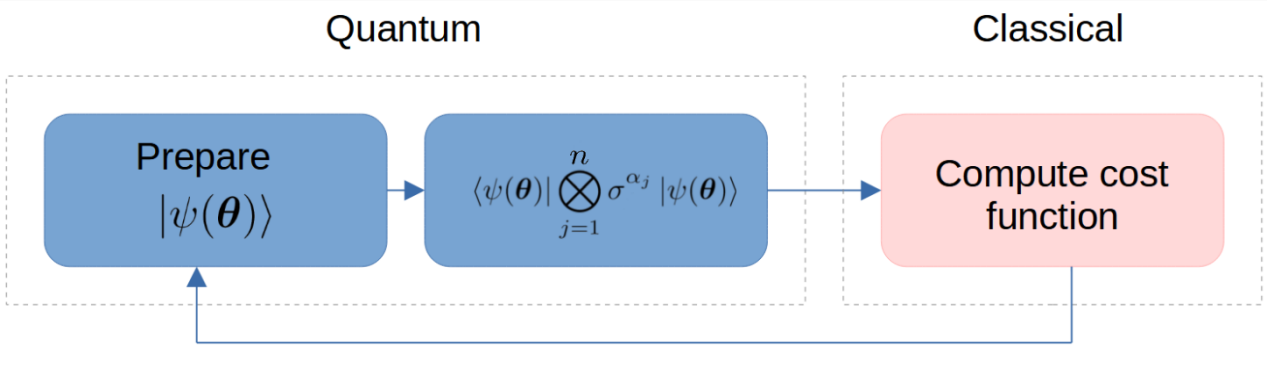
[1] **Quantum Machine Learning Tensor Network States**
Andrey Kardashin, Alexey Uvarov and Jacob Biamonte
Frontiers in Physics 8, 586374 (2021)

Space-time quantum circuit

figure from Google AI



NISQ strategy



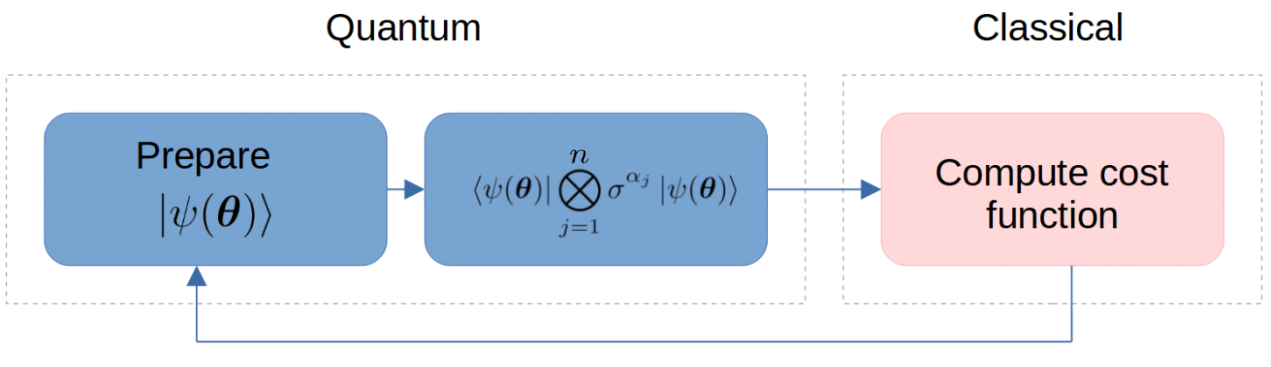
$|\psi(\boldsymbol{\theta})\rangle$ represents variational states parameterized by a vector of real parameters $\boldsymbol{\theta}$

Expected value of a Pauli-string, $\bigotimes_{j=1}^{n}\sigma^{\alpha_j}$ and $\alpha_j \in \{0,1,2,3\}$, is computed for $|\psi(\boldsymbol{\theta})\rangle$

Classical optimization routines iteratively updates parameters $\boldsymbol{\theta}$ to minimize an arbitrary Pauli-string

NISQ strategy

NISQ era quantum computation is focused on quantum circuits that are short enough to be executed without quantum error correction (length is limited by gate fidelity).



By adjusting parameters in an otherwise fixed quantum circuit, low-depth noisy quantum circuits are pushed to their ultimate use case.[2][3]

[2] **Variational Quantum Eigensolver for Frustrated Quantum Systems**
Alexey Uvarov, Jacob Biamonte, Dmitry Yudin
Physical Review B 102, 075104 (2020)

[3] **Certified variational quantum algorithms for eigenstate preparation**
Andrey Kardashin, Alexey Uvarov, Dmitry Yudin, Jacob Biamonte
Physical Review A 102, 052610 (2020)

Ansatz states

Definition (Variational statespace)
The variational statespace of a l-parameterized n-qubit state preparation process is the union of $|\psi(\boldsymbol{\theta})\rangle$ over all possible assignments of real numbers $\boldsymbol{\theta}$.

1. Example: a quantum circuit (ansatz) with $\boldsymbol{\theta} \in (0, 2\pi]^{\times l}$ tunable parameters as
$$|\psi(\boldsymbol{\theta})\rangle = \Pi_l U_l |0\rangle^{\otimes n} \tag{1}$$
where U_l is adjusted by θ_l
2. Example: time-dependent and appropriately bounded parameters ($\theta_k(t)$ corresponding to Hermitian $A^{(k)}$) as
$$|\psi\rangle = \mathcal{T}\{e^{-\imath \sum \theta_k(t) A^{(k)}}\} |0\rangle^{\otimes n} \tag{2}$$
where $\mathcal{T}$ time orders the sequence and superscript k indexes the kth operator $A^{(k)}$

Variational principles

Definition
A variational principle is a problem specific reduction to that of finding extrema of an objective function.

Variational quantum computation most often reduces problems to the normalized minimization,

$$\min_{\theta} \langle \mathcal{H} \rangle \geq \min \langle \mathcal{H} \rangle . \tag{3}$$

Alternative NISQ approaches[4] might minimise the variance

$$\min(\langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2) \geq 0 \tag{4}$$

which vanishes if and only if $|\psi\rangle$ is an eigenstate of $\mathcal{H}$.

[4] **Certified variational quantum algorithms for eigenstate preparation**
Andrey Kardashin, Alexey Uvarov, Dmitry Yudin, Jacob Biamonte
Physical Review A 102, 052610 (2020)

Implementation of $\mathcal{H}$ using expected values

The expected value of a sum is a sum of expected values.

Consider $\mathcal{H}$ expressed in the sigma basis and evaluate

$$\langle\psi|\,\mathcal{H}\,|\psi\rangle = \langle\psi|\sum_j h_j B_j\,|\psi\rangle = \sum_j h_j\,\langle\psi|\,B_j\,|\psi\rangle \tag{5}$$

where h_j is a real number and B_j is a Pauli string (a tensor product of Pauli matrices).

The right hand side can be evaluated term wise, preparing $|\psi\rangle$ and measuring B_j repeatedly gives an estimate for each $\langle\psi|\,B_j\,|\psi\rangle$ separately.

The convolutional step takes these expected values, scales each by the corresponding h_j and produces an estimate for $\sum_j h_j\,\langle\psi|\,B_j\,|\psi\rangle$.[5]

[5] While the objective function can be evaluated term-wise, achieving tolerance $\sim \epsilon$ requires $\sim \epsilon^{-2}$ measurements—see Hoeffding's inequality.

Variational quantum algorithm experiments

Experiment	Organization	Qubits	Ansatz	Depth	year
QAOA	Google	23	Split operator	4	2021
High energy model	MSU/SkT	2	Checkerboard	3	2021
Supremacy	Google	53	HEA	20	2019
Lattice model	Innsbruck	20	Split operator	6	2019
Chemistry	IBM	6	HEA	1	2017

Theoretical Milestones

Milestone	**organization**	**Year**
Barren plateaus induce limitations on training convergence	Google	2018
Depth 1 QAOA solved exactly	Microsoft	2019
Arbitrary depth QAOA proven to be universal	MIT	2019
Reachaility deficits discovered as an underparametrization induced by constraint to valuable ratio	Skoltech	2020
Barren plateaus discovered as limiting freature of moderate depth circuits	Los Alamos	2020
Piecewise trainability conjecture proven falce for specific cases (a.k.a. abrupt transitions)	Skoltech	2021
QAOA parameters for fixed depth shown to concentrate independent of the problem size, leading to a reduction on traiin time	Numerics: community Analytics: Skoltech	2019-21
Variational quantum computation proven universal	Skoltech	2021

End of lesson 1

Quantum Approximate Optimization Algorithm

1. Given $\mathcal{V} \geq 0$ an optimization problem instance.
   We define the p-depth QAOA **variational state space** as

$$\Omega = \bigcup_{\boldsymbol{\gamma},\boldsymbol{\beta}} \{|\psi(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle\} \subseteq \mathbb{C}_2^{\otimes n} \tag{6}$$

   where

$$|\psi(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle = \prod_{k=1}^{p} \exp\left(-i\beta_k \mathcal{H}_x\right) \cdot \exp\left(-i\gamma_k \mathcal{V}\right) |+\rangle^{\otimes n} \tag{7}$$

   and $\mathcal{H}_x = \sum_i \sigma_x^{(i)}$ is the standard one-body mixer Hamiltonian.

Quantum Approximate Optimization Algorithm

2. An approximate optimization terminates as

$$\min_{\psi \in \Omega} \langle \psi | \mathcal{V} | \psi \rangle \geq \min_{\phi \in \mathbb{C}_2^{\otimes n}} \langle \phi | \mathcal{V} | \phi \rangle . \tag{8}$$

3. Classical outer loop optimization iteratively finds parameters $\boldsymbol{\gamma}^*$ and $\boldsymbol{\beta}^*$ such that

$$\langle \psi(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) | \mathcal{V} | \psi(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \rangle \approx \min(\mathcal{V}).$$

Variational Grover's Search

1. Variational Grover search can be thought of as QAOA with the following setting,

$$\mathcal{V} = |\omega\rangle\langle\omega|$$

and

$$\mathcal{H}_x = (|+\rangle\langle+|)^{\otimes n},$$

where $|\omega\rangle \in \mathbb{C}_2^{\otimes n}$ is the objective state we are searching for.

2. The p-depth variational ansatz then take the form,

$$|\psi_p(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle = \mathrm{A}_p \frac{1}{\sqrt{\mathrm{N}-1}} \sum_{x \neq \omega} |x\rangle + \mathrm{B}_p |\omega\rangle ,$$

3. Amplitudes A_p and B_p can be calculated analytically and therefore,

$$E_g^{\mathrm{QAOA}} = 1 - |B_p|^2.$$

Variational Grover's Search

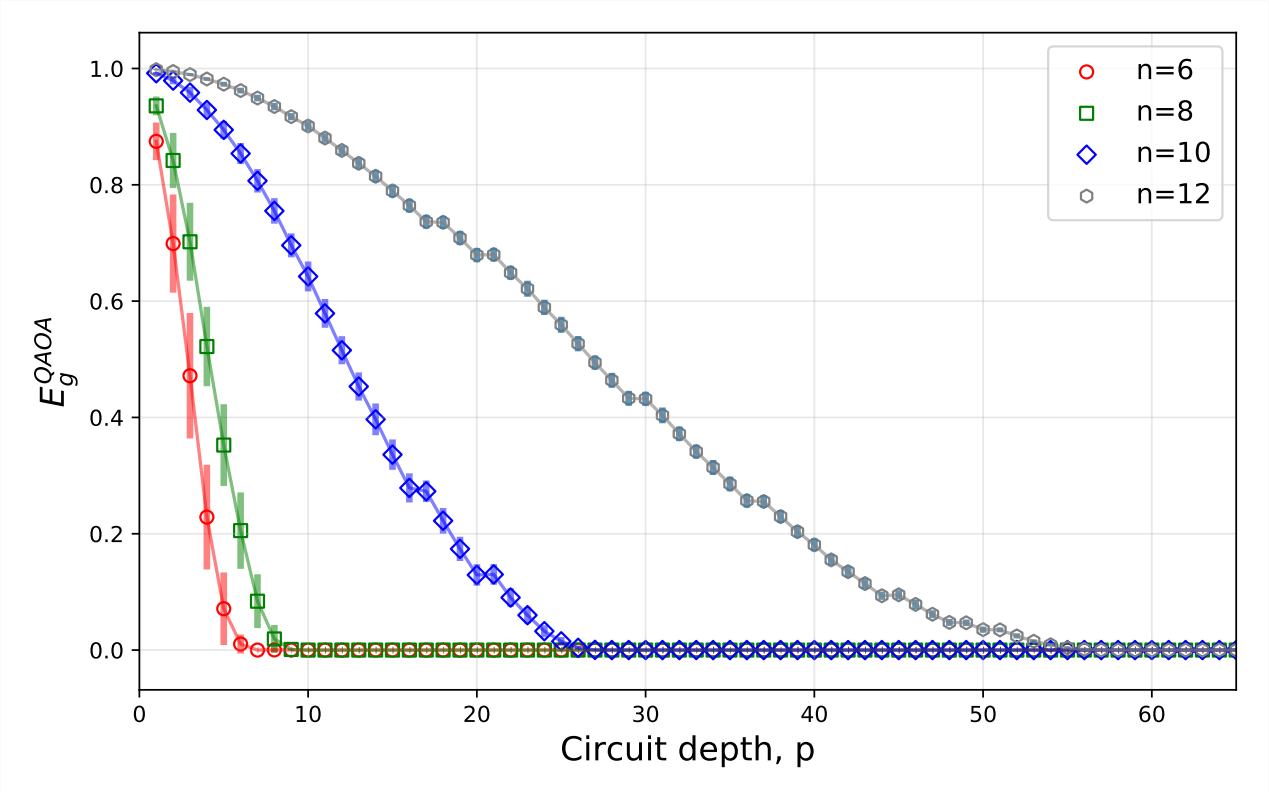


Figure 1: Convergence to exact ground state energy as a function of circuit depth for the variational Grover search on search space sizes, $n = 6, 8, 10$ and 12.

Variational Grover's Search

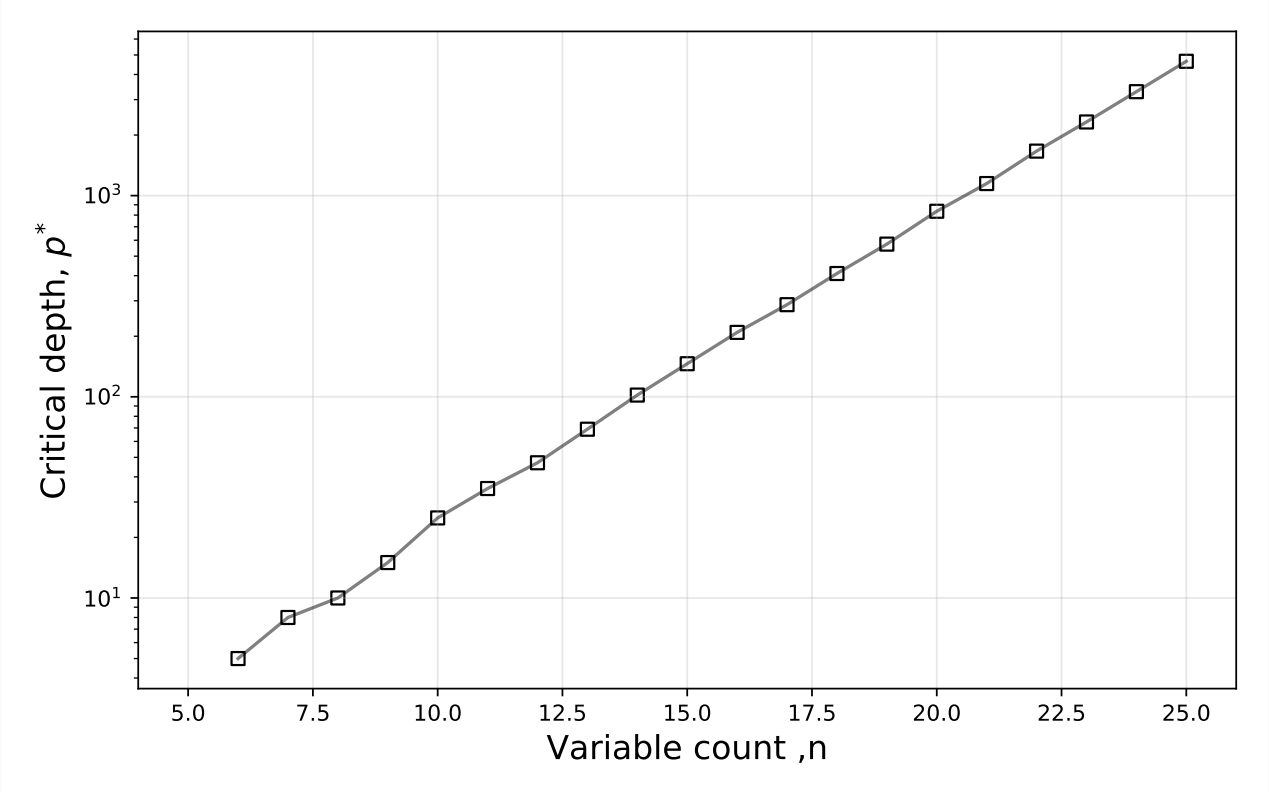


Figure 2: Recovering Grover scaling for critical circuit depth, P^*.

k- Satisfiability

1. *k*-SAT is the decision problem of determining satisfiability (truth value = 1) of Boolean expressions written in Conjunctive Normal Form (CNF) .
2. *k*-SAT clauses are randomly generated to form random instances by uniformly selecting unique *k*-tupels from the union of a variable set (cardinality $N > k$) and its element wise negation.
3. *k*-SAT is **NP**-Complete for $k \geq 3$.

3-Satisfiability

1. Consider the 3-SAT problem instance

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (x_1 \vee x_4 \vee \neg x_5). \tag{9}$$

   The instance (9) has 5 variables, 2 clauses and is satisfiable, as $x_1 = 1,\ x_2 = 1,\ x_3 = 0,\ x_4 = 0,\ x_5 = 0$.

2. An order parameter for random generation is given by the clause density defined by the simple fraction of clause to variable count, $\alpha = m/n$.

Maximum Satisfiability: MAX-*k*-SAT

1. MAX-*k*-SAT is the optimization version of *k*-SAT.
2. The objective is to find assignments that maximize the number of satisfiable clauses in a given instance.
3. MAX-*k*-SAT problems are **NP**-Hard even for $k = 2$.
4. Polynomial time approximation algorithms exist that achieve provable guarantees or approximation factor (r) e.g. for MAX-2-SAT, $r \leq \frac{21}{22}$.

QAOA on MAX-SAT

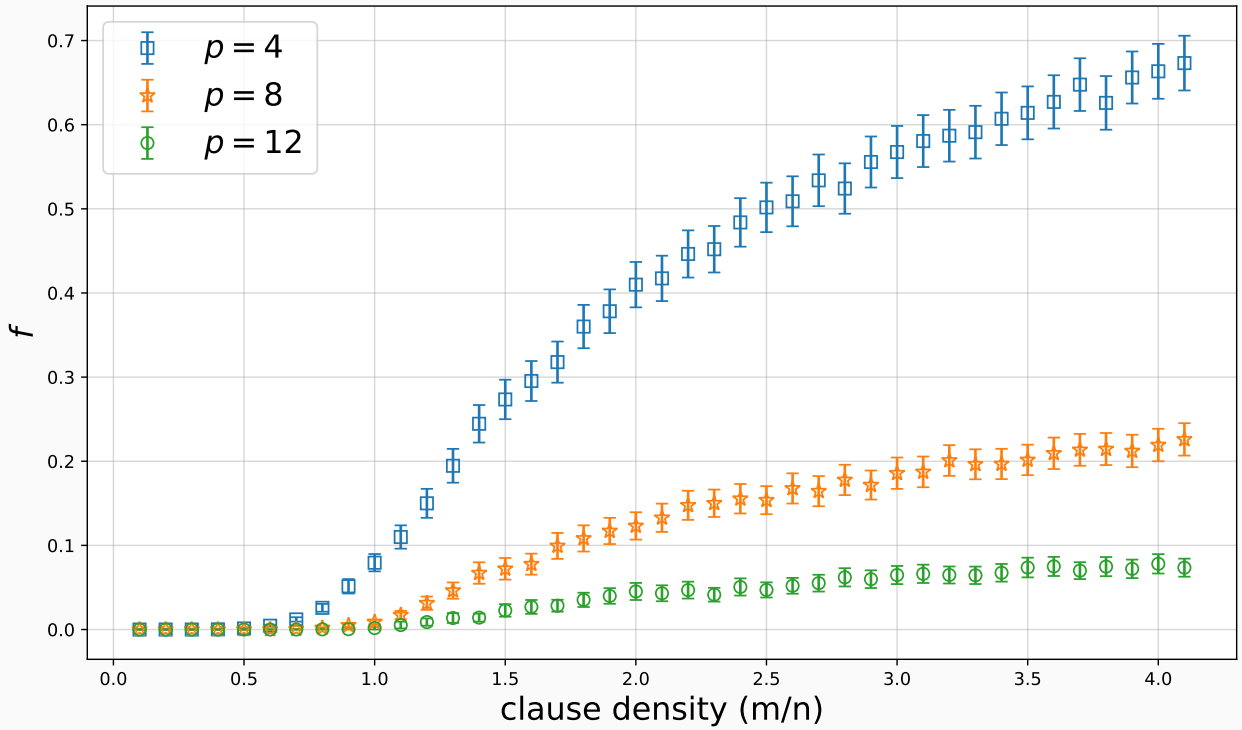


Figure 3: $f = E_g^{\text{QAOA}} - \min(\mathcal{H}_{SAT})$ vs clause density for 10 variables MAX-2-SAT for differing QAOA depths. Higher depths achieve better approximations.

1. Let $\mathcal{V} \geq 0$ and $|\psi\rangle = \prod_{i=1}^{p} \mathcal{U}(\gamma_i, \beta_i) \, |+\rangle^{\otimes n}$.
   We define,

$$\mathcal{F}(p, \alpha, n) = \min_{\psi \in \Omega} \langle \psi | \, \mathcal{V} \, | \psi \rangle - \min_{\phi \in \mathcal{H}} \langle \phi | \, \mathcal{V} \, | \phi \rangle \geq 0, \tag{10}$$

   where α is the problem density, p is the circuit depth and n is the problem size.
2. **Reachability deficits:** $\mathcal{F}(p, \alpha, n)$ is non vanishing for instances beyond a certain problem density for fixed problem size.
3. Mathematically, for depth $p \in \mathbb{N}$ and fixed problem size $n \in \mathbb{N}$, $\exists$ $\alpha \geq \alpha_c$ such that $\mathcal{F} > \epsilon$, for some $\epsilon \in \mathbb{R}_+$.

Density Dependent QAOA Depth

Fixed depth QAOA is limited by *reachability deficits*; Higher depth (or increasing depth) versions are needed to break free from this limitation.

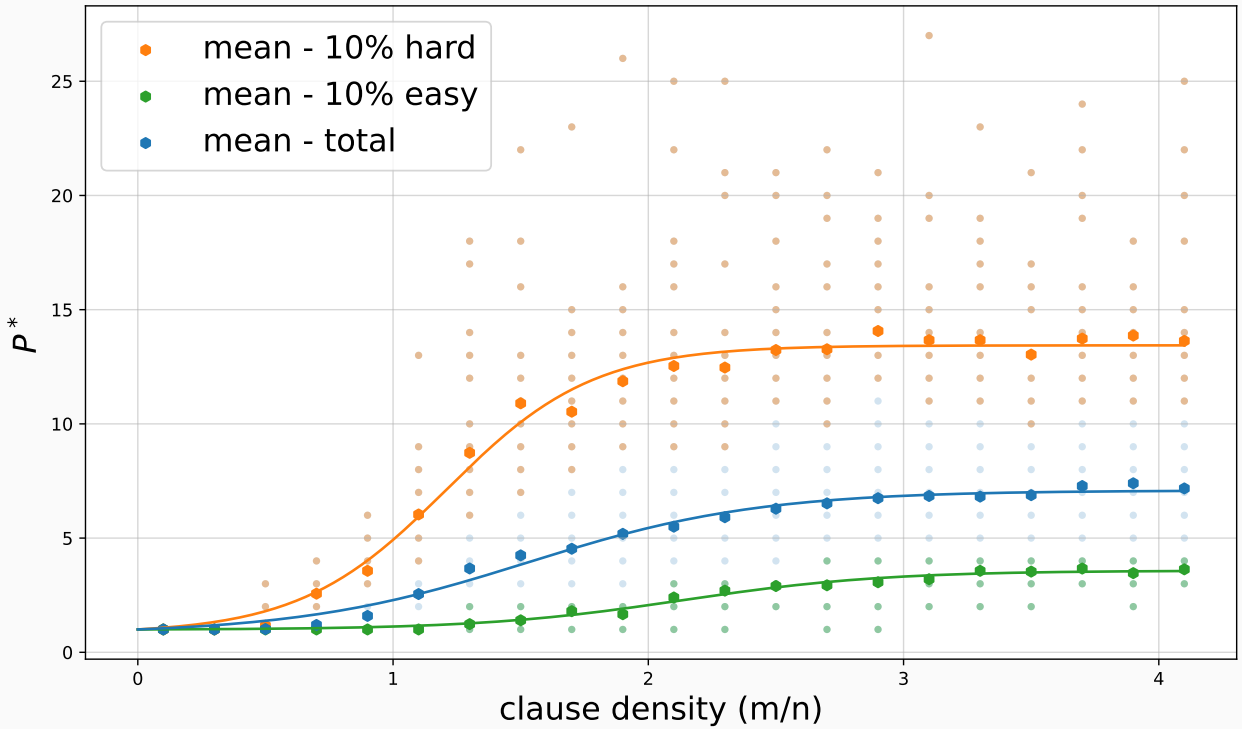


Figure 4: Critical depth, P^* for achieving $\epsilon \leq 0.3$ vs clause density for MAX-2-SAT.

Parameter Concentrations in QAOA [6]

1. Let $|t\rangle$, be some target state of interest in the computational basis.
2. To prepare the target state using QAOA, the ansatz state $|\psi(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle$

$$|\psi(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle = \prod_{k=1}^{p} e^{-i\beta_k \mathcal{H}_x} e^{-i\gamma_k |t\rangle\langle t|} |+\rangle^{\otimes n}, \tag{11}$$

   is variationally tuned to maximize the overlap function $|\langle t|\psi(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle|^2$.
3. Indeed it is observed that, parameters do concentrate with optimal parameters behaving as,

$$n\beta_i \to \pi, \gamma_i \to \pi, \tag{12}$$

   as $n \to \infty$.

[6] **Parameter Concentrations in Quantum Approximate Optimization**
V.Akshay, D.Rabinovich, E.Campos, J.Biamonte
Phys. Rev. A Letters accepted(2021)

Parameter Concentrations in QAOA

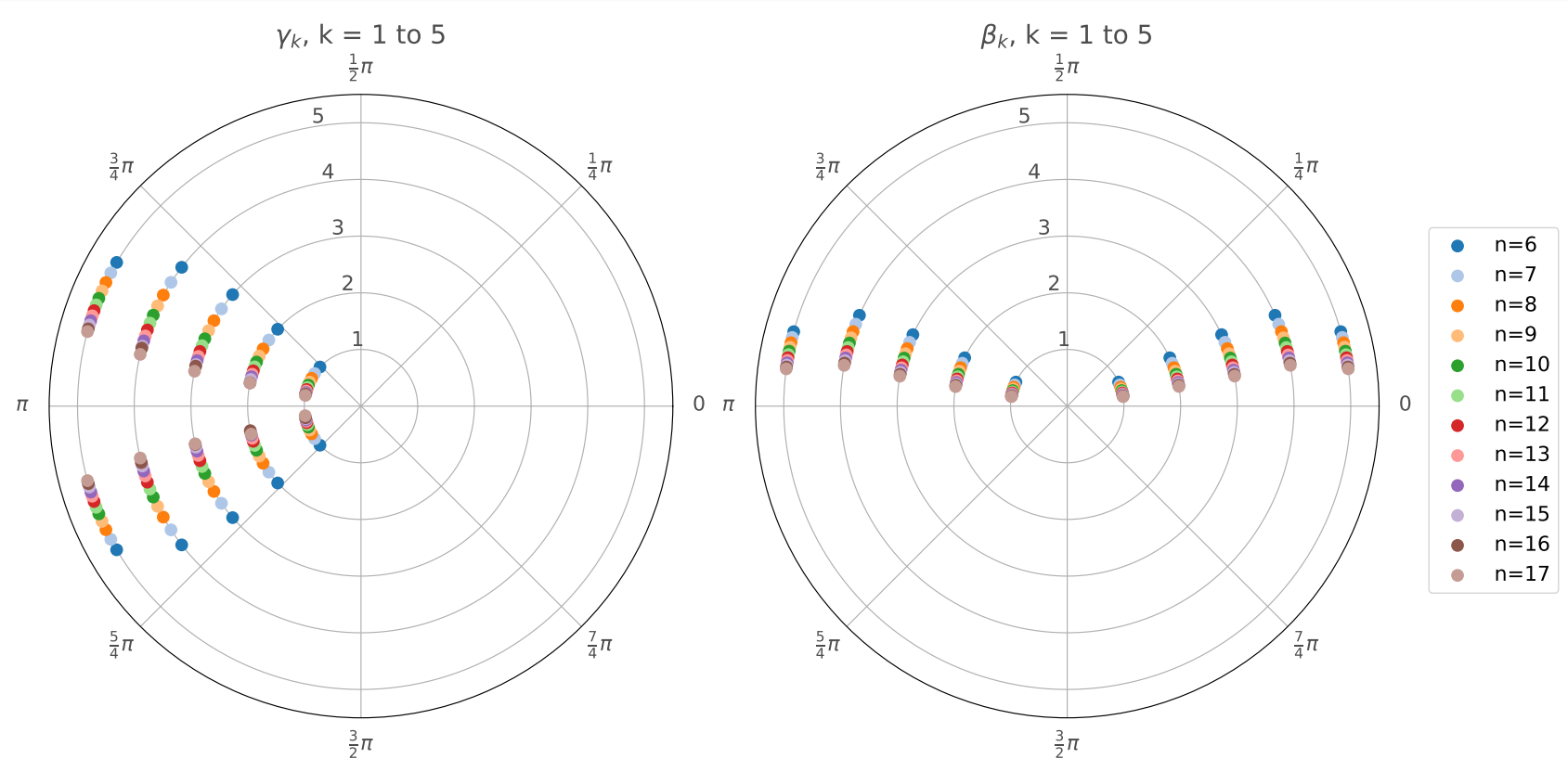


Figure 5: Parameter concentrations visualized for $p = 5$ deep QAOA and varying number of qubits n.

Further Scientific Discussion

Jacob Daniel Biamonte and Ernesto Campos
Skolkovo Institute of Science and Technology
Moscow Russia

J.Biamonte@Skoltech.ru
Ernesto.campos@Skoltech.ru
http://Quantum.Skoltech.ru

Tutorial session

Tutorial: Question 1.1 (10 minutes)

Prove that $e^{-\imath\beta\sigma_x}|+\rangle = e^{-\imath\beta}|+\rangle$.

Hints:

- Taylor expansion $e^A = \sum_{j=0}^{\infty} \frac{A^j}{j!}$
- $\sigma^x = |0\rangle\langle 1| + |1\rangle\langle 0| = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
- $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$

Tutorial: Question 1.2 (15 minutes)

Prove that: $e^{-\imath\beta \sum_j^n \sigma_x^j} = \bigotimes_{j=1}^n e^{-\imath\beta\sigma_x^j}$

Hints:

- $e^A e^B = e^{A+B}$ for $[A, B] = 0$
- if A and B are defined in different spaces, product reduces to $\otimes$

Tutorial: Question 1.3 (5 minutes)

Deduce that $e^{-\imath\beta \sum_{j=1}^{n} \sigma_x} |+\rangle^{\otimes n} = e^{-\imath\beta n} |+\rangle^{\otimes n}$

Tutorial: Question 2 (5 minutes)

Prove that $e^{-\imath\gamma|t\rangle\langle t|} = \mathbb{1} + (e^{-\imath\beta} - 1)|t\rangle\langle t|$

Tutorial: Question 3 (15 minutes)

Compute the overlap $\langle t|\psi(\beta,\gamma)\rangle$ where $|\psi(\beta,\gamma)\rangle = e^{-i\beta_k \mathcal{H}_x} e^{-i\gamma_k |t\rangle\langle t|} \, |+\rangle^{\otimes n}$

Hint: $e^{-\imath\sigma_x\beta} = \mathbb{1}\cos(\beta) - \imath\sigma_x \sin(\beta)$