Hot-Start Optimization for Variational Quantum Eigensolver

Belozerova Polina\textsuperscript{1}, Shangareev Arthur\textsuperscript{2}, Zotov Yuriy\textsuperscript{3}, Yung Manhong\textsuperscript{3}, and lv Dingshun\textsuperscript{3}

\textsuperscript{1}Moscow Institute if Physics and Technologies, Moscow, Russia
\textsuperscript{2}Moscow State University, Moscow, Russia
\textsuperscript{3}Central Research Institute, Huawei Technologies

May 3, 2021

Abstract

The Variational Quantum Eigensolver (VQE) is one the most perspective algorithms for simulation of quantum many body physics that have recently attached a lot of attention and believed would be practical for implementation on the near term quantum devices. However, its feasibility and accuracy critically depend on the ansatz structure, which can be defined in different ways and appropriate choosing the structure presents a bottleneck of the protocol. In our work we present the optimization approach allowing shallow circuit solution. The major achievement of our Hot-Start method is the requiring gates number restriction by several times. In result it increases the fidelity of the final results what is important for the current noise quantum devices implementation. We suggest Hot-Start optimization to be a good new methods beginning for the modern quantum computational devices functionality development.

1 INTRODUCTION

One of the main problem of physics and quantum chemistry is to find the many body system ground state. There is no analytical solution for the general case and the computation methods are limited as the complexity of problem in Hilbert space grows exponentially with the number of involved particles. The classical polynomial-time solver does not exist.

Density-functional theory (DFT) \cite{16},\cite{17} and wavefunction- based methods such as configuration interaction (CI) \cite{19},\cite{18} or coupled-cluster (CC) \cite{21}, \cite{20} are used worldwide for quantum mechanical modeling. For strongly correlated systems tensor network states \cite{22} are often used. Each of these methods face exponential computational cost. One of the possible reason is that the number of Slater determinants might be large. Therefore, modern classical computers are limited for tasks, which have applications in new chemical structures search, as well material physical properties study.

VQE algorithm became the embodiment of Richard Feynman’s proposal to simulate one quantum system using another, more controllable \cite{23}. This is possible due to qubits property to form the superposition and keep the entanglement. As a result, the problem is elegantly reduced from the whole ground state $O(2^n)$ superposition coefficients search to just finding the $O(n)$ angles of qubits vectors rotation and their entanglement.

Despite the significant simplification from an analytical computational point of view, in the practical implementation many difficulties appear. They are related to both the complex goal function surface and the high dimensionality when the number of parameters has to increase.

2 QUANTUM CHEMISTRY PROBLEM

2.1 Hamiltonian of the System

In quantum mechanics systems on the atomic scale can be mathematically described by the wavefunction, which accounts all the nuclei and electrons: $\Psi(r_1, r_2, \ldots, r_N, R_1, R_2, \ldots, R_I)$, where $r_N$ are electrons positions including spin information and $R_I$ are nuclei positions. For condensed matter study time dependence is usually omitted and the wavefunction
evolution is described by the Schrödinger equation for many-body Hamiltonian:

\[ i\hbar \frac{\partial}{\partial t} \Psi = \hat{H} \Psi \]  

We will use physical constants are set to unity for convenience. The hamiltonian operator is a sum of kinetic energy and potential operators, \( \hat{T} \) and \( \hat{V} \). The many-body energy operators can be expressed as:

\[ \hat{H} = \hat{T} + \hat{V} \]  

\[ \hat{T} = -\frac{1}{2} \left( \sum_N \nabla_N^2 + \sum_I \frac{1}{M_I} \nabla_I^2 \right) \]  

\[ \hat{V} = \sum_{N,I} \frac{Z_I}{|R_N - R_I|} + \sum_{N \leq K} \frac{1}{|R_N - R_K|} + \sum_{I \leq J} \frac{Z_I Z_J}{|R_I - R_J|} \]  

Where \( M_I \) are the masses of the nuclei and \( R_I \) are their coordinates, and \( R_N \) are coordinates of electrons. The potential energy operator is made up of the Coulomb interactions, neglecting external fields. The three terms are the electron-nuclear, electron-electron and nuclear-nuclear interactions, respectively. The molecular geometry implicitly specification assumes the Born-Oppenheimer approximation with fixed point charges. Thus the ground state electronic energy is a parametric function of their positions. The popular methods to find the wavefunction satisfying the Schrödinger equation, is based on DFT [2].

Thereby the canonical ground state problem may be formulated as an optimization problem with \( O(2^n) \) complexity, where \( \psi \) is a coefficients vector of the resulting superposition in Hilbert space:

\[ \min \langle \psi | \hat{H} | \psi \rangle \]  

\[ s.t. \quad \| \psi \|_2 = 1 \]  

2.2 Second quantization form

After molecule was specified, a particular basis was chosen and the physical anti-symmetry of electrons was enforced, the electronic structure problem may be written exactly in the form of a second quantized electronic Hamiltonian as:

\[ H(R) = \sum_{ij} h_{ij}(R)a_i^\dagger a_j + \sum_{ijkl} \frac{1}{2} h_{ijkl}(R)a_i^\dagger a_j^\dagger a_k a_l \]  

Where \( R \) is a vector of pairwise distances between particles in the considered system, \( a_i^\dagger \) and \( a_i \) are creation and annihilation operators, and \( h_{ij} \) and \( h_{ijkl} \) are one-electron and two-electron integrals correspondingly [3]:

\[ h_{ij} = \int \phi_i^\ast(r) \left( -\frac{\nabla^2}{2} - \sum_I \frac{Z_I}{|R_I - r|} \right) \phi_j(r) \, dr \]  

\[ h_{ijkl} = \int \phi_i^\ast(r_1) \phi_j^\ast(r_2) \phi_k(r_1) \phi_l(r_2) \, dr_1 \, dr_2 \]  

For the computational proposes it can be accomplished by OpenFermion and by supported interfaces Psi4 [3] and PySCF [4]. The chosen basis set such as e.g. sto-3g or 6-31g defines the interacting integrals coefficients \( h_{ij} \) and \( h_{ijkl} \).

2.3 Qubit mapping

After the problem has been written in the second quantized representation, it remains to map it to qubits, as the original problem is formulated for electrons, which are anti-symmetric indistinguishable particles, while qubits are distinguishable particles. There are various of mapping, called encoding, that respect the correct particle statistics. The most popular encodings are the Jordan-Wigner (JW) [6], Bravyi-Kitaev (BK)[7] and Bravyi-Kitaev super fast (BKSF) [8] transformations supported in OpenFermion. In the present work we mostly worked with Bravyi-Kitaev encoding.

3 VQE BACKGROUND

3.1 Problem reformulation

The original idea of VQE is the superposition realization on qubits capable to hold a quantum state [9]. It restricts the number of optimization parameter to \( O(n) \) angles of qubits rotation and well-tuned entanglement, where \( n \) is a number of qubits. Thus, the optimization problem is reformulated to:

\[ \min \langle \psi_{HF} U^\dagger(\bar{\theta}) | H | U(\bar{\theta}) \psi_{HF} \rangle = \lambda_{gs} \]  

\[ s.t. \quad \{ \theta_1, \ldots, \theta_N \} \in [0, 2\pi] \]  

Where \( | \psi_{HF} \rangle \) is a Hartree Fock state, the \( H \) is problem Hamiltonian and \( \lambda_{gs} \) is a ground state energy. The unitary evolution \( U(\bar{\theta}) \) usually is represented as a Pauli strings exponentiation:

\[ U(\bar{\theta}) = U_1 \ldots U_N \]  

\[ U_k = U(\theta_k) = \exp (-i\theta_k P_k) \]  

\[ P_k = \bigotimes_{i=1}^n \sigma_{ik} \]  

\[ \sigma_i \in \{ I, X, Y, Z \} - \text{Pauli matrices} \]
forming the evolution of trial state to the ground state.

3.2 The existing methods and the bottleneck

One of the VQE research direction is to find a proper structure of ansatz for particular Hamiltonian\[^{[12,13]}\]. The final circuit has to be comprehensive for optimization in terms of rotation angles and entanglement. As there is no analytical solution for the circuit, this problem remains to be unsolved and to be a wide field for the research.

However, there are two generally accepted methods for quantum chemistry simulations via VQE:

- UCC-based ansatzes\[^{[14]}\] and
- hardware efficient\[^{[9]}\].

The UCCSD ansatz is obtained by replacing the traditional Hermitian operator terms in coupled cluster’s theory with anti-Hermitian operators. It has a complex structure and it is complicated to implement on quantum hardware due to the large number of gates.

Hardware efficient ansatzes were proposed for NISQ use. The main idea is to stack similar layers - depths - of rotation and entanglement operators several times\[^{[9,10,11]}\]. Increasing the number of depths repeatedly one gives the additional chances for the algorithm to converge and in this way the final optimization error has a trend to decrease. The full form of layer circuit is:

\[
\min \langle \psi_{HF} | U_D^\dagger \ldots U_1^\dagger \psi_{HF} \rangle \\
\text{s.t.} \quad \{ \vec{\theta}_1, \ldots, \vec{\theta}_D \} \in [0, 2\pi]^{\otimes nd}
\]

But the computational cost grows as a number of optimization parameters \(O(nd)\). It makes complicated to use VQE for big quantum systems due to optimization methods limitations:

- first order optimization methods require thousands of iteration but
- second order methods become unable to deal with number of parameters (angles) due to the large number Jacobian elements and fails because of accumulative computational errors.

Such ambiguity in the number of parameters in the classical part of VQE forms the bottleneck. Also there are two methods - analytical and computational - to estimate gradient for the optimization but each of them also has disadvantages for the described problem.

The analytical method includes the derivative of subject function:

\[
f(\vec{\theta}) = \langle H \rangle_{\vec{\theta}} = \langle 0 \ U_{\theta_1}^\dagger \ldots U_{\theta_D}^\dagger \ | \ H \ | \ U_{\theta_1} \ldots U_{\theta_D} \ 0 \rangle 
\]

\[
\frac{\partial}{\partial \theta_k} U_{\theta_k} = -i \theta_k P_k \cdot \exp(-i \theta_k P_k) = V_{\theta_k}
\]

Introducing the convenient notation one can derive:

\[
U_- = U_{\theta_1} \ldots U_{\theta_{k-1}}, \\
U_+ = U_{\theta_{k+1}} \ldots U_{\theta_D}
\]

\[
\frac{\partial}{\partial \theta_k} f(\vec{\theta}) = \frac{\partial}{\partial \theta_k} \langle H \rangle_{\vec{\theta}} = \langle 0 \ U_-^\dagger \ | \ [V_{\theta_k}, U_+^\dagger H U_+] \ | \ U_- \ 0 \rangle
\]

Thereby, corresponding to Poisson bracket the analytical method requires to estimate expectation value of the quantum scheme twice for each parameter of optimization.

The computational method is exactly the explicit difference scheme which requires to call expectation function \(nd + 1 = D + 1\) times for first order accuracy method:

\[
\frac{\partial f}{\partial \theta_k} = \frac{f(\theta_1, \ldots, \theta_k + \Delta \theta, \ldots, \theta_D) - f(\theta_1, \ldots, \theta_k, \ldots, \theta_D)}{\Delta \theta_k}
\]

If one uses the second-order accuracy scheme, it will appropriately require \(2D\) iterations to compute gradient each time. As a result all approaches are clumsy and do not simplify optimization of any order.

Methods mentioned above have one thing in common: they do not take into account any property of the original objective function.

4 HOT-START METHOD

4.1 New Approach Idea

Previously not an optimization methods but ansatzes were build basing on the main property of Hamiltonian and its Pauli strings based structure:

\[
H = \sum_k h_k P_k, \quad P_k = \bigotimes_{i=1}^{n} \sigma_{ik}
\]

\[
|U_k| = |\exp(-i \theta_k P_k)| = 1
\]

But in our work we focused on the analogous property of Pauli strings - eigenvalues \(\lambda\) modulo value:
Figure 1: Maximum possible integral coefficients contribution to the ground state value computed as FCI energy. For all molecules sto-3g orbital basis was used. (a) Ordinary 4 qubits H2. (b) Restricted 6 qubits LiH version was received by [0, 1, 2] electrons freezing. (c) 8 qubits LiH has [0, 4] frozen electrons

\[ \lambda(P) = \pm 1 \]  \hspace{1cm} (22)

This property might be interpreted as a maximum possible Hamiltonian sum member contribution to the target ground state.

In one of the last works [15] authors demonstrated the adiabatic method using a linear approximation of the problem: \( H(t) = A(t)H_0 + B(t)H \). In this case all Pauli strings will contribute to the resulting expectation value simultaneously and continuously over time. It makes the convergence dependence on Hamiltonian structure complex, because in adiabatic approach one adds all Pauli strings at the same time uncontrolled. But in comparison the elementary parts contribution is very clear and equal to \( \langle \psi | P_k | \psi \rangle \).

In contrast, Hot-Start method proposed in this work approximates the objective function step by step, complicating its structure gradually. Our idea is the discrete approach with limited function representation iterations. It is based on the set of Pauli strings and corresponding integral coefficients. Hamiltonian construction by the algorithm allows to manage matrix sparsity slowly.

First, we studied the absolute values distribution for the near ground state Hamiltonians for \( \text{H}_2 \) and \( \text{LiH} \) molecules. As it can be seen from the Figure 1, the received distributions are highly inhomogeneous. The largest absolute integral coefficient tends to the ground state energy value. The analytical explanation this prevalent effect is that Number and Coulomb/exchange operators contribute to the ground state more than correlation terms as Excitation, Number-excitation and Double excitation operators, which was mentioned in [8]. These operators define the significance of the integral influence to the ground state.

One may notice that in sorted absolute values of Hamiltonian coefficients the received distribution first members slope is especially sharp. This means that the largest elements may have the great contribution to the ground state value. We propose that solving the problem for the largest elements in the first place may help to avoid local minimum caused by the rest part of the Hamiltonian. The fastest sorting algorithm has \( O(K \log K) \) complexity, where max \( K \) is a Pauli string basis size equal to 4^n. This is a high complexity, but as in practice the number of Hamiltonian sum is much less wherein computational cost is not so resource demanding for many systems of interest.

After the sorting is successfully done the smooth approximation of the objective function to the final one demonstrated meaningful results. The physical interpretation of this approach allows us to call this method a Hot-Start - we gradually rotate the qubits to their best angles, focusing first on the most important Hamiltonian elements, starting from point supposedly close to the absolute minimum. Thus Hot-Start is a method based on the Hamiltonian properties, and this is the objective function based optimization.
4.2 The algorithm

Hot-Start method is represented in Algorithm 1 for encoded Hamiltonian.

**Input:** Starting parameters are fixed random angles of initial gates rotations $\vec{\theta}$

**Output:** Ground state of $H$

**Data:**
Encoded $H = \sum_i^n h_i \otimes \sigma_i$ of the given system, where $\sigma_j \in \{I, X, Y, Z\}$

- Sort $H_i$ in the descending order of the $h_i$ absolute values.
- Receive the sequence of $\{\hat{H}_i\}$, where $|\hat{h}_1| \geq |\hat{h}_2| \geq \cdots \geq |\hat{h}_N|$.
- Start iterative optimization:

$$k = 1$$

**while** $k \leq N$ **do**

$$A_k = \sum_1^k \hat{H}_i$$

Solve by chosen method:

$$\theta_{k+1} = \arg\min \langle \psi | U^\dagger(\vec{\theta}_k) A_k U(\vec{\theta}_k) | \psi \rangle$$

$k = k + 1$

**end**

**Result:** $\theta_{N+1}$ is a final set of angles for the original VQE problem

**Algorithm 1:** Hot-Start method

Since sorting part of the algorithm may have high asymptotic complexity firstly we studied the importance of members selection method: we additionally conducted tests with Hamiltonian elements sorted in ascending order and also added randomly. The results are shown in the Figure 2.

5 RESULTS

The second order algorithms such as BFGS and LBFGS-B were used due to their high efficiency to test the ordinary approach and Hot-Start optimization. To research whether Hamiltonian-based optimization can work better than the Hamiltonian-based ansatz we compare the average error of convergence to the known ground state. Corresponding ground state values were computed by eigen-decomposition as number of qubits could be run efficiently on simulator for chemistry problem is not high. It was possible to test the algorithm only on a narrow range of qubits such as:

- 4 qubits H2, sto-3g basis;
- 6 qubits restricted LiH, sto-3g basis;
- 8 qubits restricted LiH, sto-3g basis.

The reduction of LiH to receive 6-qubits model was done by freezing $[0, 3, 4]$ orbitals as it corresponds to chemical configuration priority. This orbital combination produces 118 Pauli strings. As Hot-Start method has a high sensitivity to the number of Hamiltonian elements, hence we also designed artificial problems with less number of Pauli strings:

- 6-qubits LiH with frozen $[0, 1, 2]$ orbitals and Hamiltonian with 34 sum members;
- 8 qubits LiH with frozen $[0, 3]$ orbitals and Hamiltonian with chosen only 10 elements - each 20th member of it.

Bravyi-Kitaev encoding, Ry ansatz and CNOT full-entanglement gates were used as quantum circuits parameters. The optimization was conducted by 10 randomizations to estimate the uniform convergence ability of methods.

The results with desired chemical accuracy of convergence are represented in the Figure 2. We would interpret results for 4 qubits in a way that the original problem is simple enough by itself and that is why it was not a high difference for ordinary and Hot-Start optimization.

For 6 qubits LiH model with 34 Pauli strings we found that Hot-Start optimization works much better than usual approach: the ground state of the simple BFGS method error remains to be about $10^{-1}$ Hartree while Hot-Start approach converges well to $10^{-10}$ Hartree beating the chemical accuracy at the very shallow depth. For the ordinary approach at least
6 depths are required to converge when Hot-Start optimization requires three times less - only 2 depths.

Results for 8 qubits model demonstrate, that for the ground state is hardly achievable for any kind of optimization method.

Thereby there is a precedent case when convergence is achievable even for shallow circuit. And despite the Hot-Start method can not be widely universal due to its linear dependence on number of Pauli strings, it demonstrates the ability to optimize problem with less number of gates. This is important observation for quantum hardware implementation.

6 DISCUSSION AND CONCLUSIONS

We have presented series of results to demonstrate the possibility of Hamiltonian based optimization approaches. The Hot-Start method is the most direct one and might be computationally hard but it proved an important idea: if ordinary optimization methods do not converge to the ground state well, the problem might be not in the chosen ansatz, but in the optimization method for the object function overall.

The represented Hot-Start results can be a good beginning for such a type of optimization, when one does not need anymore to use UCCSD ansatzes to converge, because this type of Hamiltonian-based circuits is large, noisy and complicated for physical implementation. But a new methods require a guess about the optimization process itself. We would propose that may be there is an optimal batch optimization exists, where the batch is a group of Hamiltonian Pauli strings united with some general property.

The most obvious and intuitive would be expect physical features to work well. We divided the Hamiltonian into groups as: Number, Coulomb/exchange, Excitation, Number-excitation and Double excitation operators. This approach has an advantage that in this way Hamiltonian has a fixed number of groups - five - and complexity does not depend on the $O(n^4)$ possible Pauli strings.

We implemented the realization of these experiments but revealed that it does not work well for described molecules and other cases. Accordingly to these results the formulation of the batch optimization remains to be an open question.

References

[1] Peruzzo, A. "A variational eigenvalue solver on a quantum processor. eprint." arXiv preprint arXiv:1304.3061 (2013).

[2] Grimme, S. (1996). Density functional calculations with configuration interaction for the
excited states of molecules. Chemical physics letters, 259(1-2), 128-137.

[3] Babbush, R., Berry, D. W., Kivlican, I. D., Wei, A. Y., Love, P. J., Aspuru-Guzik, A. (2016). Exponentially more precise quantum simulation of fermions in second quantization. New Journal of Physics, 18(3), 033032.

[4] Sun, Q., Berkelbach, T. C., Blunt, N. S., Booth, G. H., Guo, S., Li, Z., Chan, G. K. L. The Python-based Simulations of Chemistry Framework (PySCF), 2017. Google Scholar There is no corresponding record for this reference.

[5] Parrish, R. M., Burns, L. A., Smith, D. G., Simmonett, A. C., DePrince III, A. E., Hohenstein, E. G., Gonthier, J. F. (2017). Psi4 1.1: An open-source electronic structure program emphasizing automation, advanced libraries, and interoperability. Journal of chemical theory and computation, 13(7), 3185-3197.

[6] Jordan, P., Wigner, E. P. (1993). über das paulische äquivalenzverbot. In The Collected Works of Eugene Paul Wigner (pp. 109-129). Springer, Berlin, Heidelberg.

[7] Bravyi, S. B., Kitaev, A. Y. (2002). Fermionic quantum computation. Annals of Physics, 298(1), 210-226.

[8] Setia, K., Whitfield, J. D. (2018). Bravyi-Kitaev Superfast simulation of electronic structure on a quantum computer. The Journal of chemical physics, 148(16), 164104.

[9] Kandala, A., Mezzacapo, A., Temme, K., Takita, M., Brink, M., Chow, J. M., Gambetta, J. M. (2017). Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. Nature, 549(7671), 242-246.

[10] Salis, G., Moll, N. (2019). Short-depth trial-wavefunctions for the variational quantum eigensolver based on the problem Hamiltonian. arXiv preprint arXiv:1908.09533

[11] Mitarai, K., Yan, T., Fujii, K. (2019). Generalization of the output of a variational quantum eigensolver by parameter interpolation with a low-depth ansatz. Physical Review Applied, 11(4), 044087.

[12] Grimsley, H. R., Claudino, D., Economou, S. E., Barnes, E., Mayhall, N. J. (2019). Is the Trotterized UCCSD Ansatz Chemically Well-Defined?. Journal of Chemical Theory and Computation, 16(1), 1-6.

[13] Grimsley, H. R., Economou, S. E., Barnes, E., Mayhall, N. J. (2019). An adaptive variational algorithm for exact molecular simulations on a quantum computer. Nature communications, 10(1), 1-9.

[14] Romero, J., Babbush, R., McClean, J. R., Hempel, C., Love, P. J., Aspuru-Guzik, A. (2018). Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. Quantum Science and Technology, 4(1), 014008.

[15] Garcia-Saez, A., Latorre, J. I. (2018). Addressing hard classical problems with adiabatically assisted variational quantum eigensolvers. arXiv preprint arXiv:1806.02287

[16] Engel, E., Dreizler, R. M. (2013). Density functional theory (p. 65). Springer-Verlag Berlin An.

[17] Sholl, D., & Steckel, J. A. (2011). Density functional theory: a practical introduction. John Wiley & Sons.

[18] Pople, J. A., Head-Gordon, M., & Raghavachari, K. (1987). Quadratic configuration interaction. A general technique for determining electron correlation energies. The Journal of chemical physics, 87(10), 5968-5975.

[19] Shavitt, I. (1977). The method of configuration interaction. In Methods of electronic structure theory (pp. 189-275). Springer, Boston, MA.

[20] Crawford, T. D., & Schaefer, H. F. (2000). An introduction to coupled cluster theory for computational chemists. Reviews in computational chemistry, 14, 33-136.

[21] Bartlett, R. J., & Musial, M. (2007). Coupled-cluster theory in quantum chemistry. Reviews of Modern Physics, 79(1), 291.

[22] Murg, V., Verstraete, F., Legeza, Ö., & Noack, R. M. (2010). Simulating strongly correlated quantum systems with tree tensor networks. Physical Review B, 82(20), 205105.
[23] Trabesinger, A. (2012). Quantum simulation. Nature Physics, 8(4), 263-263.