Implementation of Quantum Linear Solver Algorithm: Review

Mohd Harith Akmal Zulfaizal Fadillah, Bahari Idrus, Mohammad Khatim Hasan

Centre for Artificial Intelligence Technology (CAIT), Faculty of Information Science & Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor

Abstract

Solving linear systems of equations is fundamental in many areas of science and engineering. As data sets get bigger, the task of classically solving linear systems and especially inverting a matrix gets computationally harder. Quantum linear solver algorithm (QLSA) provides an algorithm that promises speed-up in solving linear systems which could prove beneficial. However, the algorithm does have its own caveats which has pushed researchers to improve existing quantum linear solver algorithms. In this study, we will look at the original Harrow-Hassidim-Lloyd (HHL) algorithm and its improvements as well as their implementations. We will also be looking at more recent developments on hybrid quantum linear solver and variational quantum linear solver (VQLS) algorithms and their implementations on IBM Quantum machines.

1 Introduction

As the Moore’s law is coming to an end with the ever increasingly huge data sets making problem solving computationally harder, the field of quantum computation has proved, by theory, that it is able to overcome existing limitations observed in current classical computing [1, 2]. One of the most important and fundamental problem to be solved in many areas of science and engineering is the problem of solving linear systems of equations. This relatively simple problem by classical standard opens up new possibilities in a variety of fields such as machine learning and differential equations among others. The importance of solving linear systems has pushed quantum computing and algorithm researchers to devise quantum linear solver algorithms (QLSA) which has existed for some time and contributed to the field of quantum machine learning.

However, researchers are still developing new improvements to existing QLSA and devising new potential algorithms that could prove to be more viable with current Noisy Intermediate Scale Quantum (NISQ) computers. We have looked theoretically and in depth at the original Harrow-Hassidim-Lloyd (HHL) algorithm in our previous report [3] on quantum algorithms. In this report, we will be focusing more on the implementations that have been realized of several quantum linear solver algorithms on circuit based quantum computers. We will look at the original HHL and its improvements as well as the implementation of an optimized circuit of the algorithm. Afterwards, we will focus on recent developments on hybrid QLSAs and their implementations on IBM Quantum (IBM-Q) machines and observe how different hybrid algorithms vary with each other in their construction and outcomes. Finally, we will look at a variational approach of solving linear systems which are more reliable in current NISQ quantum computers.
2 General quantum linear solver algorithm

The quantum algorithm for linear systems of equations devised initially by Harrow, Hassidim and Lloyd (2009) is generally known as the Harrow-Hassidim-Lloyd (HHL) algorithm [4]. This algorithm yields a scalar measurement of the solution vector as its output and does not give the actual values of the solutions although it is possible through state tomography, it is very inefficient and expensive. The algorithm provides a significant speedup over its classical counterpart on the condition that the linear system is sparse and the condition number $\kappa$ which is the ratio of the maximum eigenvalue over the minimum is low. The algorithm runs at time $O(\kappa^2 s^2 \log_2 N)$ where $s$ is the sparseness and $\epsilon$ is the precision, which trumps over the fastest classical conjugate-gradient method runtime of $O(N\kappa s \log(1/\epsilon))$ providing exponential speedup in $N$ but linearly slower in $s$ and $\kappa$.

The subroutine that makes up the HHL algorithm includes quantum phase estimation which is the algorithm used for the matrix inversion needed in solving linear equations. Tweaks and alterations have been made to improve HHL algorithm in terms of its runtime complexity with respect to the condition number $\kappa$ and the precision $\epsilon$. Ambainis (2010) managed to improve by reducing the dependence of condition number from $\kappa^2$ to $\kappa \log^3 \kappa$ by introducing variable time amplitude amplification resulting in the matrix inversion being BQP-complete [5]. Further improvements by Childs, Kothari and Somma (2017) managed to reduce the precision number dependence from $O(poly(1/\epsilon))$ to $O(poly \log(1/\epsilon))$ [6]. Improvement has also been made in making the algorithm possible for dense matrices where Wossnig, Zhao and Prakash (2017) introduced an algorithm based on quantum singular value estimate (QSVE) providing polynomial speedup for dense matrices, $O(\kappa^2 \sqrt{n \log(n/\epsilon)})$ [7].

![Figure 1: a. Original HHL algorithm. b. Optimized quantum circuit for HHL algorithm solving a 2 x 2 linear system of equations. [8]](image)

An experimental implementation of HHL algorithm is reported by Cai, Weedbrook et al (2013) by demonstrating and solving a 2 x 2 linear equations for various input vectors [8]. The circuit for the HHL algorithm is optimized and compiled into a 4-qubit circuit composed of the ancilla, phase and input registers as in Figure 1. The paper demonstrated the experiment on
the matrix,

\[
\hat{A} = \begin{pmatrix}
1.5 & 0.5 \\
0.5 & 1.5
\end{pmatrix}
\]

for input vectors,

\[
|b_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad |b_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad |b_3\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}
\]

The algorithm is implemented and the measurement of expectation values of the Pauli observables X, Y and Z for each input vectors \(|b\rangle\) are obtained. Figure 2 shows the ideal and experimental expectation values for each observables, \(\langle x | M | x \rangle\). The fidelity can be obtained by calculating the inner product between \(|x\rangle\) and the density matrix from the output state, \(\rho_x\), \(\langle x | \rho_x | x \rangle\). Relative to the ideal outcomes, the yielded output from the algorithm have fidelities of 0.993, 0.825 and 0.836 for \(|b_1\rangle\), \(|b_2\rangle\) and \(|b_3\rangle\) respectively. The difference in the fidelities lies in the construction of the algorithm circuit on the optical setup causing high-order photon emission events and post-selection in CNOT gates.

![Figure 2: Experimental results on the expectation values for X, Y and Z observables for |b_1\rangle, |b_2\rangle and |b_3\rangle respectively. [8]](image)

In HHL algorithm and its subsequent improvements, the problem of encoding vector and matrix into quantum states still persist [9]. The encoding, while not a part of the algorithms itself, are crucial to solving the linear systems but yet it is rarely touched upon in the literature. The problem of encoding these data are often considered as a separate non-trivial problem which can make or break the QLSA. Theoretically, the encoding can achieve exponential compression with \(O(\log n)\) qubits and to make the HHL algorithm reliable, the running time of the encoding algorithm should be in \(O(\text{polylog}(n))\) time [9].

The oracle quantum random access memory (QRAM) is the standard model for memory allowing queries in quantum superposition which have studied extensively as a solution to the encoding problem. Real-valued vector are manipulated into amplitudes of the quantum state which is the elements of the vectors scaled by the norm using operation \(R\) which should run at most polylog in \(N\). Figure shows the vector state preparation for 4-dimensional state \(|\phi\rangle\) in the amplitudes of the quantum state. An improved augmented QRAM which pre-process all the vector elements classically before loading is proposed by Prakash (2018) [10]. It utilizes quantum key value map to insert the norm elements into the superposition via conditional rotation and amplitude amplification. Another contending solution to the encoding problem is using the technique by Grover and Rudolph (2002) technique which generates a quantum superposition that encodes an approximate version of a classical probability distribution provided its density...
is efficiently integrable [11]. However, this solution has not been explored much as QRAM has shown to be more viable option in most cases.

### 3 Hybrid quantum linear solver algorithm

In this section, we will be focusing on modified versions of the original HHL algorithm and their implementations on IBM Quantum. From the previous section, we have learned that the subroutine, quantum phase estimation, involved in the algorithm responsible for matrix inversion is the most complex and expensive part. The hybrid quantum linear solver algorithms we will be looking will attempt to alter this subroutine in order to reduce the overall circuit complexity yielding results identical to those of the HHL algorithm.

Lee, Joo and Lee (2019) proposed a hybrid quantum algorithm for linear systems, referred to as Hybrid Reduced HHL algorithm, based on the HHL algorithm with the idea of removing unnecessary quantum part with prior classical information which can then be feed into a shortened reduced HHL circuit [12]. Based on Figure 4, the matrix $\hat{A}_\lambda$ and vector $b$ undergo a prior QPE algorithm which would yield an output that dictate the type of reduced ancillary quantum encoding (AQE) in the reduced HHL. The hybrid algorithm is supposed to solve specific linear equations in the form of,

$$\hat{A}_\lambda x = b$$

Consider that the hybrid HHL algorithm is applied to the linear equation, $\hat{A}_\lambda x = b$ where $\lambda = 1/4, 2/4, 3/4$ for 2-qubit register. By undergoing a repeated QPE algorithm beforehand, we are able to obtain a probability distribution of its measurement outcomes. The probability distribution can give insight to the eigenvalues of $\hat{A}_\lambda$ such that,

$$Pr(j) = \frac{1}{16} e^{6\pi i \lambda} (1 + e^{4\pi i \lambda})^2((-1)^j + e^{2\pi i \lambda})^2, \quad \text{for } j = 0, 1$$

From the probability distribution for the measurement outcome, it is also possible to obtain the knowledge that $\hat{A}_\lambda$ is perfectly 2-estimated where the matrices $\hat{A}_{1/4}$ and $\hat{A}_{3/4}$ have fixed eigenmean $\bar{m}_2 = 1$ and the matrix $\hat{A}_{2/4}$ has fixed eigenmeans $\bar{m}_1 = 1, \bar{m}_2 = 0$. Using the knowledge of eigenvalues and eigenmeans, they are able to implement a reduced AQE part.
for the reduced HHL algorithm. In the case of $\hat{A}_1/4$ and $\hat{A}_3/4$, the AQE parts to be used are controlled-unitary operation

$$|0\rangle_A |0\rangle_{r_1} \rightarrow (\sqrt{1-c^2} |0\rangle_A + c |1\rangle_A) |0\rangle_{r_1}$$

$$|0\rangle_A |1\rangle_{r_1} \rightarrow (\sqrt{1-c^2} |0\rangle_A + \frac{c}{3} |1\rangle_A) |1\rangle_{r_1}$$

and the AQE part for matrix $\hat{A}_2/4$ is given by a single-qubit unitary operation.

By performing the reduced HHL part based on the reduced AQE parts, the normalized solution of the linear equation in the qubit system $V$ is obtained. This hybrid algorithm is able to solve the linear equation given that the matrix $\hat{A}_3$ is perfectly n-estimated and has fixed eigenmeans. The experimental results for a 2x2 system on IBMQX4 is shown in Figure 5. The hybrid algorithm with reduced HHL was able to give a slightly better result compared to the original HHL as seen in Figure 5 where the probability obtained by reduced HHL is closer to the theoretical value compared to the original HHL.

Another quantum hybrid algorithm for linear systems of equations is proposed by Perelshtein et al. (2019) [13]. This hybrid algorithm, referred to as the Hybrid $U_C$ QLSA, has been implemented on a fairly large scale where it has been able to experimentally solve a $2^{17}$-dimensional problem on IBM Quantum which is a record as of yet. Based largely on the previous hybrid algorithm by Lee et al. (2020) where it takes the advantage of the fact that some bits of $A^{-1}$ eigenvalues could be the same for any eigenvector. The proposed hybrid algorithm alters the QPE part in HHL by introducing correcting single qubit gate $U_C$ which produces an exponential speedup instead of the original $U^{2j}$ [13].

Generally, this hybrid algorithm uses a lesser number of qubits for the phase register as compared to the original HHL algorithm which would make the circuit less complex as the width and depth of the circuit decreases. The construction of the operator $\hat{U} = e^{iA}$ is expressed as a tensor product of local operators. The physical operators $\hat{U}$ are divided into three groups:

**TP$_1$:** $\hat{U}_{TP_1}$ is the tensor product of single-qubit gates which does not entangle qubits.

**TP$_2$:** $\hat{U}_{TP_2}$ is the tensor product of two-qubit gates which leads to the emergence of entangled two-qubit clusters.

**NTP: ** $\hat{U}_{NTP}$ is not a tensor product of single-qubit or two-qubit gates which entangle all qubits.
Consider a 2x2 linear systems of equations, the paper tries to solve the system by investigating all operators of differing matrix types, \( \hat{U}_{TP_1} \), \( \hat{U}_{TP_2} \) and \( \hat{U}_{NTP} \) on four IBM-Q machines, Burlington (5 qubits), Yorktown (5 qubits), Melbourne (15 qubits) and Johannesburg (20 qubits). Consider the system of linear equations,

\[
\begin{pmatrix}
1 & 2 \\
-3 & 8
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2
\end{pmatrix}
=
\begin{pmatrix}
1 \\
0
\end{pmatrix}
\]

By running through the hybrid algorithm, the normalized solution is encoded into a pure single-qubit quantum state \( |\tilde{x}\rangle = \frac{4}{5} |0\rangle + \frac{3}{5} e^{-i\pi/2} |1\rangle \). The density matrix \( \rho \) obtained through full state tomography is shown in Figure 7(a) where the solutions are able to be approximated with high precision and fidelity. Figure 7(b) shows the fidelity of the algorithm with only NTP matrices, averaged over 70 NTP matrices for varying circuit widths. It shows relatively similar fidelity with respect to width of the circuit for the first three quantum processors with the same quantum volume, \( V_Q = 8 \) while the fidelity for Johannesburg with \( V_Q = 16 \) is slightly higher for corresponding widths.

The paper goes on to consider a large-scale implementation of the hybrid HHL algorithm. However, since full state tomography for a large-scale system requires \( 3^n \) experiments for n-qubit
circuit, the tomographical fidelity is not reliably attainable for this scale. A cross entropy that estimates the fidelity based on single Z-projective measurement which gives us the probability distribution of outcomes. The fidelity is defined as,

\[ F_{XEB} = \frac{\sum_{j=1}^{M} (p_{j}^e - p_{j}^{u} \cdot p_{j}^{t})}{\sum_{j=1}^{M} (p_{j}^{t} - p_{j}^{u} \cdot p_{j}^{t})} \]

where \( p_{j}^e \) and \( p_{j}^t \) are the probability distribution of outcomes for the noiseless implementation and experimental run of \( j \)th circuit respectively while \( p^{u} \) is the uniform probability distribution. This cross entropy fidelity shows the averaged proximity of the obtained projection of vector to the ideal solution. Another approach in a metric to measure the algorithm is the digital error model (DEM) which is essentially the product of the fidelity of readout errors, single-qubit gates and two-qubit gates. However, this metric failed in predicting the hybrid HHL algorithm performance due to correlation errors for experiments conducted on real IBM-Q machines.

Figure 9 shows that the measured fidelity is independent of the matrix types used in (b) and (d). While DEM can be useful in predicting simulations, it is not useful in predicting experimental results as gate errors are correlated. In this work, they are able to experimentally implement a quantum linear solver algorithm that finds the quantum state projection corresponding to the solution of a \( 2^{17} \times 2^{17} \) system of linear equations, which is as of now, a record in quantum linear solving problems for a gate-based quantum processing unit.

The third hybrid classical-quantum algorithm is devised specifically for Noisy Intermediate Scale Quantum (NISQ) machines by Chih-Chieh Chen et al. (2019) which outlines three main issues with existing HHL algorithms and its improvements: the input and output are in quantum states, loading data and \( \hat{A} \) is non trivial and expensive and it is inefficient in getting classical solution via tomography [14]. This hybrid algorithm, from now on will be referred as the Hybrid QW QLSA, is based on the quantum random walk (QRW) algorithm where only \( \hat{A} \) is encoded in quantum registers. Vectors \( \mathbf{b} \) and \( \mathbf{x} \) are classical and hence it solves one of the big contending issues in the complexity of loading and unloading data into quantum states. \( \hat{A} \) is encoded using Hamming cube structure – a square matrix of size \( N \) requiring \( O(\log N) \) qubits where QRW takes \( O(\log N) \) to obtain a component of \( \mathbf{x} \).

The Hybrid QW QLSA considers the usual linear system of equations in the form of \( \hat{A}\mathbf{x} = \mathbf{b} \) where \( \hat{A} \) is rewritten as a combination of identity matrix and a Markov chain transition matrix with a variable \( \gamma \), \( \hat{A} = \hat{I} - \gamma \hat{P} \). We will not go deeper into the intricacies of this algorithm as it is non-trivial and not of our focus in this study. As in Figure, since the algorithm is able to
Figure 8: (a) The fidelity obtained on a simulator based on the noise profile of Melbourne for 3 different matrix type. (b) The experimental results measured on Melbourne. The same type of measurements are done on Johannesburg in (c) & (d). [13]

yield the classical solutions of the linear equations, they are able to observe the relative error of the solutions which generally goes lower as the sampling number increases.

Figure 9: Relative error of the classical solution with respect to the prediction as a function of the sampling number on simulators and IBM-Q Tokyo. [14]

Across the three implementations of hybrid QLSA, all of them are able to output the solution
\(|x\rangle\) to varying degrees. Only the third implementation, Hybrid QW QLSA via quantum walk was able to retrieve the classical solution for the linear systems whereas the others would need a completely different algorithm, state tomography algorithm, to obtain their classical solutions. It would be inaccurate to compare them side by side as each of the algorithms differ quite significantly in its methodology and no definite big \(O\) notation and complexity are given in the papers. However, we are most interested in the first two hybrid algorithm, Hybrid Reduced HHL algorithm and Hybrid \(U_C\) QLSA which are based on the HHL algorithm. The Hybrid Reduced HHL algorithm by Lee et al. (2020) are devised and implemented for small scale implementation while the Hybrid \(U_C\) QLSA has the capability to be applied to larger scale problems nonetheless, only the former documented a concise implementation of the algorithm that will be of use for us in our study.

4 Variational quantum linear solver (VQLS) algorithm

Recently, a variational approach for solving quantum linear system problems proposed by Bravo-Prieto et al. (2020) has been demonstrated to perform well in noisy intermediate-scale quantum (NISQ) computers up to 1024 \(\times\) 1024 problem size (10 qubits) [15]. The Variational Quantum Linear Solver (VQLS) utilizes variational quantum eigensolver to solve linear equations. The output yielded by this algorithm is similar to the original HHL algorithm however the algorithm takes the advantage of its variational nature allowing it to be executed on current NISQ quantum machines and reducing the complexity of the algorithm by utilizing classical optimization.

Based on Figure 10, the input of the algorithm is the linear combination of unitaries \(A_l\) and a quantum circuit, \(U\) which prepares the state \(|b\rangle\). The cost function, \(C(\alpha)\) will be estimated via an ansatz \(V(\alpha)\) with parameters \(\alpha\) in a quantum-classical loop. The cost function, \(C(\alpha)\) quantifies how much the component \(\hat{A} |x\rangle\) is orthogonal to \(|b\rangle\). The value of the cost function is then returned to the classical computer to adjust a new value of \(\alpha\) via optimization algorithm to continue reducing the cost. Once the cost reaches the desired threshold, the optimal parameters \(\alpha_{opt}\) will be passed through the ansatz, \(V(\alpha_{opt})\) to form \(|x\rangle\).

There are two reasonable cost functions discussed in the paper – the global cost function \(C_G\) and the local cost function \(C_L\). However, the global cost functions show a plateau as the number of qubits \(n\) increases. Figure shows that as \(n\) increases, it gets harder for the global cost function to optimize. The local cost function for 50 qubits shows significantly better result in its optimization hence making local cost functions better for large scale implementations. The ansatz \(V(\alpha)\) are chosen to be a fixed hardware ansatz for convenience as the structure of the ansatz is fixed and the optimization is done over the varying \(\alpha\).

Figure 12 shows the implementation on Rigetti’s quantum hardware for a problem up to 1024 \(\times\) 1024 size. As observed the cost function obtained by training in the quantum computer are very close to the values obtained from a simulator. The value of the cost function approaches
near zero value which indicates a good solution to the linear system however due to noise present in the quantum computer the cost does not go to zero. To guarantee a precision of $\epsilon$, the cost must be lowered as the condition number increases. This will require more iterations of the variational hybrid classical-quantum loop to achieve. Essentially, this variational approach replace the gate complexity of the usual approach in exchange for the number of iterations for a fixed circuit depth. Since this is essentially a different algorithm altogether from the HHL algorithm, we will not go much further than this.

Figure 12: Implementation of VQLS on Rigetti’s quantum hardware. The global cost function $C_G$ is plotted against the number of optimization steps. [15]

5 Conclusion

In this report, we have looked at the difference between the general quantum linear solver algorithm, hybrid quantum linear solver algorithms and variational quantum linear solver (VQLS) algorithm and their implementations on circuit based quantum computers. This review acts as a precursor in finding a reliable algorithm for linear solving that could be implemented on a large scale and using NISQ machines for further works in implementing least squares data fitting algorithm that will be largely based on some variation of QLSA.
Among the implementations that we have analysed, our main focus is on the algorithms that are based on the original HHL algorithm – the general QLSA and hybrid QLSA. The latter shows a lot more usability in that one of the hybrid algorithm, the Hybrid $U_C$ QLSA are able to be implemented for large scale problems as opposed to the more general QLSA which are limited by the circuit complexity and current quantum computers. Of the hybrid algorithms mentioned in Section 3, the first two, the Hybrid Reduced HHL algorithm and Hybrid $U_C$ QLSA could prove to be useful provided we are able to implement them. While VQLS has proved to be able to be executed on current NISQ machines, the sheer difference between its variational approach to the conventional HHL algorithm has made it difficult for us to consider them as most of the data fitting algorithms we have looked at are largely based on HHL algorithm. However, VQLS is still highly reliable on its own for linear solving problems.

The problem here on lies in the reconstruction of the algorithms in IBM-Q as there are limited documentation and guides given in the literature for the hybrid algorithms. Reconstructing and implementing these hybrid algorithms by ourselves would be beneficial in helping us to understand the subject matter and will help in preparing for our next step in understanding implementations of least squares and linear regression algorithms in circuit based quantum computation.

Acknowledgements

This work is related to the ongoing research by Ts. Dr. Bahari Idrus and Assoc. Professor Dr. Mohammad Khatim Hasan funded by GUP-2020-061.

References


