Quantum-Assisted Graph Clustering and Quadratic Unconstrained D-ary Optimisation

Sayantan Pramanik and M Girish Chandra

TCS Research and Innovation
India
sayantan.pramanik, m.gchandra@tcs.com

Abstract—Of late, we are witnessing spectacular developments in Quantum Information Processing with the availability of Noisy Intermediate-Scale Quantum devices of different architectures and various software development kits to work on quantum algorithms. Different problems, which are hard to solve by classical computation, but can be sped up (significantly in some cases) are also being populated. Leveraging these aspects, this paper examines unsupervised graph clustering by quantum algorithms or, more precisely, quantum-assisted algorithms. By carefully examining the two cluster Max-Cut problem within the framework of quantum Ising model, an extension has been worked out for Max 3-Cut with the identification of an appropriate Hamiltonian. Representative results, after carrying out extensive simulation studies, have been provided including a suggestion for possible futuristic implementation with qutrit devices. Further extrapolation to more than 3 classes, which can be handled by qudits, has also been touched upon with some preliminary observations; quantum-assisted solving of Quadratic Unconstrained D-ary Optimisation is arrived at within this context. The paper also demonstrates how quantum description/formulation can sometimes lead to a different perspective and way of solving problems by providing the results for subgraph identification in graphs.

Index Terms—Ising model, graph Clustering, hamiltonian, quantum annealing, quantum approximate optimisation algorithm, qudits.

I. INTRODUCTION

We are progressing through an exciting period in Quantum Technologies and with the small-scale commercial quantum computers becoming increasingly available [1], Quantum Information Processing is witnessing spectacular developments. Before quantum processors become scalable devices capable of error correction and universality [2], the current and near-term devices, referred to as the Noisy Intermediate-Scale Quantum (NISQ) [2] devices are getting explored for solving certain hard problems to achieve significant speedups over the best known classical algorithms [4]. Promising results are already reported for solutions in the areas of optimisation, chemistry, machine learning, among others. Needless to say, hybrid quantum algorithms which use both classical and quantum resources to solve potentially difficult problems [5] are worked out and put into action.

It has been brought out that unsupervised machine learning and the associated optimisation strategies can be elegantly handled by quantum or hybrid quantum algorithms. In this paper, we consider clustering, an important unsupervised task. Clustering consists of assigning labels to elements of a dataset based only on how similar they are to each other - like objects will have the same label, unlike objects will have different labels [6]. In order to represent dissimilarity (or similarity), we need to define a distance measure between two data samples [6]. The distance between every possible pair of data samples can be captured in a matrix. This matrix can be interpreted as an adjacency matrix of a graph, where each vertex or node represents an element of data set and the the weight of edge between vertices is the corresponding distance [6]. In clustering, the main assumption is that distant points belong to different clusters; hence maximizing the overall sum of all weights (distances) between nodes with different labels represents a natural clustering algorithm for two-cluster case [6]. The mathematical formulation of this is a well known Maximum-Cut (Max-cut) problem and it can be easily translated to an optimisation objective [6]. The Max-cut problem is an example of the class of NP-complete problems, which are notoriously hard to solve. Many other combinatorial problems can be reduced to Max-cut, e.g., machine scheduling, computer-aided design, traffic message management problems, image recognition, Quadratic Unconstrained Binary Optimisation Problems (QUBO) [7] and many more [6]. One approach to solving Max-Cut is to construct a physical system typically a set of interacting spin - particles (two state particles) whose lowest energy state encodes the solution to the problem, so that solving the problem is equivalent to finding the ground state of the system [6].

Two main approaches have been identified to find the ground state of interacting spin systems (quantum optimisation) in NISQs [2], [6]: Quantum Annealing (QA) and Quantum Approximate Optimisation Algorithms (QAOA) [8]. QA is a form of analog computation that has been developed theoretically in the early nineties but realized experimentally in a programmable device only in 2011 by D-Wave Systems. QAOA, invented in 2014 and recently generalised for constrained combinatorial optimisation, requires digital gate-model quantum computing; it can be seen in some parameter range as a digitised version of QA [2].

In this paper, our starting point is a graph and we examine clustering on this abstraction (the graph itself can be constructed from the data points as cursorily mentioned in the beginning). To start with, we briefly touch upon the 2-cluster
Max-Cut problem in terms of the usual Ising model of interacting spins, but report some additional results/observations related to graph components (independent subgraphs). Then, we propose a simple way to extend the strategy to address 3-cluster problem on graphs. The requisite 3-state particles interaction and the associated Hamiltonian are brought out. Apart from providing typical results, remarks on how to go about implementations are also made, including on the hypothetical qutrit computing device. Extrapolation to more than three cluster case involving qutdis is also suggested, culminating in the Quadratic Unconstrained $D$-ary Optimisation (QUDO).

The paper is organized as follows: In Section \[\text{VI}\] Ising model and the Max-Cut clustering is presented; the results related to identifying independently partitionable subgraphs is delegated to Section \[\text{VII}\]. Clustering into 3 classes is covered in detail in Section \[\text{II}\]. Graph clustering into d classes is brought out in Section \[\text{IV}\]. In both Section \[\text{III}\] and Section \[\text{V}\] relevant results and differentiation with respect to some existing literature and possible implementations are interspersed. Initial remarks related to implementation are provided in Section \[\text{V}\]. Conclusions are provided in Section \[\text{VII}\].

II. TWO-GROUP CLUSTERING MAX-CUT PROBLEM

As mentioned in the previous section, one way to solve the two-cluster graph maxcut problem is to have a model of two-state interacting particles and solve for the lowest energy state. This interaction model (for spins) is the Ising model, originally developed to describe ferromagnetism, but subsequently extended to more problems \[\text{[2]}\].

A. Ising Model

The Ising model can be formulated on any graph as follows: consider an undirected graph $G = (V, E)$, where $V = \{v_1, \ldots, v_N\}$ is a set of $N$ sites, and $E$ is a set of edges representing the interactions between these sites. Every site $i$ has a corresponding spin variable $s_i$ \[\text{[10]}\]. These spins are binary-valued, taking values $\pm 1$ for up or $-1$ for down. Two spins $s_i$ and $s_j$ may interact with each other \[\text{[10]}\]. The energy of such an interaction depends on whether the values of the participating spins are the same or different: it is given by $J_{ij} s_i s_j$, where $J_{ij}$ is the strength of the interaction \[\text{[10]}\]. For each pair of interacting spins $s_i$ and $s_j$ (i.e., $J_{ij} \neq 0$), there exists a corresponding edge $(i, j) \in E$. The state of the model, $\sigma$, is an assignment of all $N$ variables $s_i$, $1 \leq i \leq N$. The set of all possible configurations is $\xi = \{-1, 1\}^N$ \[\text{[10]}\]. As well as pair-wise interactions, there can also be an external field that affects each site $i$ with energy $h_i s_i$. Thus, in the general case, the energy of a configuration $\sigma \in \xi$ is given by the so-called Edwards-Anderson Hamiltonian \[\text{[10]}\]:

\[
H(\sigma) = \sum_{(i,j) \in E} J_{ij} s_i s_j + \sum_{i \in V} h_i s_i \tag{1}
\]

When $h_i = 0 \ \forall \ i \in V$, the system is said to have no external field (also called zero magnetic field condition), in which case the energy of the configuration $\sigma$ becomes:

\[
H(\sigma) = \sum_{(i,j) \in E} J_{ij} s_i s_j \tag{2}
\]

In this paper, we mostly consider the Hamiltonian with zero external field. The system prefers lower energy states, i.e., those $\sigma$ that minimise $H(\sigma)$. An important task is to find configurations that minimise the energy of the system; such a configuration is known as ground state.

In order to obtain the Quantum Mechanical description of the Ising model of Equation \[\text{(1)}\] and Equation \[\text{(2)}\], one has to replace each $s_i$ by the Pauli-$Z$ matrix given by $\sigma_z^i$. As the state corresponding to the quantum mechanical interaction of one or more particles is given by the tensor products of the corresponding individual states, the $\sigma_z^i \sigma_z^j$ terms in $H(\sigma)$ denote the tensor product between $\sigma_z^i$ and $\sigma_z^j$, where $\sigma$ is the cumulative spin configuration of the complete system. It must be noted that these terms need to be appropriately constructed through tensor products of $\sigma_z$ and $I$ matrices, as discussed in subsection \[\text{II-B}\] to capture the pair-wise interaction between the $i^{th}$ and $j^{th}$ spins. With these considerations in mind, the Equation \[\text{(1)}\] and Equation \[\text{(2)}\] convert to \[\text{(11)}\]:

\[
H(\sigma) = \sum_{ij} J_{ij} \sigma_z^i \sigma_z^j + \sum_i h_i \sigma_z^i \tag{3}
\]

\[
H(\sigma) = \sum_{ij} J_{ij} \sigma_z^i \sigma_z^j \tag{4}
\]

B. Max-Cut Problem Using Ising Model

Little more formally, the Max-Cut problem starts with an undirected graph $G(V, E)$ with a set of vertices $V$ and a set of edges $E$ between the vertices. The weight $w_{ij}$ of an edge between vertices $i$ and $j$ is a positive real number, with $w_{ij} = 0$ if there is no edge between them. A cut is a set of edges that separates the vertices $V$ into two disjoint sets $V_1$ and $V_2$, such that $V_1 \subseteq V$ and $V_2 = V \setminus V_1$, and the cost of a cut is defined as the sum of all weights of edges connecting vertices in $V_1$ with vertices in $V_2$. One can connect the cut to the Ising model by identifying the vertices with $s_i$ and $w_{ij}$ with $J_{ij} s_i s_j$; $s_i = 1$ suggesting that node $i$ belongs to $V_1$, and $s_i = -1$ corresponds to $V_2$ (of course, $V_1$ and $V_2$ can be interchanged). The cost of the cut can, in fact, be written in terms of the objective function:

\[
C = \sum_{ij} \frac{w_{ij}}{2} (1 - s_i s_j) \tag{5}
\]

The Max-Cut problem aims at partitioning the nodes in such a way that the cost of the resulting cut is maximised. On comparing Equation \[\text{(2)}\] and Equation \[\text{(5)}\], it must be noted that the sign of the $s_i s_j$ term changes. Thus, the Max-Cut then boils down to finding the lowest energy state, instead of the maximum, of Ising model with $J_{ij} = w_{ij}$:

\[
\max C = \min H(\sigma) \tag{6}
\]

by noting that scaling the cost function by a constant multiplicative factor does not change the optimal solutions. As remarked earlier, the Max-Cut problem is equivalent to QUBO formulation where the two-state variable is 0 or 1 binary valued instead of 1 and $-1$; if the QUBO variables are denoted
by $x_i$, the two formulations are related by $s_i = 2x_i - 1$ \[7\].

The classical Ising Max-Cut can be extended to the quantum framework by replacing $s_i$ by $\sigma_i^z$, as discussed in Section II-A. This gives us the final quantum Ising Hamiltonian, given by Equation \[7\], which needs to be minimised to get the optimum cut \[12\]. In carrying out the optimisation based on Equation \[7\] below, if two adjacent nodes fall into the same cluster, then $w_{ij}$ is added to the cost function, else it is subtracted.

$$H(\sigma) = \sum_{ij} w_{ij}\sigma_i^z\sigma_j^z$$ \[7\]

As discussed in the previous subsection, the $\sigma_i^z\sigma_j^z$ terms represent the interaction of the nodes through the use of tensor products, as shown in the following example. Let us consider a graph having 5 vertices. If there is an edge between the nodes $v_1$ and $v_3$ having weight $w_{13}$, then the interaction between the respective vertices is given by:

$$H_{13} = w_{13}(I \otimes I \otimes \sigma_z \otimes I \otimes I)$$ \[8\]

The Hamiltonian thus formed is always diagonal, as there are no $\sigma_x$ terms involved, owing to the tensor product between diagonal matrices.

The cost Hamiltonian $H(\sigma)$ of the graph $G$, being diagonal, has orthogonal eigenvectors that form a complete standard basis. The system settles to the state having the lowest energy and the eigenvector corresponding to it is given by a unit vector along a standard basis state. The least-energy eigenvector is a $2^N \times 1$ vector, whose ket representation gives us an $N$-length bit-string. The nodes of the graph, $G$, are labelled as 0 or 1, according to the digits in the bit-string, with the most significant bit representing the label of the first vertex. This gives us a binary-clustered graph. The result of partitioning a graph with the given algorithm has been shown in Figure 1 \[12\] Figure 1(b). The minimum energy state corresponding to the partition is given by $|010101100101\rangle$.

For different and more complex cost functions, the transverse Ising Model may be used. If the Hamiltonian is not diagonal, the matrix can be diagonalised with the eigenvectors obtained from it. The standard eigenvector of the diagonalised Hamiltonian, corresponding to the least eigenvalue, is then used to partition the graph, as stated before.

**III. SOLVING MAX 3-CUT USING ISING MODEL**

In the direction of arriving at the quantum-assisted solution for Max 3-cut problem, few observations are put in place. Consider the nodes 2 and 3 of a graph with 5 nodes, shown in Figure 2 for binary clustering:

The cost for the connection, in the Ising model, is given by:

$$H_{23} = 5 \times (I \otimes I \otimes \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \otimes I \otimes I)$$ \[9\]

The $4 \times 4$ matrix, in the equation above, incorporates all the possible cluster combinations of the nodes 2 and 3, which are represented along the diagonal of the matrix. If the two cluster are named 0 and 1, then their combinations and energies are shown in Figure 3.

The energy is 1 when both the nodes are classified into the same cluster, $|00\rangle$ or $|11\rangle$, and is $-1$ otherwise. The system will settle for the lower energy state, $-1$, thus providing the optimum clustering.

The idea can be extended to clustering the nodes of a graph into 3 classes, with class labels 0, 1 and 2. The interaction-matrix between adjacent nodes of the graph should ideally
Thus, the Hamiltonian, being an observable, must be hermitian. This can be taken care of by modelling the interaction between two nodes as the tensor product between $\Omega_3$ and $\Omega^\dagger_3$ and taking only the real part of the elements of the resultant matrix. The final form of the interaction is given as:

$$H_{23} = 5 \times I \otimes I \otimes \frac{1}{2} (\Omega_3^2 \otimes \Omega_3^{3^\dagger} + \Omega_3^{2^\dagger} \otimes \Omega_3^3) \otimes I \otimes I \quad (14)$$

It is interesting to note that this is completely analogous to the $k = 2$ case, where the Pauli-$Z$ matrix is Hermitian and thus, $\sigma_z \otimes \sigma_z^\dagger = \sigma_z^\dagger \otimes \sigma_z$.

The term $\tilde{H}_{23} = \frac{1}{2}(\Omega_3^2 \otimes \Omega_3^{3^\dagger} + \Omega_3^{2^\dagger} \otimes \Omega_3^3)$ evaluates to:

$$\tilde{H}_{23} = \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & -0.5 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & -0.5 & 0 \\ 0 & 0 & \ldots & 0 & 1 \end{bmatrix} \quad (15)$$

If adjacent nodes are placed in dissimilar clusters, the interaction energy is $-0.5$, and 1 for similar clusters. This is in contrast to the desirable energy values of $-1$ and $1$ for dissimilar and similar clusters, respectively. But this does not have any effect on the clustering result since the energy for adjacent vertices in different clusters is still lower than that for similar clusters, and the former will be energetically favoured.

To provide an easy visualisation, the cluster classes can be modelled along the cube roots of unity as shown in Figure 5. The requirement for such a visualisation will become evident in the next section.

In the binary case, the nodes were represented by $2 \times 2$ Pauli-$Z$ matrices, the tensor product of which gave the required Hamiltonian matrix. Thus, for the 3-class problem, we need $3 \times 3$ matrices for each node, whose tensor product with another such matrix can give the requisite $9 \times 9$ Hamiltonian.

One can think of using a matrix with the cube roots of unity placed along the diagonal of the $3 \times 3$ matrix [14]:

$$\Omega_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{2\pi i/3} & 0 \\ 0 & 0 & e^{4\pi i/3} \end{bmatrix} \quad (11)$$

In an interaction between two nodes, the first node is represented by $\Omega_3$ and the second node is represented by the complex conjugate transpose of $\Omega_3$, $\Omega_3^\dagger$. The 3 in subscript signifies that the cube roots of unity are used to form the matrix. Thus, the energy of interaction of nodes 2 and 3 is:

$$H_{23} = \ldots \otimes \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & e^{2\pi i/3} & 0 & 0 & 0 & 0 \\ 0 & 0 & e^{4\pi i/3} & 0 & 0 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & e^{-2\pi i/3} & 0 & 0 & 0 & 0 \\ 0 & 0 & e^{-4\pi i/3} & 0 & 0 & 0 \end{bmatrix} \otimes \ldots \quad (12)$$

$$\Rightarrow H_{23} = \ldots \otimes \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & e^{-2\pi i/3} & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & e^{2\pi i/3} & 0 \\ 0 & 0 & \ldots & 0 & e^{4\pi i/3} \end{bmatrix} \otimes \ldots \quad (13)$$

But, the Hamiltonian, being an observable, must be hermitian. This can be taken care of by modelling the interaction between two nodes as the tensor product between $\Omega_3$ and $\Omega_3^\dagger$ and taking only the real part of the elements of the resultant matrix. The final form of the interaction is given as:

$$H_{23} = 5 \times I \otimes I \otimes \frac{1}{2} (\Omega_3^2 \otimes \Omega_3^{3^\dagger} + \Omega_3^{2^\dagger} \otimes \Omega_3^3) \otimes I \otimes I \quad (14)$$

It is interesting to note that this is completely analogous to the $k = 2$ case, where the Pauli-$Z$ matrix is Hermitian and thus, $\sigma_z \otimes \sigma_z^\dagger = \sigma_z^\dagger \otimes \sigma_z$.

The term $\tilde{H}_{23} = \frac{1}{2}(\Omega_3^2 \otimes \Omega_3^{3^\dagger} + \Omega_3^{2^\dagger} \otimes \Omega_3^3)$ evaluates to:

$$\tilde{H}_{23} = \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & -0.5 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & -0.5 & 0 \\ 0 & 0 & \ldots & 0 & 1 \end{bmatrix} \quad (15)$$

If adjacent nodes are placed in dissimilar clusters, the interaction energy is $-0.5$, and 1 for similar clusters. This is in contrast to the desirable energy values of $-1$ and $1$ for dissimilar and similar clusters, respectively. But this does not have any effect on the clustering result since the energy for adjacent vertices in different clusters is still lower than that for similar clusters, and the former will be energetically favoured.

To provide an easy visualisation, the cluster classes can be modelled along the cube roots of unity as shown in Figure 5. The requirement for such a visualisation will become evident in the next section.

In the binary case, the nodes were represented by $2 \times 2$ Pauli-$Z$ matrices, the tensor product of which gave the required Hamiltonian matrix. Thus, for the 3-class problem, we need $3 \times 3$ matrices for each node, whose tensor product with another such matrix can give the requisite $9 \times 9$ Hamiltonian.

One can think of using a matrix with the cube roots of unity placed along the diagonal of the $3 \times 3$ matrix [14]:

$$\Omega_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{2\pi i/3} & 0 \\ 0 & 0 & e^{4\pi i/3} \end{bmatrix} \quad (11)$$

In an interaction between two nodes, the first node is represented by $\Omega_3$ and the second node is represented by the complex conjugate transpose of $\Omega_3$, $\Omega_3^\dagger$. The 3 in subscript signifies that the cube roots of unity are used to form the matrix. Thus, the energy of interaction of nodes 2 and 3 is:

$$H_{23} = \ldots \otimes \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & e^{2\pi i/3} & 0 & 0 & 0 & 0 \\ 0 & 0 & e^{4\pi i/3} & 0 & 0 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & e^{-2\pi i/3} & 0 & 0 & 0 & 0 \\ 0 & 0 & e^{-4\pi i/3} & 0 & 0 & 0 \end{bmatrix} \otimes \ldots \quad (12)$$

$$\Rightarrow H_{23} = \ldots \otimes \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & e^{-2\pi i/3} & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & e^{2\pi i/3} & 0 \\ 0 & 0 & \ldots & 0 & e^{4\pi i/3} \end{bmatrix} \otimes \ldots \quad (13)$$

But, the Hamiltonian, being an observable, must be hermitian. This can be taken care of by modelling the interaction between two nodes as the tensor product between $\Omega_3$ and $\Omega_3^\dagger$ and taking only the real part of the elements of the resultant matrix. The final form of the interaction is given as:

$$H_{23} = 5 \times I \otimes I \otimes \frac{1}{2} (\Omega_3^2 \otimes \Omega_3^{3^\dagger} + \Omega_3^{2^\dagger} \otimes \Omega_3^3) \otimes I \otimes I \quad (14)$$

It is interesting to note that this is completely analogous to the $k = 2$ case, where the Pauli-$Z$ matrix is Hermitian and thus, $\sigma_z \otimes \sigma_z^\dagger = \sigma_z^\dagger \otimes \sigma_z$.

The term $\tilde{H}_{23} = \frac{1}{2}(\Omega_3^2 \otimes \Omega_3^{3^\dagger} + \Omega_3^{2^\dagger} \otimes \Omega_3^3)$ evaluates to:

$$\tilde{H}_{23} = \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & -0.5 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & -0.5 & 0 \\ 0 & 0 & \ldots & 0 & 1 \end{bmatrix} \quad (15)$$

If adjacent nodes are placed in dissimilar clusters, the interaction energy is $-0.5$, and 1 for similar clusters. This is in contrast to the desirable energy values of $-1$ and $1$ for dissimilar and similar clusters, respectively. But this does not have any effect on the clustering result since the energy for adjacent vertices in different clusters is still lower than that for similar clusters, and the former will be energetically favoured.

To provide an easy visualisation, the cluster classes can be modelled along the cube roots of unity as shown in Figure 5. The requirement for such a visualisation will become evident in the next section.
traditionally been known as Max $d$-Cut \[15\]. Before moving onto the problem, it is necessary to introduce some additional mathematical machinery to facilitate the solution.

In section III, the matrix $\Omega_3$ can be recognised as the clock matrix for three dimensions. This is not surprising as the clock and shift matrices, $U_d$ and $V_d$, have been used to generalise the Pauli-$Z$ and $X$ matrices \[16\], respectively, for $d$ dimensional qudits \[17\]. These matrices, which are zero-trace and symmetrical, but not hermitian, can be utilised to solve $d$-ary optimisation problems where the variables can take $d$ number of discrete values, similar to a $d$-level system. In this section we extend the idea of Quadratic Unconstrained Binary Optimisation, QUBO, to such $d$-ary problems, and dub the technique as Quadratic Unconstrained $D$-ary Optimisation, QUDO. The Ising Hamiltonians for such problems can be given by:

\[
H = \sum_{ij} \frac{w_{ij}}{2} (U_d^i U_d^j + U_d^j U_d^i)
\]  \hspace{1cm} (17)

where $U$ is the clock matrix in $d$ dimensions and the subscript refers to the number of levels in the optimisation problem, and are defined as in Equations \[18\] and \[19\] \[18\]. Here, again, more complicated Hamiltonians can be modelled via the interaction of $U_d$ and $V_d$ terms, that lead to non-diagonal Hamiltonian matrices.

\[
V_d = \sum_{j=0}^{d-1} |(j+1) \text{ mod } d\rangle \langle j| 
\]  \hspace{1cm} (18)

\[
U_d = \sum_{j=0}^{d-1} \omega^j |j\rangle \langle j| 
\]  \hspace{1cm} (19)

where $\omega = e^{2\pi i/d}$ is the $d^{th}$ root of unity \[18\].

Such an optimisation problem can be demonstrated via the $d$-class Max-Cut partitioning of a graph, called the Max $d$-Cut problem. The cost function (Hamiltonian) is taken as given in Equation \[17\]. For the $d = 4$ case, the interaction Hamiltonian matrix for two adjacent nodes has been shown in Figure 7. It should be noted that the $|aa\rangle$ elements are 1, while the $|ab\rangle$ elements are 0 or $-1$. This happens because there are two possible angles between the $4^{th}$ roots of unity, i.e., $\pi/2$ and $\pi$. If the angle between the classes is $\pi/2$, then the interaction term is 0, and it is $-1$ if the classes are $\pi$ radians apart. This means that having a larger angular difference between the classes is more favorable. The effect of clustering still remains the same, however. The system settles for a state that ensures the highest angular difference between the classes. This was not apparent for the $d = 2$ or $d = 3$ cases because there was only one possible angle between the classes.

For a $d$-cluster problem, if the classes are numbered from 0 to $(d - 1)$, then the interaction energy term between nodes of classes $a$ and $b$ is given by:

\[
z_a \cdot z_b = \frac{1}{2} (\bar{z}_a z_b + z_a \bar{z}_b)
\]  \hspace{1cm} (20)

where $z_a = e^{2\pi a/d}$ and $\bar{z}_a$ is the complex-conjugate of $z_a$.

Figure 8 shows the vector representation of the four classes along the fourth roots of unity. Using the clock matrices to replace the Pauli-$Z$ matrix, we arrive at results identical to the vector Potts model (also referred to as the clock or planar Potts model) \[19\], \[20\], where the spin states of a $d$ level system were distributed uniformly on a unit circle, at angles specified by:

\[
\Theta_n = 2\pi n/d, \quad n = 0, 1, \ldots, d - 1
\]  \hspace{1cm} (21)
As in Equation (20), the interaction energy depends on the relative angle between the two vectors:

$$H = -\sum_{ij} J(\Theta_{ij})$$  \hspace{1cm} (22)$$

where \( J(\Theta_{ij}) \) is \( 2\pi \) periodic, and \( \Theta_{ij} = \Theta_{ni} - \Theta_{nj} \) is the angle between the two spins at neighbouring sites \( i \) and \( j \). For the vector Potts model [20]:

$$J(\Theta) = -\epsilon_1 \cos \Theta$$  \hspace{1cm} (23)$$

Thus, the proposed method can be considered as a quantum mechanical extension of the vector Potts model. A typical result of quaternary clustering for a simple graph is shown in Figure 9. Again, extensive simulation studies confirmed the correctness of the methodology.

Additionally, for \( d = 4 \), an interesting behaviour was observed. For many graphs, the result was bipartite, i.e., the resultant partition had only two classes. The partitioning for such a graph has been shown in Figure 10. Other graphs had solution states for \( 2, 3 \) and \( 4 \) classes, all having the minimum energy eigenvalue. This might have been due to the fact that adjacent classes have \( 0 \) interaction energy. Further investigations are necessary for more than \( 4 \) classes and to arrive at the nature of the behaviour for general \( d \) classes.

V. PROPOSITION FOR FUTURE IMPLEMENTATION OF QUDO ALGORITHMS

With the NISQ devices already being available, it is really essential to consider the implementation of the quantum algorithms on the hardware. Few pointers are put forth in this section. The binary Max-Cut problem can be solved using quantum annealing such that the system settles to the final Hamiltonian given by the Ising function. The D-Wave annealer Hamiltonian may be represented as [21]:

$$H_{ising} = -\frac{A(s)}{2} \left( \sum_i \sigma_i^z \right) + \frac{B(s)}{2} \left( \sum_i h_i \sigma_i^z + \sum_{i>j} J_{ij} \sigma_i^z \sigma_j^z \right)$$  \hspace{1cm} (24)$$

The cost function can easily be mapped to the D-Wave Ising Hamiltonian and the minimum energy states can be obtained by appropriate sampling.

Alternatively, the QAOA approach can also be utilised to find the solution of binary Max-Cut problem by evolving the appropriately prepared wavefunction using unitary operators towards the Hamiltonian obtained by casting the problem into Ising model (or the equivalent QUBO) [22], [23]. Thus, QA and QAOA have formulation of the Ising model in common but they require different hardware. Ruslan Shaydulin, et al have compared the two approaches for a different, but similar, problems [24]. In QAOA, hybrid quantum and classical processing is utilised and the quantum computation is in terms of the gate-circuit model, as mentioned earlier [22], [25].

For Max \( d \)-Cut (or QUDO) problems, these approaches cannot be used directly. The existing quantum annealers are inherently binary in nature, due to the presence of \( 2 \times 2 \) Pauli matrices. Ushijima-Mwesigwa, et al [12], have proposed a concept of super nodes for Graph Partitioning into \( d \) classes. The problem formulation is quite similar to that of Max-Cut, and the same approach can be used for the latter. The drawback is that for a graph with \( N \) vertices, \( dN \) qubits are required to model the problem onto an annealer and the matrix blows up by a factor of \( d^2 \).

If useful, future annealers can be designed in such a way that they utilise Equation (17) as the final Hamiltonian, with the addition of cross terms containing \( U_d \) and \( V_d \). Specifically, one can think of annealers based on qutrits [26] to address Max 3-Cut based on Equation (16).
Moving on to QAOA for $d$ classes, to find the expectation value of a $d^N \times d^N$ cost operator using Equation (25), $d$ dimensional qudits may be used. The expectation value may then be minimised through the use of classical optimisation algorithms as described in the original QAOA paper [8].

$$\langle H \rangle = \langle \psi | H | \psi \rangle$$

As another possibility, $m$ qubits may be utilised to simulate $N$ qudits, where $m$ is given by [27]:

$$m = N(\log_2 d)$$

VI. ADDITIONAL RESULTS RELATED TO MAX-CUT

We have spelt out the usual quantum mechanical formulation to carry out Max-Cut in Section II-B. For some graphs, we may end up having more than one eigenstate for the same (lowest) energy, i.e., the state exhibits degeneracy. In this section, we carefully examine the degeneracy and bring out some novel observations and results in the context of graph partitioning. Specifically, we consider how degeneracy provides a different perspective towards the examination of certain subgraph identification in the quantum formulation. Some symmetries make it possible to cluster some of the vertices independently from the others, i.e., the partitioning of a subset of vertices has no effect on the partitioning of the remainder. Such sets of vertices, along with the edges connecting the vertices within the sets are referred to as independently partitionable subgraphs of the graph $G$ within the context of this work. We shall define this observation more rigorously and illustrate it with examples in the following paragraphs.

While carrying out this exercise, we have restricted ourselves to graphs having all edge weights equal to unity.

Let $G(V, E)$ be a graph with $V$ as the set of vertices, having cardinality $N$, and $E$ as the set of edges. Max-Cut partitioning allows us to cluster the set of vertices into two subsets of $V$, subject to minimising the cost function. It may happen that there are multiple solutions corresponding to the minimum cost. Then these degenerate solutions can be used to further partition the graphs vertices into $M$ sets, $\bar{V}_1, \bar{V}_2,$ and so on, where $\bar{V}_i \subseteq V$, and $M < N$. The clustering of the vertices in $\bar{V}_i$ does not influence the clustering of the vertices in any other set $\bar{V}_j$. The edges between the vertices of the set $\bar{V}_i$ forms the set $\bar{E}_i$, and the edges connecting vertices of $\bar{V}_i$ to that of set $\bar{V}_j$ are discarded, where $i \neq j$. Thus $\bar{V}_i$ and $\bar{E}_i$ together form a
graph \( \bar{G}_i \) which is a subgraph of the original graph \( G \). The set of 
these subgraphs can be called independently partitionable 
subgraphs of \( G \) as the Max-Cut clustering of \( \bar{G}_i \) has no effect 
on the clustering of \( \bar{G}_i \), given \( i \neq j \).

The dual of a binary number is given by converting the 
0s to 1s and 1s to 0s. Dual binary numbers, thus, represent 
equivalent partitions of the graph into two clusters, and will 
have the same energy or cost value. But if there exist solutions 
that are not duals of each other, and yet have the same 
clustering cost, then the graph can be said to have independent 
subgraphs within it. It can easily be seen that if the number 
of such independently partitionable subgraphs is \( L \), then there 
will be \( 2^L \) solutions with the minimum eigenvalue.

For example, if the minimum energy computational states 
are:

\[
\begin{align*}
0011010 & \rightarrow 1 \\
0011110 & \rightarrow 2 \\
1100001 & \rightarrow 3 \\
1100101 & \rightarrow 4
\end{align*}
\]

1 and 4 are duals of each other, as are 2 and 3. They 
represent the same partitions of the graph. Though 1 and 
2 have the same energy, they represent different partitions. 
Thus, it does not change the cost function whether node 4 is 
classified as into class 0 or 1. Similarly with 3 and 4.

The bitwise XOR of 1 and 2 gives 1 on the fifth place from 
the left. As the bits in the solution refer to the clusters that 
the nodes of the graph are placed in, solutions 1 and 2 place 
all the nodes in similar clusters, except for the fourth node. 
Thus, the node 4 can be said to be independent of the rest 
of the graph, as their respective clustering do not affect each 
onther, as well as the overall cost function. The graph then 
contains 2 independent subgraphs: \( G_1 \) and \( G_2 \), having vertex 
sets \( V_1 = \{0, 1, 2, 3, 5, 6\} \) and \( V_2 = \{4\} \), respectively.

Thus, the non-transverse Ising Hamiltonian gives the inde-
pendent and isolated subgraphs of a graph. A different cost 
function will provide different subgraphs corresponding to 
some other property represented by that cost function. This 
can be visualised by the graph shown in Figure 11a. On Max-
Cut partitioning the graph in 11a the graph 11b is obtained, 
with the nodes clustered into either the red or blue cluster.

The solutions with the minimum energy values are given 
by:

\[
\begin{align*}
0010110 & \rightarrow 1 \\
0110010 & \rightarrow 2 \\
0110011 & \rightarrow 3 \\
0110110 & \rightarrow 4 \\
1001001 & \rightarrow 5 \\
1001100 & \rightarrow 6 \\
1001101 & \rightarrow 7 \\
1101001 & \rightarrow 8
\end{align*}
\]

The solutions 5, 6, 7 and 8 are the duals of 4, 3, 2 and 1, 
respectively, and thus, do not incorporate any extra information 
about the partitioning of the system. The bitwise XOR of 1 
and 2 gives nodes 1 and 4; 1 and 3 gives 1, 4 and 6; 1 and 

![Fig. 11: Example graph to show independent subcomponents](image)

VII. CONCLUSIONS

The paper, by appropriately combining mathematical as-
pects and simulation, studies proposed quantum-assisted graph 
clustering for three or more clusters. Even though the presenta-
tion is heavily biased towards algorithmic aspects, pointers are 
provided for possible architectures for implementation. While 
firm footing appears to have been established for Max 3-Cut 
case, additional study for the case of more than 3 clusters 
is required to concretize the observations and results; aspects 
related to measurements also need to be looked into.
ACKNOWLEDGMENT

The authors sincerely thank Mr. Mahesh Rangarajan, Dr. Arpan Pal and Dr. Balamuralidhar P of TCS R&I for their support and encouragement.

REFERENCES

[1] L. Tse, P. Mountney, P. Klein, and S. Severini, “Graph cut segmentation methods revisited with a quantum algorithm,” arXiv e-prints, p. arXiv:1812.03050, Dec 2018.

[2] M. Kim, D. Venturelli, and K. Jamieson, “Leveraging quantum annealing for large mimo processing in centralized radio access networks,” in Proceedings of the ACM Special Interest Group on Data Communication, pp. 241–255, ACM, 2019.

[3] J. Preskill, “Quantum computing in the nisq era and beyond,” Quantum, vol. 2, p. 79, 2018.

[4] C. H. Bennett, E. Bernstein, G. Brassard, and U. Vazirani, “Strengths and Weaknesses of Quantum Computing,” arXiv e-prints, pp. quant-ph/9701001, Jan 1997.

[5] D. Zhu, N. M. Linke, M. Benedetti, K. A. Landsman, N. H. Nguyen, C. H. Alderete, A. Perdomo-Ortiz, N. Korda, A. Garfoot, C. Brecque, et al., “Training of quantum circuits on a hybrid quantum computer,” Science advances, vol. 5, no. 10, p. eaaw9918, 2019.

[6] J. Otterbach, R. Manenti, N. Alidoust, A. Bestwick, M. Block, B. Bloom, S. Caldwell, N. Didier, E. S. Fried, S. Hong, et al., “Unsupervised machine learning on a hybrid quantum computer,” arXiv preprint arXiv:1712.05771, 2017.

[7] M. Lewis and F. Glover, “Quadratic unconstrained binary optimization problem preprocessing: Theory and empirical analysis,” Networks, vol. 70, no. 2, pp. 79–97, 2017.

[8] E. Farhi, J. Goldstone, and S. Gutmann, “A quantum approximate optimization algorithm,” 2014.

[9] R. [https://physics.stackexchange.com/users/369/traskolnikov], “Ising model for dummies,” Physics Stack Exchange. URL:https://physics.stackexchange.com/q/1525 (version: 2010-12-01).

[10] D. Kamenetsky et al., “Ising graphical model,” 2010.

[11] A. Tan, “Quantum ising models,” 2018.

[12] H. Ushijima-Mwesigwa, C. F. Negre, and S. M. Mniszewski, “Graph partitioning using quantum annealing on the d-wave system,” in Proceedings of the Second International Workshop on Post Moore’s Era Supercomputing, pp. 22–29, ACM, 2017.

[13] T. Gowda, “Introduction to quantum optimization using D-Wave 2X,” 2018.

[14] D. Ghosh, T. Jennewein, and U. Sinha, “Entanglement certification and quantification in spatial-bin photonic qutrits,” arXiv preprint arXiv:1909.01367, 2019.

[15] A. Friese and M. Jernee, “Improved approximation algorithms for max k-cut and max bisection,” in Integer Programming and Combinatorial Optimization (E. Balas and J. Clausen, eds.), (Berlin, Heidelberg), pp. 1–13, Springer Berlin Heidelberg, 1995.

[16] S. Prakash, A. Jain, B. Kapur, and S. Seth, “Normal form for single-qutrit clifford+ t operators and synthesis of single-qutrit gates,” Physical Review A, vol. 98, no. 3, p. 032304, 2018.

[17] G. Grimmett, The Random-Cluster Model, ch. 10, pp. 320–340. Springer Berlin Heidelberg, 2006.

[18] F.-Y. Wu, “The potts model,” Reviews of modern physics, vol. 54, no. 1, p. 235, 1982.

[19] F.-Y. Wu, “The potts model,” Reviews of modern physics, vol. 54, no. 1, p. 235, 1982.

[20] D.-Y. Wu, “The potts model,” Reviews of modern physics, vol. 54, no. 1, p. 235, 1982.