MIT iQuHACK 2024 QuEra Challenge — The Qubit Questers

Matthew Marsicano, Adam Godel, Maximilian Rieger, Alexander Poulin, Archit Kalra

February 4, 2024

1 Defining the problem

The challenge posted by QuEra was to explore a special case of the "maximal independent set" problem. The original formulation of the problem includes a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and the task to find a maximal subset $\mathcal{M} \in \mathcal{V}$ of vertices, such that $\forall\, x_i, x_j \in \mathcal{M} :\ (x_i, x_j) \notin \mathcal{E}$. We consider the special case in which all vertices are placed on a regular square lattice spacing a and size $N \times N$. Furthermore, we remove 20% of the vertices from the lattice (selected at random). It has been demonstrated recently that if the edges connect nearest and next-nearest neighbors on the lattice, the problem of finding the maximal independent set (MIS) can be efficiently tackled by classical computational methods [Andrist et al., 2023]. Here we consider an interaction range that spans three lattice sites a, and finding the MIS is, therefore, arguably "harder" than in the initial problem. The task is to approach the problem by simulating it with trapped neutral atoms.

2 Theory

2.1 Overview

We consider the neutral atoms to be sufficiently described by a two level system of a ground state $|g_i\rangle$ and an excited state (Rydberg state) $|r_i\rangle$ (the atoms are labeled by indices i). Transitions between the ground state and the Rydberg state can be driven optically with a laser. Choosing appropriate values of frequency, phase, and pulse duration leads to an effective absorption/emission term $\sum_i \frac{\Omega(t)}{2}\left[e^{i\phi(t)}\,|g_i\rangle\langle r_i| + e^{-i\phi(t)}\,|r_i\rangle\langle g_i|\right]$ in the Hamiltonian. By detuning the frequency of the driving laser by $\Delta(t)$, one can control the energy gain/cost from adding excited atoms $-\Delta(t)\sum_i n_i$, where the on-site Rydberg density operator $n_i = |r_i\rangle\langle r_i|$ has been defined. What makes the neutral atoms a unique platform for quantum experiments is the dipole-dipole interaction between excited atoms. This is captured by the term $\sum_{i<j} V_{ij} n_i n_j$, where $V_{ij} = \frac{C_6}{|x_i - x_j|^6}$ and C_6 is a constant that depends on the choice of neutral atom. The Aquila quantum computer uses ^{87}Rb atoms with a value of $C_6 = 862690 \times 2\pi$ MHz μs^6. The Rydberg interaction has a significant impact on the excitability of atoms close to already excited atoms. Due to the interaction, the energy cost of exciting a neighboring atom is lifted, and the drive is thus off-resonance. This effect is called the Rydberg blockade. The total Hamiltonian can be summarized to be

$$H = \sum_i \frac{\Omega(t)}{2}\left[e^{i\phi(t)}\,|g_i\rangle\langle r_i| + e^{-i\phi(t)}\,|r_i\rangle\langle g_i|\right] - \Delta(t)\sum_i n_i + \sum_{i<j} V_{ij} n_i n_j. \tag{1}$$

with $\hbar = 1$.

2.2 $\mathbf{R_b/a}$

An important aspect to take into account is the Rydberg blockade mechanism. Effectively, when two or more atoms are within a close enough proximity to each other (within the Rydberg blockage radius R_b), there is a distinctive suppression of the chance of a double excitation (whereby both atoms in the proximity are excited). The interaction can be expressed as a Van der Waals interaction:

$$R_b = \sqrt{\frac{C_6}{(\Omega^2 + \Delta^2)^6}}. \tag{2}$$

In this equation, the relationship between Ω, the Rabi frequency, and Δ, the detuning factor, is of crucial importance when optimizing for a large Rydberg radius. By treating the Rabi frequency as negligible, the Rydberg radius equation was simplified to a single parameter, the detuning factor. We require a Rydberg radius to be ideally around the geometric mean of the distance across three atom separations and the closest atom outside of the three atom separation radius. Using the minimal separation distance of the Aquila system as $a = 4\mu m$, the final detuning parameter was estimated using the following equation:

$$\Delta \approx \frac{C_6}{(\sqrt{3\sqrt{10}}a)^6} \approx 1.55\,\text{MHz}. \tag{3}$$

From this point, we began to explore alterations to the waveform, such that the final detuning energy was steadily optimized towards. The linear waveform was chosen as an ideal baseline to begin the simulation in Julia. Through simulation and experimentation on quantum hardware, we concluded that it is plausible that the conditions of Aquila are capable of fitting $R_b/a = 3$, as the calculated Δ value is within the possible range of resolution of the hardware (-20 to 20 $\times$ 2π MHz). This is particularly interesting because the annealing hardness spikes significantly, peaking at 3, offering potential for quantum advantage.

3 Programming

3.1 Simulations with Julia

We began our programming process by simulating the MIS problem classically in Julia on small (4x4 and 5x5) graphs. Using Julia gave us a sense of how to model the problem and we were then able to begin optimizing our delta graph from the original linear shape to a shape better suited for finding the MIS more accurately. The advantage of using Julia was the ability to check our MIS simulation results using Generic Tensor Networks, giving us confidence with our optimizations before moving onto the actual quantum hardware.

3.2 Optimization

3.2.1 Nelder-Mead Optimization

We spent significant time considering how to optimize the time dependence of the detuning $\Delta(t)$ from the initial linear graph varying from $-\Delta_{\max}$ to $\Delta_{\max}$. We used the negative Rydberg density sum in conjunction with Julia's Optim optimizer to try to maximize the density. At first, we encountered the issue of overfitting with unusual and highly variable graph shapes, which we tried to resolve by using the average of an array of optimizing factors for 20 different 5x5 lattice graphs instead of individual optimization, which we found worked more consistently for classical simulations of lattice graphs of arbitrary classically computable size. For a more sophisticated approach, it would be necessary to optimize the time dependence of the parameters with respect to various lattice topologies, but there is only so much that you can do during a 24-hour Hackathon.

3.2.2 Bayesian Optimization

After running into an overfitting error in the implementation of the Nelder-Mead optimization scheme, the Bayesian optimization method became a worthy alternative to explore. A particular benefit of Bayesian Optimization is the fact it is suited towards a black box problem. The particular interaction between atoms and the detuning towards excited states is nearly pure chaos. To be able to determine the fitting waveform for delta to achieve a high fidelity for the MIS requires some reiterative learning model. Hence, Bayesian Optimization [Finžgar et al.].

The process in general is relatively simple. It begins with setting up initial parameters, bounded within some interval. An optimizer with the given parameter sets up a surrogate model of the black box function. Now, the initial system is ready for its initial iteration. A small set amount of shots is sent with the surrogate model of the delta function to retrieve the initial set of data points. Additional points are chosen by the upper confidence bound which is determined by the sum of the mean and the product of the standard deviation with a decay rate. After each iteration, a Gaussian process,

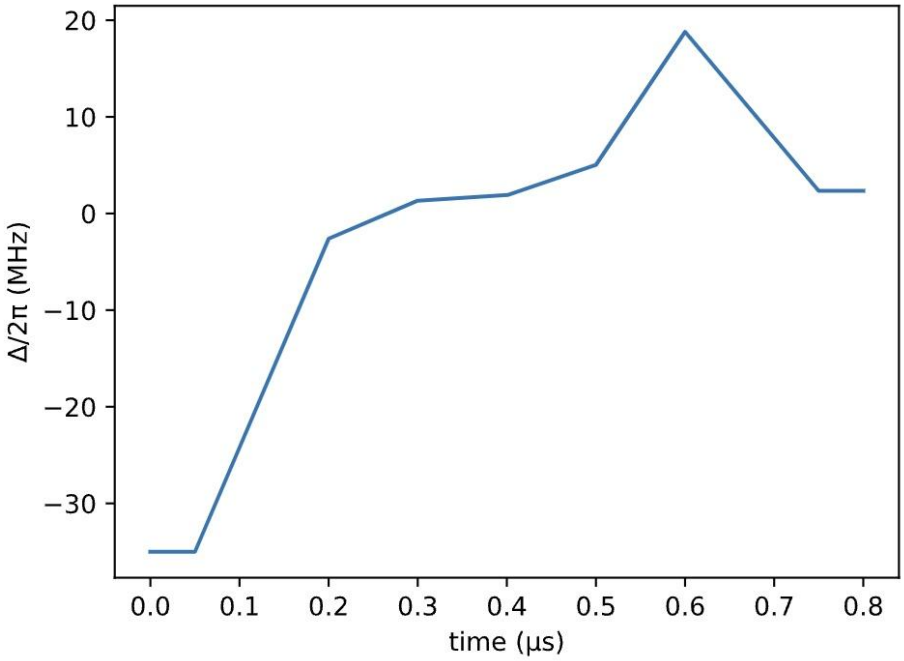

Figure 1: Piecewise Linear Graph for $\Delta/2\pi$ (MHz) vs. time (μs)

through common waveforms such as Bang-Bang, cubic, linear, and Fourier, determines more points so the uncertainty and decay decreases while the form, hopefully, tends toward the target function.

While the premise of the algorithm is accessible to learn, the actual implementation was a tough challenge. Having been exposed to Bayesian Optimization for the first time, an application of the process within several hours would have been a miracle. Nevertheless, work began in Python to implement as much of the algorithm as possible within the time constraint. In hindsight, Julia's platform and capabilities with Bloqade would have been a faster and easier to learn on the fly with. Still progress was made but a full implementation is still out of reach within such a short time frame. A basic implementation of the several of the schedule parametrizations was completed and work on the parameters had begun. To further realize this algorithm, the parametrization would need to be refined, the initial parameters would have to be chosen and possibly generalized depending on the matrix size and shape. Additionally, the upper confidence bound would have to be implemented so that the reiterative approach could work. While the Bayesian Optimization algorithm could not be tested for MIS during this Hackathon, it's results in the included reference provides great confidence in its capabilities and should be investigated further to fully test it's ability to create these blackbox waveforms.

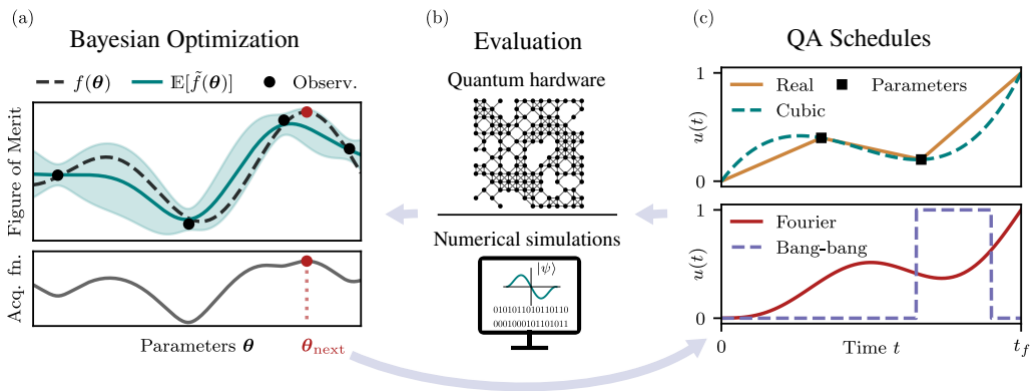

Figure 2: Overall illustration of the Bayesian Optimization Process [Finžgar et al.]

3.2.3 Other optimization techniques

A differential evolution algorithm using the BlackBoxOptim library in Julia as well as a reinforcement learning algorithm were written to help optimize the sum of the on-site Rydberg densities. These algorithms were not as successful as the classical algorithms in terms of optimization, and tended heavily toward corner solutions with parameters [-1, -1, -1, -1, 1]; this would lead to a rather steep and inaccurate piecewise linear function.

3.3 Quantum Hardware Results

Running the routine with parameters generated from the Nelder-Mead optimization on actual quantum hardware was of course the highlight of the Hackathon. The Aquila quantum computer allows for up to 100 runs within a cycle. In order to benchmark the results, we determined the MIS with the help of a classical Generic Tensor Network. The results are summarized in the following table:

Graph Size	MIS Accuracy	Excited State Count Accuracy	HP
4x4	21%	37%	0.54
5x5	2%	44%	3.0
9x9	0%	8%	3.44
10x10	0%	1%	—

(4)

The second column specifies the percentage of correctly determined MIS, and the third column the percentage of states that have exactly the same amount of excited states as an MIS, but might violate the "independence" condition.

3.4 Annealing Hardness Parameters

In combinatorial optimization problems, the difficulty of both classical and quantum annealing processes are quantized by annealing hardness parameters which reflect the annealing schedule, hardware specifications, and noise level of the system. As a result of the complexity of these interrelated conditions, calculating the hardness parameter in a generalizable way requires repeated experimentation via a tropical tensor network algorithm. In the process of building our tensor network algorithm to optimize for a possible hardness parameter, we realize that the time complexity would quickly scale out of control for larger lattices. This scaling issue was the product of the annealing hardness parameter calculation that we chose to match the equation below.

$$HP := \frac{N_{|MIS|-1}}{|MIS| \cdot N_{|MIS|}}$$

In this equation, N represents the degeneracy of the independent set configurations that have a size that is less than or equal to the optimal number of maximum independent sets. We have calculated the annealing hardness for the 4×4 and 5×5 lattice and listed the results in the above table.

4 Business

4.1 Biological Applications of MIS

There are a variety of biologically relevant applications of the maximal independent set problem. Chiefly, MIS finds great usage in computational biology, molecular simulations and drug design. In particular, MIS is integral to the reconnaissance metadynamics (RMD) technique for protein-ligand interaction analysis [Söderhjelm et al., 2012]. It allowed for analysis of protein structure by simulating each alpha carbon in any peptide chain as a vertex of a certain graph. In doing so, MIS allowed for the simulated separation of protein-ligand interactions such that there were a maximum number of surface interactions without interfering with the overall structure of the protein itself. With the introduction and adoption of new technologies including DeepMind's structure prediction software AlphaFold, the use of MIS remains a significant factor in the study of protein structure as well as protein-ligand and protein-protein interactions [Ahn et al., 2020].

A similarly significant application of MIS to biology includes its use in drug discovery. In particular, modern early drug discovery techniques heavily emphasize the principle of diversity selection,

to prevent a large number of potential molecules having similar domains (and thus redundant functionalities) [Meinl et al., 2011]. Though most diversity selection techniques incorporate the static structures of the molecules, this is an NP-hard problem. Meinl et al. pioneered a technique called Maximum-Score Diversity Selection, which additionally uses MIS to analyze the molecular dynamics of different molecules to determine if they have similar pharmacological interactions (and thus pre-screen them before the high-thoroughput screening universally conducted in drug discovery labs). Similar techniques have since become prevalent in computational drug design, with MIS and graph theory as a whole playing a critical role in unsupervised learning for pharmacophore discovery [Xu et al., 2023] Training times for such algorithms were as little as 10 to 30 seconds.

With specific regard to the use of MIS in drug discovery, it is important to note that some pharmacophores have a mandatory minimum diversity selection score; in order to achieve a score such as this, there may be a constraint on how far "happy" vertices have to be from each other [von Korff and Sander, 2022; Ghirardi and Salassa, 2022]. Indeed, the particular restriction $R_b/a = 3$ would be crucial in small molecules (molecular weight < 1 kilodalton) as opposed to larger molecules such as biologics; this is because two hypersimilar functional domains would not be as effective when directly next to each other, but may be able to exert greater pharmacological impact if they are farther away. This is especially seen with Proteolysis Targeting Chimeras (PROTACs), which are small molecules with heterodomains that are attached to each other with an ubiquitin ligase [Békés et al., 2022]. With a value of $R_b/a = 3$, the radius is not too large such that there is only one domain on the molecule, but also not too small such that there are extensively similar domains adjacent to each other, which would be heavily limiting on the efficiency of the molecule to degrade unwanted proteins in the body. The particular PROTACs designed by one of the authors of the aforementioned paper, Dr. Craig Crews, successfully launched Proteolix in 2003, which was acquired by Onyx Pharmaceuticals for $851 million, which itself was later acquired by Amgen.

4.2 Logistical Application of MIS

A further logistical application of MIS in solving real-world problems includes the use of Hopfield neural networks, especially in the context of wireless sensor networks [Serpen and Li, 2011]. In particular, Hopfield neural networks involve self-referential feedforward propagation and can simulate processes such as human memory; the use of parallelism in a neural net combined with the structure of an MIS can allow for extremely quick computation. Moreover, combined with the power of microelectronic sensor nodes, the technology allows for quick sensing as well as solid learning and information processing capabilities.

4.3 Vertices necessary for real-life problems

A particularly interesting application of MIS and its relevance to vertex modeling for real-life problems further concerns protein structure modeling [Söderhjelm et al., 2012]. As an example, the use of MIS to model the alpha carbons of trypsin resulted in a 223-vertex graph, which could easily operate on a 256-qubit quantum computer or any other quantum computer with qubits on the order of a few hundred, such as QuEra's Aquila [QuEra]. Other applications of MIS could include far smaller graphs to model local systems of "happy" vertices such as the distribution of pencils in an elementary school classroom. Indeed, graphs like these could easily have as few as 10 to 15 vertices and are perhaps simulatable by the algorithms that our team worked on this weekend.

5 Conclusion

We found this challenge over the past 24 hours to be very interesting and certainly worthy of further investigation. If we had more time, we would try to implement some of the other optimization techniques we looked into as described above but did not have time to fully implement, as well as further investigating the values of annealing hardness for different size graphs. In addition, we would have spent more time looking into more values of R_b/a smaller than 3 and investigating their viability with Aquila. We deeply thank QuEra for giving us the opportunity to work on this interesting and novel problem using their state of the art quantum hardware.