Cardiff University

School of Computer Science and Informatics

CM3203 – Final Year Project

Reinforcement Learning Algorithms for Controlling Quantum Spin-$_{1/2}$ Networks

---

Author: Anastasia Ugaste

Supervisor: Dr F.C.Langbein

Table of Contents

1. **Project Description 3**
2. **Project Aims & Objectives 4**

2.1 Aims 4

2.2 Milestones 4

2.3 Risks 4

3. **Work Plan 5**

Project Description:

Quantum control offers methods to steer the dynamics of quantum systems; this is particularly useful for building devices for quantum computing, simulation and networking.

This is done using optimisation algorithms such as L-BFGS (which approximates the second derivative in the Newton algorithm that working with a vector rather than a matrix) and stadnard optimisation algorithms for control like GRAPE or KROTOV.

KROTOV is based on a sequential update: for a piecewise constant control signal, the controls are updated for each time step, optimising the controls per time step using L-BFGS or similar methods.

GRAPE, even though it also uses L-BFGS, is based on a concurrent update - a piecewise constant control signal, which is updated for all time intervals. This is closest to the standard gradient-based optimisation used elsewhere.

For this project, we will consider a quantum-½ spin system, which can be described by a Hamiltonian[1]:

$$H_{\text{full}} = \sum_{n=1}^{N} \Delta_n Z_n + \sum_{m \neq n} J_{mn} [X_m X_n + Y_m Y_n + \kappa Z_m Z_n]$$

The model is interacted with by the in-signal, which is transmitted using continuous time intervals. The probability of the final state at that particular time is defined by the solution of the Schrodinger's equation[1]:

$$p(t) = \max_{\Delta} |\langle \text{out} | e^{-itH_\Delta} | \text{in} \rangle|^2$$

Once the maximum probability is found, that's the out-signal of the system. If out -signal of the model is corresponding to the in-function, it means that the probability of successful signal transmission is high.

The aim of the project is to optimize an in-signal in the time period so that the probability of needed final state was maximal. In order to do so, we will use Reinforcement Learning algorithms - we hope the reinforcement learning approach will be able to deal with the uncertainties in the system that are hard to put in a model for open loop control. We will simulate the uncertainties for the project to test how far reinforcement learning algorithms can improve upon the open loop algorithms. As a starting point, we will try a REINFORCE algorithm on the system, and later on will be looking into other learning techniques, such as deep Q learning and an actor-critic approach[2].Our approach will be model free, and the learning element will be using deep neural networks.

At the end of my project, I would conclude if reinforcement learning algorithms tested during this project can give us more robust controls under uncertainties compared to the gradient based methods and bring a benefit to the currently used methods in quantum

optimization problems. This will be demonstrated by showing how much the fidelity of the controls varies under uncertain parameters in the Hamiltonian and dephasing in simulation.

Project Aims & Objectives:

1. Research several reinforcement learning algorithms and decide on the most appropriate one to be used for quantum problems based on the comparison of 2-3 reinforcement learning algorithms with L-BFGS algorithm results that we already have.
2. Implement REINFORCE reinforcement learning algorithm to solve the quantum energy landscape control problem.
3. Implement another reinforcement learning algorithm (either a temporal difference learning or alpha-zero approach – depending on the initial experiments) to see if the results improve
4. Evaluate the algorithms based on their robustness compare the results of the reinforcement algorithms created versus the L-BFGS

Milestones:

1. The working environment is built, the L-BFGS algorithm results are obtained using the existing code (completed by Week 2)
2. A basic Reinforcement Learning Algorithm is produced (completed by Week 4)
3. Reinforcement Learning Algorithm's results are evaluated versus L-BFGS algorithm and documented (completed by Week 5)
4. Advanced approach has been selected (completed by Week 6)
5. Advanced approach has been implemented (completed by Week 7)
6. Advanced approach's results have been analysed and compared with a basic Reinforcement Learning Algorithm and with L-BFGS algorithm, the conclusion has been drawn. (completed by Week 10)

Risks:

The main concern associated with this project is that nobody has tried to use Reinforcement Learning algorithms for this purpose before, so we may end up with a situation where the results obtained are not promising. If this would be the case, we can continue testing other various Reinforcement Learning algorithms, or even use alpha-zero or mu-zero (which we know were already successfully used for quantum problems). We can also consider going into higher excitation subspaces for the energy landscape control problem as an alternative.

Second concern is partly linked with a first one. Due to the nature of the project, the progress needs to be seen from Week 3 at latest – which limits the time for understanding the theory behind the algorithms used and implementing them.

Work Plan

The project will be completed over the scope of 12 weeks, with weekly meetings with my supervisor.

The plan below shows the detailed scope of work to be done within this period. As the outcomes of each week depends on the results of the previous, the weekly activities that I will be talking about below will contain a general approach to the problem but will not go into deeper detail.

We presume that nothing has been done regarding the project activities, apart from some background reading during the Christmas Period. Milestone weeks are highlighted in the work plan below in a different colour and have a link to the corresponding milestone from the check list above.

During the development of the project, some changes will likely be made, thus some weeks below will not be assigned in full – to allow some time for write-up and reflection on the work done, or to catch up with work from the previous week if needed.

Week 1 (01/02/2021 – 07/02/2021)	I will be focusing on writing my initial plan, doing some background reading, familiarizing myself to the existing L-BFGS algorithm code, building a conda environment and researching Reinforcement Learning Algorithms existing and picking up the one that I think suits the project best.
Week 2 (07/02/2021 – 14/02/2021)	Milestone (1) from above Once the initial plan is out of the way, after I have done my research on Reinforcement Learning algorithms, I will present to my supervisor an algorithm I want to implement that I think works the best (REINFORCE), as well as show him my environment with some results from L-BFGS algorithm code that is widely used in research society.
Week 3 (14/02/2021 – 21/02/2021)	This week will be dedicated to the development of the algorithm from Week 2. As a starting point, I will assume that algorithm's "actions" will be the

	changes to controls, and the “reward” - the change in fidelity of the system.
Week 4 (21/02/2021 – 28/02/2021)	Milestone (2) from above This is another milestone week – by this stage I should have a working REINFORCE algorithm that can provide some insights towards the final conclusion if the reinforcement learning algorithms can bring something new in or not. Some thought can already be given towards the possible improvements of the algorithm.
Week 5 (01/03/2021 - 08/03/2021)	Milestone (3) from above Week is dedicated towards evaluating and comparing the results in terms of fidelity robustness. Some methods may include: ➔ plotting a scatter map ➔ Probability Density Functions(PDFs) of the fidelity under the uncertainties for a particular controller ➔ sensitivity of the system to noise in comparison with the other algorithm
Week 6 (08/03/2021 - 15/03/2021)	Milestone (4) from above Based on conclusions made in the previous week and the progress so far, me and my supervisor will decide upon the improvements on the reinforcement algorithm that is implemented by now.
Week 7 (15/03/2021 - 22/03/2021)	Week 7, 8 and Easter will be dedicated to improving the algorithm. This will also include a research phase (to see how people tackled the problem using other algorithms before), coding phase and refinement phase (look at the improvements made, analyse the results with the improved program, try a different approach/extend an existing one if needed)
Week 8 (22/03/2021 - 29/03/2021)	
Easter (29/03/2021 - 18/04/2021)	
Week 9 (19/04/2021 - 26/04/2021)	Milestone (5) from above In Week 9, I will be finalizing the improved version of the algorithm, ensuring the version of the improved algorithm is stable and is ready to be presented to the supervisor next week.
Week 10 (26/04/2021 - 03/05/2021)	Milestone (6) from above By this week I should have a stable improved algorithm that can deal with more complex

	problems. I will gather all available evidence and discuss all available results with the supervisor to draw a reliable conclusion.
Week 11 (03/05/2021 - 10/05/2021)	Even though I intend on making extensive notes throughout the project, I will use last two weeks to polish my write up, and to finish off some bits of the program if it is not completed by Week 10.
Week 12 (10/05/2021 - 15/05/2021)	

Bibliography:

[1] F.C. Langbein, Sophie G Schirmer, Edmond Jonckheere - Time optimal information transfer in spintronics networks (https://langbein.org/wp-content/uploads/2015/11/CDC2015.pdf)

[2] Richard S. Sutton and Andrew G. Barto - Reinforcement Learning: An Introduction (SuttonBartoIPRLBook2ndEd.pdf (stanford.edu))