Machine Learning for Quantum Crosstalk Calibration (MLCA)
Project Authors Name: Hisen Zhang & Dan Fiumara
Project Evaluator Name: Adithya Vijayan & Swastik Kanjilal

- Guidelines to review the report for Question 1-3 below: https://dl.acm.org/journal/dgov/reviewer-guidelines
- Guidelines to review the code artifacts for Question 4-8 below: https://conferences.sigcomm.org/sigcomm/2022/cf-artifacts.html

1. **Summary**

Provide a brief summary of the project in your own words.
This project written by Hisen Zhang & Dan Fiumara applies Machine Learning (ML) to improve the calibration of superconducting qubits affected by crosstalk. When conducting experiments across all 5 qubits, their method achieves greater than 85% success rate. This shows that their system has reliable performance.

2. **Strengths**

Provide strengths or positive aspects of the project.

- We liked how everything was set up: equations, images, subsections, etc.
- Additionally, the captions under the images were really detailed, allowing us (as readers) to really understand what was going on

3. **Weakness**

Provide any weakness or aspects that can be further improved.

- The code was a little difficult to go through and run
    - To properly produce the results, a decent amount of steps to go through.

4. **Documentation: Is the artifact/code sufficiently documented?**

   Rate from 0% to 100%, where 0% means "documentation is completely insufficient" and 100% means "documentation is absolutely sufficient". If you need to assess both a dataset and tools, please take the average and comment below. In assessing tools, please consider if they are easy or difficult to install/set up and get to run. In assessing datasets, please consider if the meta data is sufficient.

   Choices are:

- **1.** 0%
- **2.** 20%
- **3.** 40%
- **4.** 60%
- 5. 80%
- **6. 100%**

Documentation: Comment on/explain your choice above:

The documentation was well written, and the ReadMe file was also good, just a little difficult for us to follow through. Additionally, all different files were organized into the various folders well (e.g. experiments, data collection, etc)

Completeness: Do the submitted artifacts/code include all of the key components described in the report?

Rate from 0% to 100%, where 0% means "does not include any key components" and 100% means "includes all key components".

Choices are:

- **1.** 0%
- **2.** 20%
- **3.** 40%
- **4.** 60%
- **5.** 80%
- **6. 100%**

Completeness: Comment on/explain your choice above

Everything mentioned in their report was also detailed in their code, with proper comments.

6. **Exercisability: Do the submitted artifacts/code include the scripts and data needed to run the experiments described in the paper, and can the software be successfully executed?**

   Rate from 0% to 100%, where 0% means "the scripts/software cannot be successfully executed and/or no data is included" and 100% means "the artifact includes all necessary scripts/software and data, and scripts/software (if present) can be successfully executed".

   Choices are:

- **1.** 0%
- **2.** 20%
- **3.** 40%
- **4.** 60%
- **5.** 80%
- **6. 100%**

Exercisability: Comment on/explain your choice above

All the scripts were present in the folders, the ReadMe file told us how to run the code itself.

7. **Results attainable: Does the artifact/code make it possible, with reasonable effort, to obtain the key results from the artifact/code?**

   Rate from 0% to 100%, where 0% means "no results can be obtained" and 100% means "all results can be obtained".

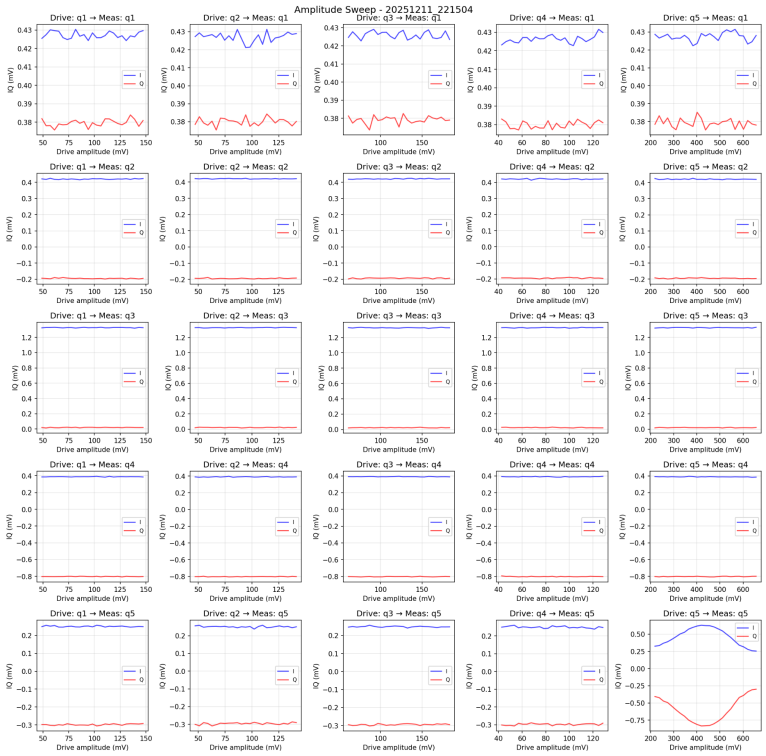
   Choices are:

- **1.** 0%
- **2.** 20%
- **3.** 40%
- **4.** 60%
- **5.** 80%
- **6. 100%**

Results attainable: Comment on/explain your choice above

The graphs all look good, it was also nice that they uploaded the results in addition to the code as well, so that we can compare the results with the actual

results.They also provided the final .csv file with all the data they captured, which was nice to visualize.



8. **Results completeness: How many key results of the paper/report is the provided code meant to support?**

   Rate from 0% to 100%, where 0% means "the artifact is meant to support no key results" and 100% means "the artifact is meant to support all key results".

   Choices are:

- **1.** 0%
- **2.** 20%
- **3.** 40%
- **4.** 60%
- **5.** 80%
- **6. 100%**

Results completeness: Comment on/explain your choice above

The results all look good, I liked how all 25 graphs were on one image, so it was easy to compare with each other.

Reviewer Team member1 Name, Signature
Adithya Vijayan

Reviewer Team member2 Name, Signature
Swastik Kanjilal