Classifiers for Improving Superconducting Qubit Readout Fidelity from Homodyne Detection

Fischer Moseley (fischerm@mit.edu)
Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology
Department of Physics, Massachusetts Institute of Technology
(Dated: March 1, 2022)

Superconducting qubits provide a feasible modality for implementing many-qubit algorithms, but suffer from an incorrect determination of qubit state during readout to room-temperature electronics. This is due to coupling to a noisy environment, which produces a noisy readout signal as well as potential for qubit relaxation, dephasing, and thermalization. Process improvements seek to decrease noise and increase qubit lifetimes by physical modifications to the qubit, but the fidelity with which the qubit state is determined is also dependent on the manner in which the readout data is processed. This work compares three classifiers for determining qubit state given a measurement of a single field quadrature provided by a homodyne measurement. It is shown that direct applications of networks with Multilayer Perceptron-based classifiers produce subpar results when compared a simple averaging classifier, but providing MLP classifiers with time-averaged data increases readout fidelity by nearly four percent.

I. INTRODUCTION

Quantum computation provides a means of solving computational problems that would be intractable for classical supercomputers. Using a quantum computer to factor prime numbers with Shor's algorithm is the canonical example, but such devices have applications in cryptography, machine learning, and fundamental science.

The current state of quantum computing has been described as being in the Noisy Intermediate Stage Quantum (NISQ) era.[1] Present processors have enough quantum bits ('qubits') to be useful for a number of tasks, but are still relatively noisy. Quantum error-correction techniques mitigate this by using additional qubits for redundancy, at the expense of using those qubits for computations. This is being mitigated by next-generation quantum processors with impressive qubit densities [2], but current processors still lack the qubit density to broadly implement error-correcting techniques. This forces users of quantum computers to make a tradeoff between complexity and utility; more useful algorithms require more gates to be applied to a qubit, but this takes time, during which the qubit state could relax or decohere. As a result, focus is given to improving readout fidelity.

Fidelity can be improved by decreasing noise, which arises from a coupling between the qubit and noise sources in the environment. In superconducting qubits one such source is thermal noise, which is abated by cooling the processor down to 10mK in a cryostat, reducing the amplitude of the noise source. Purcell filters have also been placed between the qubit resonator and the microwave transmission lines used for its control. This reduces the coupling to the noise sources, rather than reducing the magnitude of the noise source itself. These techniques (and many others) serve to better physically isolate qubits and their associated circuity from the environment, decreasing the amount of time needed to obtain sufficient information about the qubit state, as well as increasing the qubit relaxation and coherence times.

While qubit lifetimes can't be increased by novel processing of the readout data, the amount of information obtained can be. If were less sensitive to disturbances to the qubit during readout, then longer operations could be run while still providing meaningful results. The readout process behaves as a classifier, taking microwave tones from resonator cavities as inputs, and giving a binary assignment of the qubit state into either eigenstate of σ_z as an output. This represents the crossover over the interpretation of qubit state from quantum to classical, and faster

improvements to the classification process can readout fidelity without modification to hardware.

In the case of superconducting qubits, homodyne detection is produced by mixing the microwave readout tone from a qubit-coupled resonant cavity with a local oscillator set to the qubit frequency. This downshifts a 5-6GHz readout tone to a few hundred MHz, which is digitized and demodulated into I and Q quadratures. Since the qubit state is dependent on the frequency of the readout tone, the components of the downshifted tone encode a measurement of the qubit state. Homodyne detection measures one field quadrature, in this case $< a + a\dagger >$. If the detection were heterodyne, then the other quadrature $< j(a-a\dagger) >$ would also be recorded during readout. The readout signal is represented by this field quadrature summed with a noise term dW, a Wiener increment sampled from a Gaussian distribution:

$$r(t) = < a + a\dagger > + dW \tag{1}$$

Classifying this continuous measurement trajectory through phase space into a determination of the qubit state in the σ_z basis is the classification problem investigated in this work. Three classifiers are presented, evaluated with simulated qubit readout data, and their merits compared.

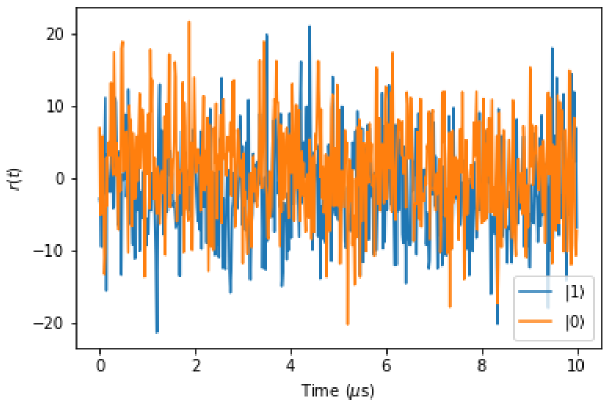

FIG. 1. Examples of $r(t)$ trajectories for a qubit in the $|0\rangle$ and the $|1\rangle$ states. The noise on each is similar, but a slight upward bias can be seen in the readout of the $|0\rangle$ state.

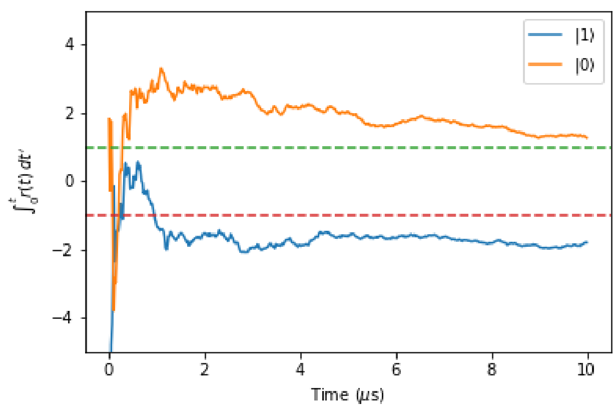

FIG. 2. The same $r(t)$ curves from Figure 1, cumulatively integrated. The separation between trajectories becomes clear, as $|0\rangle$ states will converge to near -1, and $|1\rangle$ states will converge to near +1.

II. RELATED WORK

This problem has been addressed before with a variety of classifiers. Linear filters have been designed to convolve with $r(t)$ trajectories, such that their final value is better representative of the qubit state.[3] These have also been optimized quantitatively [4], and various machine-learning approaches exist with RNNs. [5]

III. METHODS

A single-box classifier, a double-box classifier, and a Multi-Layer Perceptron-based classifier were evaluated for their ability to classify an arbitrary qubit readout trajectory $r(t)$ into either $|0\rangle$ or $|1\rangle$, the positive and negative eigenstates of σ_z This was evaluated quantitatively by defining the readout fidelity as:

$$F = 1 - P(1|0) - P(0|1) \tag{2}$$

Where $P(1|0)$ and $P(0|1)$ are the probabilities of interpreting a qubit prepared in $|0\rangle$ as being in $|1\rangle$, and vice versa. [3]

A. Sample Data Generation

Since data from a real quantum computer was not available, $r(t)$ trajectories were simulated with QuTiP's stochastic master equation solver, which was programmed the Jaynes-Cummings Hamiltonian to describe the qubit-resonator system:

$$H_{JC}/\hbar = \omega_r \left(a \dagger a\right) - \frac{1}{2}\omega_q \sigma_z + g\left({}_{+}a + \sigma_{-}a\dagger\right) \tag{3}$$

The system is operated in the dispersive limit where $|\Delta| \ll g$, which simplifies the Hamiltonian:

$$H_{JC(disp)}/\hbar = \left(\omega_r + \chi\sigma_z\right) a \dagger a + \frac{1}{2}\tilde{\omega_q}\sigma_z \tag{4}$$

To compare classifier performance, 5,000 trajectories were generated for the density matrices corresponding to the $|0\rangle$ and $|1\rangle$ states. This was done to ensure that the feature space was well represented through the noise, which is of particular importance to the network-based classifiers. It's also worth noting that the system dynamics have a fixed point at the σ_z eigenstates, so a small component of the complementary eigenstate was added while preparing the density matrices, taking care to preserve normalization. [6] For example, the density matrix corresponding to $|0\rangle$ was modified as $\rho_{|0\rangle} \to \tilde{\rho}_{|0\rangle} = (1-\delta)|0\rangle\langle 0| + \delta|1\rangle\langle 1|$, with a similar modification for the $|1\rangle$ state.

The trajectories were post-processed to compensate for this. Since perfect knowledge of the qubit state is known by querying the simulator's internal state, we can compute $\langle\sigma_z\rangle$ directly. This is used to filter out trajectories that don't evolve to the desired eigenstate. About 50 of the simulated trajectories show this behavior, which makes sense given the 5,000 trajectories simulated and chosen δ of 0.01.

This Hamiltonian does not model any form of qubit decay - relaxation, decoherence, or thermalization. This implies that an infinite-horizon average of $r(t)$ will yield a perfect estimate of qubit state as the noise integrates out to zero. However, this is not realizable experimentally as practical qubit lifetimes are finite. As a result, in the analysis here trajectory is truncated at $1\mu s$ in order to represent the imperfect knowledge of the qubit state. This is an approximation, but will serve for the analysis here while the author bolsters his knowledge of QuTip's `smesolve`.

B. Description of Classifiers

Averaging Classifier:
The averaging classifier defines an integration time t_i, over which $r(t)$ is averaged as shown in Equation 5. If the average is less than zero, then the qubit is determined to be in the excited state. If the average is greater than zero, then the qubit is determined to be in the ground state. The exact numerical value of the average can be interpreted as a confidence that the qubit is in the determined state. The closer the value to 1, the greater the likelihood that the qubit is in the ground state. Similarly, the closer the value to -1, the more likely the qubit is in the excited state. In this work, only the sign of the average is used to assign the qubit state.

$$s = \frac{1}{t_i} \int_0^{t_i} r(t)\, dt \tag{5}$$

Direct MLP Classifier:
The direct Multilayer Perceptron (MLP) classifier is a densely connected neural net, comprised of an input layer, a number of hidden sigmoid layers, and an output layer. It takes the entire $\langle r(t) \rangle$ trajectory as input, and outputs a scalar between 1 and -1, assigning each qubit into either eigenstate of $\langle \sigma_z \rangle$. Again, only the sign here is used - but interestingly this didn't appear to matter. No restrictions were placed on the output layer to only output -1 or +1, but during testing no other numeric value came out of the classifier.

Averaging MLP Classifier:
The averaging MLP classifier combines characteristics from the other two classifiers. The cumulative average of the readout signal $r(t)$ is taken to produce $\overline{s}(t)$ as shown in Equation 6. This input is then fed to the classifier in exactly the same way as used by the Direct MLP method.

$$\overline{s}(t) = \frac{1}{t} \int_0^{t} r(t')\, dt' \tag{6}$$

Of the slightly less than 5,000 trajectories left after postprocessing, 4,000 of them were given to the MLP classifiers as training data, and the performance evaluated on the remaining thousand.

IV. RESULTS

When evaluated over a $1\mu s$ sampling window, it was found that the averaging classifier gave a readout fidelity of 70.41%, while direct MLP could not achieve a fidelity higher than 64.5%, even after adding four hidden layers of 100 neurons each. The averaged MLP classifier was able to achieve a fidelity of 74.3%, with just two 100-neuron layers, outperforming the other classifiers.

V. FURTHER WORK

The method in which information was removed from the readout process here is contrived. Solving real readout problems requires decay terms to be added, and is the immediate next step short of obtaining real qubit readout data.

Once experimental data becomes available, it would be useful to evaluate these classifiers across qubit frequencies. It would be interesting to see how well a MLP classifier trained on one qubit transfers to other qubits on the same chip. It would be extremely useful to train a classifier on a single qubit that was robust enough to be effective for other qubits on the same device. Standard data regularization techniques may find be able to detect overfitting, but that may not be the dominant mechanism by which the robustness of the classifier is undermined. Including the resonator power and qubit frequency as feedforward terms to the neural network may also be useful. Lastly, using stateful neural networks such as LSTMs or RNNs may be able to better capture the abrupt change in trajectory that occurs when a qubit relaxation or excitation event occurs. The success of the averaging MLP classifier suggests that these may also benefit from a few densely-connected layers upstream for initial filtering of the readout tone.

VI. SOURCE CODE

The code used for this analysis can be found at `https://github.com/FischerMoseley/readout_sims` An associated Deepnote project is linked there, which includes all the simulation, processing, and classification code. It can be executed in a web browser by anyone curious, and the author highly recommends looking through the `simulation.ipynb` notebook. More intermediate plots are available there that better tell the story outlined in this text.

VII. ACKNOWLEDGEMENTS

Many thanks are due to Patrick Harrington in MIT's EQuS group for help getting valid simulation output out of QuTip. Many more thanks are due to Kevin O'Brien and Bright Ye for offering 6.644 this semester. At the beginning I thought I was in way over my head, but I learned far more than expected and and the support throughout the term has been greatly appreciated. It's been a lot of fun, thanks for having me ;)

[1] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, Superconducting qubits: Current state of play, Annual Review of

Condensed Matter Physics **11**, 369–395 (2020).

[2] IBM Newsroom, IBM unveils breakthrough 127-qubit quantum processor, (2021).

[3] T. Walter, P. Kurpiers, S. Gasparinetti, P. Magnard, A. Potočnik, Y. Salathé, M. Pechal, M. Mondal, M. Oppliger, C. Eichler, and et al., Rapid high-fidelity single-shot dispersive readout of superconducting qubits, Physical Review Applied **7**, 10.1103/physrevapplied.7.054020 (2017).

[4] J. Gambetta, W. A. Braff, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, Protocols for optimal readout of qubits using a continuous quantum nondemolition measurement, Physical Review A **76**, 10.1103/physreva.76.012325 (2007).

[5] Élie Genois, J. A. Gross, A. D. Paolo, N. J. Stevenson, G. Koolstra, A. Hashim, I. Siddiqi, and A. Blais, Quantum-tailored machine-learning characterization of a superconducting qubit (2021), arXiv:2106.13126 [quant-ph].

[6] A. Chantasri, J. Dressel, and A. N. Jordan, Action principle for continuous quantum measurement, Physical Review A **88**, 10.1103/physreva.88.042110 (2013).

[7] K. O'Brien and B. Ye, Lecture Notes, MIT 6.644 - Quantum Measurement and Noise (2021).

[8] Qiskit Textbook, Exploring the Jaynes-Cummings Hamiltonian with Qiskit Pulse (2021).