Quantum Measurement Classification using Statistical Learning

ZACHERY UTT, University of Florida, USA
DANIEL VOLYA, University of Florida, USA
PRABHAT MISHRA, University of Florida, USA

Interpreting the results of a quantum computer can pose a significant challenge due to inherent noise in these mesoscopic quantum systems. Quantum measurement, a critical component of quantum computing, involves determining the probabilities linked with qubit states post-multiple circuit computations based on quantum readout values provided by hardware. While there are promising classification-based solutions, they can either misclassify or necessitate excessive measurements, thereby proving to be costly. This paper puts forth an efficient method to discern the quantum state by analyzing the probability distributions of data post-measurement. Specifically, we employ cumulative distribution functions to juxtapose the measured distribution of a sample against the distributions of basis states. The efficacy of our approach is demonstrated through experimental results on a superconducting transmon qubit architecture, which show a substantial decrease (88%) in single qubit readout error compared to state of the art measurement techniques. Moreover, we report additional error reduction (12%) compared to state-of-the-art measurement techniques when our technique is applied to enhance existing multi-qubit classification techniques. We also demonstrate the applicability of our proposed method for higher dimensional quantum systems, including classification of single qutrits as well as multiple qutrits.

CCS Concepts: • **Hardware → Quantum computation**.

Additional Key Words and Phrases: Quantum Computing, quantum measurement, error mitigation, statistical learning

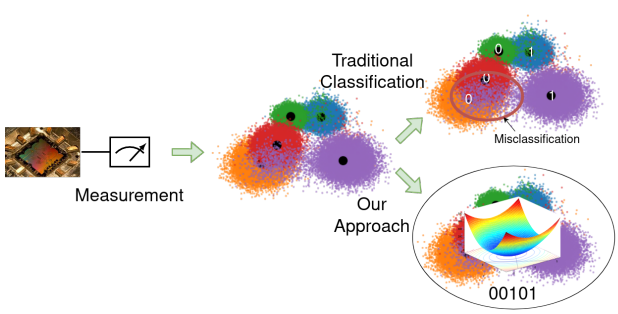
ACM Reference Format:
Zachery Utt, Daniel Volya, and Prabhat Mishra. 2024. Quantum Measurement Classification using Statistical Learning. *ACM Trans. Quantum Comput.* 1, 1, Article 1 (January 2024), 16 pages. https://doi.org/10.1145/3644823

1 INTRODUCTION

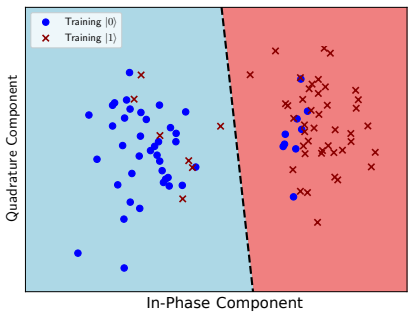
It is broadly recognized that quantum computing holds the promise to outperform classical computing in a variety of complex problem domains, thanks to quantum phenomena such as entanglement and superposition [1–7]. Unlike a classical computer, which operates in a single state at any given moment, a quantum computer can inhabit a blend of states simultaneously. However, this quantum advantage comes at the expense of increased noise and unwanted entanglement with the environment, complicating the processing of wanted information [8–12]. Furthermore, the measurement of a quantum system forces it to collapse into a specific known state with some probability, necessitating multiple computations (shots) to gather the outcomes.

Authors' addresses: Zachery Utt, University of Florida, Gainesville, FL, 32611, USA, utt.zachery@ufl.edu; Daniel Volya, University of Florida, Gainesville, FL, 32611, USA, dvolya@ufl.edu; Prabhat Mishra, University of Florida, Gainesville, FL, 32611, USA, prabhat@ufl.edu.

2643-6817/2024/1-ART1 $15.00
https://doi.org/10.1145/3644823



(a) Traditional versus proposed approach



(b) An illustrative example

Fig. 1. An overview of quantum measurement. (a) Unlike traditional approaches that consider each data point separately, our proposed approach considers all data points together. (b) Example of a linear quantum state discrimination in the IQ space. The model is trained by configuring the quantum computer to output known samples of basis states $|0\rangle$ and $|1\rangle$, which are measured and used to divide the measurement space with a hyperplane. When performing inference, a measurement that is located above the hyperplane will generate a prediction of $|1\rangle$; and a measurement located below the hyperplane will generate a prediction of $|0\rangle$. Training data was sampled from *ibm_quito*, and the hyperplane was computed using *scikit* learn.

Noise within quantum systems, manifesting as computational errors, can originate from several sources including (a) the preparation of the initial state, (b) during computation itself, and (c) at the point of result measurement [13]. The focus of this study is on the mitigation of these quantum measurement errors. Namely, to foster robust quantum computing, it is essential to alleviate measurement errors which can range from 5% to 30% in contemporary devices [14, 15]. Current quantum computers apply a combination of sensitive hardware (detectors) and standard classification techniques to deduce the properties of the quantum state. This data-driven analysis is carried out on classical computers and requires numerous repetitions of quantum circuit executions (shots) on the quantum computer to converge to an accurate solution on the classical machine. The classification used to link a quantum measurement to its deduced quantum state is referred as a *qubit discriminator*.

Figure 1a shows an overview of the measurement procedure in quantum computers. Common physical representations of quantum computers, such as superconducting transmons, characterize a collapsed qubit as a 2-component vector – the in-phase and quadrature components (IQ) of a wave observed as it passes through the resonator [16, 17]. The quantum measurement device yields sets of results, including IQ points. Subsequently, a classifier labels each point as either $|0\rangle$ or $|1\rangle$. Our approach leverages the measurement data to construct a cumulative distribution function, which is compared to the training distributions via convex optimization. As a result, this method does not disregard potential hidden statistical traits present in the measurement data.

Quantum measurement error arises due to noisy measurement readings as well as classification errors caused by imperfect discriminators. Hence, the accuracy of the quantum computer is contingent on the performance of the qubit discriminator. In response to these challenges, we suggest an enhanced classification method aimed at boosting measurement precision. Specifically, this paper makes the following key contributions.

- We propose a framework for mitigation of quantum measurement errors using cumulative distribution functions to accurately classify quantum measurements.

- We also propose a scalable enhancement to existing discriminators to significantly improve discriminator accuracy.
- Experimental evaluation demonstrates the effectiveness and efficiency of our model with real quantum data in terms of non-linearity, statistical consistency, and versatility compared to state-of-the-art qubit and qutrit discriminator approaches.

Our framework comes at modest performance reduction, yielding a computational complexity of $O(n \cdot k \cdot \log(k))$, for k being the number of quantum measurements and n denoting the number of qubits. In comparison, traditional methods have a computational complexity of $O(c \cdot k)$, since for each of the n qubits we form a partition function traversing all k samples for each qubit.

This paper is organized as follows. Section 2 surveys related efforts. Section 3 describes our proposed framework. Section 4 proves error bounds for our proposed discrimination method. Section 5 presents the experimental results. Finally, Section 6 concludes the paper.

2 BACKGROUND AND RELATED WORK

In this section, we first provide background on quantum computing. Next, we describe state-of-the art methods for qubit state discrimination. Finally, we highlight the limitations of the existing methods to identify the areas for improvement.

2.1 Quantum Computing

We briefly describe the foundational principles of quantum computing, beginning with the fundamental units that store information, followed by the processes involved in information manipulation, and finally with the retrieval of results.

2.1.1 Qubits. Quantum bits, or qubits, are the fundamental building blocks of quantum computing, similar to how classical bits are in classical computing. Unlike classical bits, which can only take on the value of 0 or 1, a qubit can be in a state that is a superposition of 0 and 1. The state of a qubit $|\psi\rangle$ is represented as $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, where α and β are complex numbers such that $|\alpha|^2 + |\beta|^2 = 1$. The probability of measuring each state is given by the square of the corresponding coefficient. Moreover, qubits have the property of entanglement, meaning that the state of one qubit can be intimately connected to the state of another, regardless of the distance between them.

2.1.2 Qudits. While qubits are a two-state quantum system, qudits generalize this concept to d-level systems where $d > 2$. In other words, if a qubit is a quantum state in a 2-dimensional complex vector space (a 2-level system), a qudit is in a d-dimensional complex vector space (a d-level system). Just like a qubit can be in a superposition of two states, a qudit can be in a superposition of d different states. Mathematically, a qudit $|\phi\rangle$ can be represented as $|\phi\rangle = \sum_i a_i |i\rangle$, where i ranges from 0 to $d-1$, the a_i's are complex numbers such that $\sum_i |a_i|^2 = 1$. This implies that there's a probability associated with measuring the qudit in any of its possible states, and these probabilities sum to 1. Qudits are of particular interest as they can represent and manipulate more information than qubits, potentially enabling more efficient quantum algorithms [18, 19].

2.1.3 Quantum Gates and Circuits. Quantum gates are the basic operations that can be applied to qubits and qudits in a quantum computing system. They are represented as unitary matrices and can manipulate the state of a quantum system in various ways, including changing a single qubit or qudit, or entangling multiple qubits or qudits. A sequence of quantum gates forms a quantum circuit, which can be used to perform complex quantum computations [20–22]. Quantum circuits operate on principles of linearity and unitarity, ensuring that the sum of probabilities of all possible outcomes always equals one.

2.1.4 Measurement. Measurement in quantum systems involves coupling a qubit or a qudit with a measurement device (apparatus), like a resonator in superconducting systems. The quantum state influences the state of the measurement device, which is then read out. This readout is processed to yield measurement signals. However, these signals do not directly reveal the quantum state – instead, they represent certain physical properties (like frequency shifts in a resonator) that correlate with the quantum state. Hence, a classifier is needed to map these measurement signals back to the quantum states. In other words, the classifier decodes the signals from the measurement device, determining whether it represents a 0 or a 1 in the case of a qubit (or 0 through d-1 for qudits). This step is more generally known as quantum state discrimination (classification) and is a crucial component in unveiling the results of a quantum computation. Furthermore, measurement itself is prone to various noise sources which are both classical and quantum in nature. For example, classical noise may be introduced as analog signals must travel through various stages of a dilution refrigerator to an analog-to-digital converter. Quantum noise may enter as the measurement device itself is a quantum system that may entangle with unwanted degrees of freedom [23].

2.2 Related Work

Machine learning techniques are widely used for qubit state discrimination [24–35]. Linear Discriminant Analysis (LDA) is one such model commonly used for qubit state discrimination. A hyperplane is selected to partition the IQ vector space into regions of $|0\rangle$ and $|1\rangle$ based on the measured IQ outputs from the training data, which the model assumes follows a Gaussian distribution. Figure 1b visualizes this method, and highlights the difficulty of the classification task. Due to noise in quantum systems (imperfect measurement devices, environmental contamination, and qubit cross-talk), sampled data contains high variance and may appear in the partition, leading to misclassifications.

Other machine learning models such as k-nearest neighbors (kNN) [26, 27], deep neural networks (DNN) [27–34], and support vector machines (SVM) [28] have been used with quantum IQ data to partition the measurement space into regions of $|0\rangle$ and $|1\rangle$. Some of these models also consider the effects of quantum "crosstalk" – a phenomena where unwanted interactions among qubits can be predicted and accounted for post-readout. While these methods offer alternative ways to partition, *they each implement the same inference workflow by mapping each qubit measurement to a single location within the partition space.* An overall qubit state is obtained by analyzing the frequencies associated with each prediction class. Beyond variations in the partitioning method, further improvements have been obtained by enabling models to tag samples as "inconclusive" [28, 35], and discarding such samples from processing. Some of these methods can be extended to classify higher energy states [34].

2.3 Limitations of State-of-the-Art Approaches

The existing quantum measurement classification methods have the following fundamental limitations.

- The existing models operate by partitioning the IQ space into regions corresponding to each basis state. Regardless of the partitioning method used, the measurement space is often inherently noisy from hardware error, leading to classifications shown in Figure 1b.
- The existing methods map a single IQ measurement tuple into a single quantum state. Since only a single measurement tuple is used for prediction, valuable statistical information encoded within the distribution of test data is neglected. For example, using data collected from the IBM Quito quantum computer, the $|1\rangle$ basis state contains higher variance than

the $|0\rangle$ state [14], and thus sample variance – a distribution property – contains information about bitstate unused by state-of-the-art methods.
- Existing methods assume properties of quantum data (Gaussian) and partition boundaries (linear, quadratic). Since quantum measurements diverge from ideal distributions, such assumptions may introduce bias.
- Some current methods operate by discarding data deemed "inconclusive". We believe a method that quantifies uncertainty without discarding data can outperform these techniques.
- Most qubit discriminator models are difficult to effectively boost (combine with other models). While ensemble techniques exist, such methods require significantly more computation and have not to date demonstrated superior results for quantum discrimination.
- It is difficult to quantify or guarantee convergence with many state-of-the-art qubit discriminator methods. Quantum circuits are often sampled for tens of thousands of iterations, since no stochastic framework exists to bound the error associated with classification. In practice, many quantum engineers evaluate the accuracy of the discriminator by using a test set. Sampling is repeated until the discriminator reaches an accuracy threshold on the test set, requiring potentially thousands of additional quantum samples. Moreover, test sets can be biased, leading to inaccurate conclusions.

We propose a sophisticated quantum discriminator that overcomes these limitations, guarantees convergence, and as an additional benefit, can produce an estimate for the number of samples needed to attain convergence within a threshold without requiring a holdout (testing) set.

3 DISTRIBUTION-BASED CLASSIFICATION FOR MITIGATING QUANTUM MEASUREMENT ERRORS

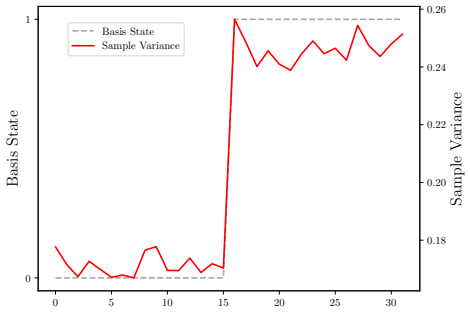
The goal of measurement classification is to take the results of measuring qubits in a quantum register (a collection of IQ points) and correctly identify the corresponding bitstring labels. For example, after measuring the quantum state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, the classifier should provide bitstrings "00" and "11", each occurring with equal probability. Traditionally, classifiers are trained to partition the IQ space, as shown in Figure 1b. We choose to use linear discriminant analysis (LDA) and k-nearest neighbors (kNN) as baselines for this work, due to their prevalence in the community and widespread use in open source libraries, such as Qiskit, and high performance among other discrimination methods [27] on IBM's quantum machines for single qubit state discrimination.

Rather than providing yet another approach to partition the IQ space, we propose an entirely novel distribution-based classification workflow that overcomes the shortcomings outlined in Section 2. Unlike previous methods which produce classifications for every measurement shot, our method directly estimates the probabilities of $|0\rangle$ and $|1\rangle$ in one task.

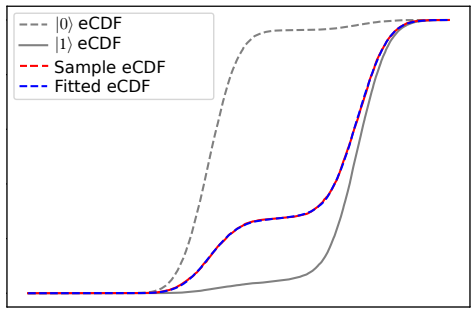
In this section, we first outline the use of cumulative distribution functions (CDF) for classification. We then show an example of classification on a single qubit using CDFs. Finally, we incorporate our CDF approach to existing classification techniques, as highlighted in Figure 3 – providing the advantages of CDF while also being scalable even for a large number of qubits.

3.1 Classification using Cumulative Distribution Functions

It is a well established fact that the Cumulative Distribution Function (CDF) uniquely characterizes a probability distribution. Since a qubit exists in a superposition of states $|0\rangle$ and $|1\rangle$, it follows that the qubit exists in a mixed distribution of basis states $|0\rangle$ and $|1\rangle$. Thus, its unique CDF can be decomposed into a linear (convex) combination of $|0\rangle$ and $|1\rangle$ CDFs, where the weights associated with the constituent $|0\rangle$ and $|1\rangle$ CDFs directly represent the true proportion of measurements that



(a) Measurement variance for single-qubit



(b) Single-qubit cumulative distribution functions

Fig. 2. Visualization of single qubit quantum measurement. (a) Quadrature measurement variance shown on IBM Quito for the $|0\rangle$ and $|1\rangle$ basis states. (b) Visualization of the distribution approach to single qubit state discrimination. The training $|0\rangle$ and $|1\rangle$ empirical CDFs (eCDFs) are shown in gray. The sample qubit eCDF is shown in blue, and the fitted eCDF estimate is shown in red. The fitted eCDF is obtained by creating a convex combination of both gray curves with weights selected to follow the blue curve as closely as possible, as outlined in Section 3.2. The coefficients used for the combination are the estimate for the quibt's state. Data obtained from ibm_quito. Each curve is computed using Algorithm 1.

collapsed into each of the $|0\rangle$ and $|1\rangle$ states. This decomposition for an arbitrary qubit superposition is depicted in Figure 2.

Properties like randomness, state stability, and variance are inherently accounted for in the CDF. Additionally, by analyzing the CDF of the qubit as a whole, we avoid the need to classify each shot individually to arrive at a state estimate. Instead, we analyze properties of the qubit measurement distribution, the $|0\rangle$ measurement distribution, and the $|1\rangle$ measurement distribution to determine the superposition. Next, we discuss methods for estimating each of these CDFs and the procedure used for the decomposition.

3.2 Empirical CDF based Single Qubit Classification

We begin by preparing a training data set of size k for both $|0\rangle$ and $|1\rangle$ quantum states. Each state is prepared, measured, and tagged with the associated $|0\rangle$ or $|1\rangle$ label. Since each measurement consists of both the in-phase and quadrature components, both the $|0\rangle$ and $|1\rangle$ data sets have dimensions $(k \times 2)$. We then construct a test distribution of dimensions $(k \times 2)$ by placing the qubit into a random mixed distribution of $|0\rangle$ with proportion α and $|1\rangle$ with proportion $1 - \alpha$, where α is randomly selected. We will evaluate our method's ability to reproduce α given the test data, the training data, and the training labels. Next, we study the distribution of both the $|0\rangle$ and $|1\rangle$ training data sets. We estimate the CDF of both sets by computing the empirical CDF (eCDF). The empirical CDF is a consistent and unbiased estimator that converges absolutely to the true CDF. Moreover, the Dvoretzky–Kiefer–Wolfowitz inequality [36] provides a closed form error bound for each of the eCDF estimators as a function of the sample size. The eCDF is computed by finding the proportion of values in the data set less than or equal to x, given by

$$\hat{P}(X \leq x) = \frac{1}{k} \sum_{i=1}^{k} I(t_i \leq x) \tag{1}$$

where t_i runs through each element of the set. Here, $I(q)$ is the indicator function, which is given as 1 if q is true, and 0 otherwise. Since this sum denotes the number of elements less than x, we

implement binary search to compute this sum in log(k) complexity using the index of x in the sorted set. We then perform a linear interpolation to transform these staircase-like empirical CDFs to smooth estimates — a technique useful for small data sets.
Algorithm 1 describes the procedure for computing the eCDF for x given a sorted array of either in-phase or quadrature quantum measurements. Using this method, we compute individual eCDFs for the in-phase and quadrature component for each basis state. Similarly, we compute eCDFs for the in-phase and quadrature component for the sample. We use least squares regression to obtain a value for $\hat{\alpha}$, under the constraint $0 \leq \hat{\alpha} \leq 1$, such that we minimize

$$(\hat{\alpha} \cdot \hat{F}_{0,\text{quad}}(x) + (1 - \hat{\alpha} \cdot \hat{F}_{1,\text{quad}}(x))) - \hat{F}_{\text{sample,quad}}(x) + (\hat{\alpha} \cdot \hat{F}_{0,\text{in}}(x) + (1 - \hat{\alpha}) \cdot \hat{F}_{1,\text{in}}(x)) - \hat{F}_{\text{sample,in}}(x) \quad (2)$$

where $\hat{F}(x)$ denotes an eCDF estimate, *in* denotes the in-phase dataset component, and *quad* denotes the quadrature component of the quantum measurement dataset. We use $\sqrt{\hat{\alpha}}$ and $\sqrt{1 - \hat{\alpha}}$ as the estimates for the qubit's state. Overall, $\hat{\alpha}$ is computed in $k \log(k)$ time complexity.

Algorithm 1 Empirical CDF (eCDF) function computation

input: x: Value to evaluate the eCDF at, *array*: Sorted array of quantum measurement data
output: Interpolated eCDF estimate for x

```
procedure eCDF-Curve(x, array)
    if x ≥ array[len(array) − 1] then
        return 1
    else if x ≤ array[0] then
        return 0
    end if
    upper ← binarySearch(x, array)
    lower ← upper − 1
    difference ← array[upper] − array[lower]
    return lower + ((x − array[lower])/difference)
end procedure
```

3.3 Empirical CDF-based Multi-Qubit Classification

In the previous section, we demonstrated how an eCDF single qubit discriminator can be used instead of a traditional discriminator. This is possible since the eCDF discriminator model decomposed the sample distributions into two basis states, representing the full state space of a qubit. In this section, we show how the method can be employed with existing methods for quantum computers with more than one qubit.

With the single qubit eCDF estimation method, it is possible to effectively decompose the mixed distribution of a single qubit into known distributions of $|0\rangle$ and $|1\rangle$ states using estimation and regression techniques. On a machine of n qubits, this method generates n constraints on the position space of 2^n basis states. To illustrate this point, suppose we have a quantum computer with $n = 2$ bits, and it is estimated from the above method that qubit 0 decomposes into $|0\rangle$ with frequency α_0 and that qubit 1 decomposes into $|0\rangle$ with frequency α_1. From this, we have the following constraints (X indicates a "don't care" bit that can take any value): (1) The frequencies of states $|X0\rangle$ given by $|00\rangle$ and $|10\rangle$ sum to α_0, (2) the frequencies of states $|0X\rangle$ given by $|00\rangle$ and $|01\rangle$ sum to α_1, and (3) the frequencies of states $|00\rangle$, $|01\rangle$, $|10\rangle$, and $|11\rangle$ sum to 1.

For a computer with n qubits, the solution space contains 2^n unique basis states, $n+1$ constraints, with $2^n - (n+1)$ remaining free variables. For time complexity purposes, a qubit discriminator can

not typically constrain all 2^n states. We demonstrate how the constraints can improve performance and enhance state of the art classification methods. To underscore this claim, we employ a linear discriminator to classify quantum measurements and measure classification performance before and after the constraints are applied. Our workflow is highlighted in Figure 3.

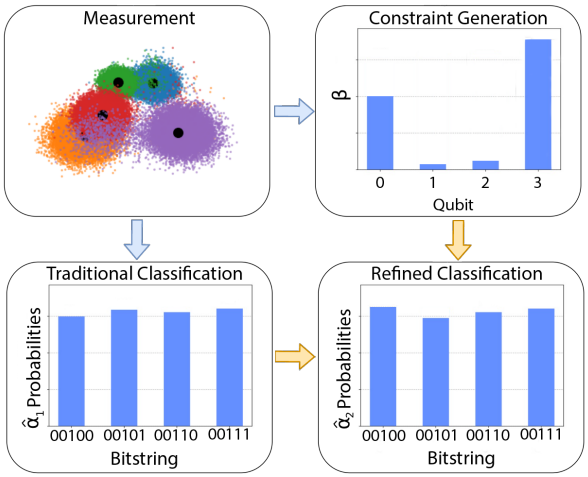


Fig. 3. A high-level overview of the multi-qubit measurement classification procedure that consists of three stages. The first stage (*Traditional Classification*) uses existing methods to produce an estimate. The second stage (*Constraint Generation*) applies the eCDF method to each qubit to generate constraints. The final stage applies the constraints to the existing estimate to produce a new estimate (*Refined Classification*).

We begin by creating a training distribution for each qubit in the $|0\rangle$ and $|1\rangle$ quantum states in the same manner as the single qubit method. We then generate a sample convex label vector $\boldsymbol{\alpha}$ by sampling a random proportion of measurements from each of the 2^n states. For example, if $\boldsymbol{\alpha} = \begin{bmatrix}0.25 & 0.75 & 0 & \dots & 0\end{bmatrix}$, 25% of our test distribution would be sampled from the 00000 state, and 75% of our test distribution would be sampled from the 00001 state.

Next, we employ an existing state of the art analysis method (i.e. LDA) to produce an estimate for the qubit's state, given as $\hat{\boldsymbol{\alpha}}_1$. This is done by first training n traditional models on each qubit's training set. In the traditional method, each IQ pair in the test set is classified independently as $|0\rangle$ or $|1\rangle$, producing an estimate bitstring. This is repeated for every qubit string in the test set and the frequencies are computed to generate $\hat{\boldsymbol{\alpha}}_1$.

Finally, we generate the constraints and produce a refined estimate $\hat{\boldsymbol{\alpha}}_2$ that adheres to each constraint. As described in the previous section, we create eCDF estimates for each qubit's $|0\rangle$, $|1\rangle$, and test distributions. We compile these constraints into $\boldsymbol{\beta}$, a vector of length n which, for each qubit, independently estimates the proportion of that qubit's test distribution measured in the $|0\rangle$ state. Due to the presence of free variables, there are many possible candidates which adhere to all $\boldsymbol{\beta}$ constraints. Rather than considering all of them, we define the refined estimate $\hat{\boldsymbol{\alpha}}_2$ as the distribution closest to $\hat{\boldsymbol{\alpha}}_1$ that adheres to all β constraints. In other words, we update the estimate $\hat{\boldsymbol{\alpha}}_1$ to satisfy the marginal probabilities given by the constraints $\boldsymbol{\beta}$ while minimizing $\|\hat{\boldsymbol{\alpha}}_1 - \hat{\boldsymbol{\alpha}}_2\|$.

3.4 Empirical CDF based Single Qudit Classification

In this section, we show how the eCDF classification method can be extended for quantum systems with more than 2 basis states. Some existing classification methods, like linear discriminant analysis, support vector machines, and binary logistic regression do not naturally extend beyond 2

prediction categories. While strategies like "one-vs-rest" can be used to train binary models for each category, such methods are associated with higher training costs, lower performance, and greater susceptibility to class imbalance.

Rather than training multiple models (as done in "one-vs-rest"), we take the approach of designing a multi-class solution for qubit state discrimination. Assume that our qudit has c basis states. Then, our convex optimization operation can be expressed as:

$$\sum_{i=1}^{c-1}(\hat{\gamma}_i \cdot \hat{F}_{|i\rangle,\text{quad}}(x)) + (1 - \sum_{i=1}^{c-1}(\hat{\gamma}_i) \cdot \hat{F}_{|i\rangle,\text{quad}}(x)) - \hat{F}_{\text{sample,quad}}(x) + \sum_{i=1}^{c-1}(\hat{\gamma}_i \cdot \hat{F}_{|i\rangle,\text{in}}(x)) + (1 - \sum_{i=1}^{c-1}(\hat{\gamma}_i) \cdot \hat{F}_{|i\rangle,\text{in}}(x)) - \hat{F}_{\text{sample,in}}(x) \quad (3)$$

The above optimization leads to the value of $\hat{\gamma}_i$ when the function is minimized. We demonstrate the efficacy of this method in Section 5.3 on a quantum computer with a single qutrit.

3.5 Empirical CDF based Multiple Qudit Classification

The single qudit method can be extended to machines with multiple qudits. Rather than exploring the exponential basis state space formed from a sequence of qudits, we build a scalable estimation pipeline by performing traditional classification (see Section 3.3) with constraints produced from the marginal distribution of each individual qudit (see Section 3.4). We demonstrate the efficacy of such a pipeline in Section 5.4 on a quantum computer with two qutrits.

4 ERROR BOUNDS OF ECDF-BASED DISCRIMINATION

Even after a traditional model has been trained, reliably evaluating its performance can prove challenging. Strong model performance on the training set does not necessarily translate to strong performance on new data, due to over-fitting. Thus, model performance is often measured by computing the accuracy of the model on a holdout set. In this section, we discuss a statistics-driven approach to provide an accurate performance estimate without the use of a holdout set.

4.1 Error Bounds

Our goal in this section is to produce a confidence interval for α such that for a given confidence-level p and a number of shots k, we compute a sufficient upper and lower estimate for α.

Lemma: Let $1 - p$ be a confidence level, and let $\{\hat{F}_n(x)\}$ be a sequence of eCDFs, and let $\{k_n\}$ denote the number of data points used to fit each eCDF.

Then, for all x and for $0 \leq i \leq n$,

$$P(\hat{F}_i(x) - \epsilon_i \leq F_i(x) \leq \hat{F}_i(x) + \epsilon_i) \geq 1 - p \quad (4)$$

where

$$\epsilon_i = \sqrt{\frac{\ln \frac{2}{n \cdot p}}{2 \cdot k_i}}$$

Proof: This is derived from previous work by Dvoretzky, Kiefer, Wolfowitz, and Massart [37] (later enhanced by Kolmogorov and Smirnov) which proved for a *single* distribution with a true cCDF $F(x)$, and an eCDF $\hat{F}(x)$, a $1 - p$ confidence interval is given by

$$P(\hat{F}(x) - \epsilon \leq F(x) \leq \hat{F}(x) + \epsilon) \geq 1 - p$$

where

$$\epsilon = \sqrt{\frac{\ln \frac{2}{p}}{2 \cdot n}}$$

In our method, we are separately computing n empirical CDFs. We can apply the Bonferroni adjustment method

$$p_{\text{family}} = \frac{p}{n}$$

to create a *family*-wide confidence interval. This allows us to extend our expression for a single confidence interval to a set of confidence intervals in which we are $(1 - p)$ confident that *all* confidence intervals in the sequence are correct. ■

4.2 Confidence Interval Construction

In light of the three error bounds in Equation 4, we view each bound as a space of CDFs, where all three spaces contain their respective true CDF with $1 - p_{\text{total}}$ confidence, as shown in Figure 4. Since we view the sample CDF as a sum of CDFs from both the $|0\rangle$ and $|1\rangle$ weighted by α, we can attain error bounds on α by exploring the maximum and minimum values such a weight can take while staying within the bounds.

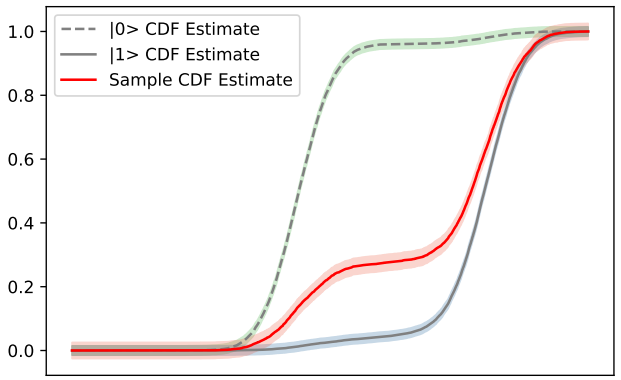


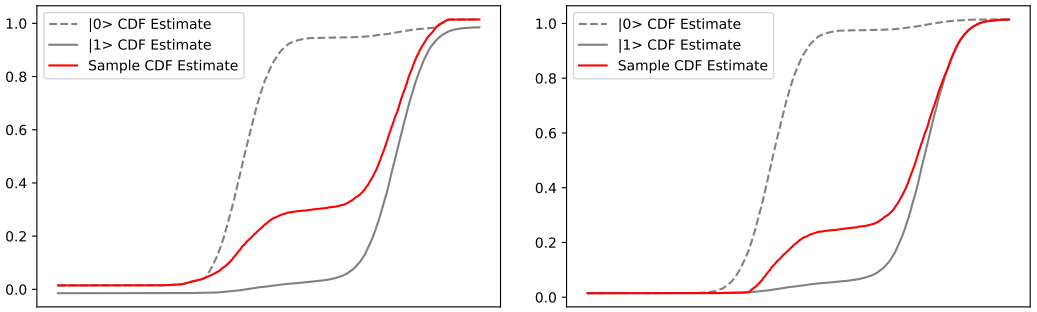
Fig. 4. Visualization of the 99% confidence interval for the family of in-phase eCDFs. If many samples were taken, and all 3 eCDF confidence intervals were generated for each sample, we would find for at least 99% of samples, all 3 eCDF intervals correctly contain the true CDF for all x.

From its construction, to attain the lowest possible value for α, we would select the $|0\rangle$ curve from the family of $|0\rangle$ curves to be as similar (from a least squares perspective) as possible to the sample curve. We would likewise select the $|1\rangle$ curve to be as far from the sample curve as possible.

We assume that we have a working quantum computer in which measured readout values are statistically correlated with quantum state. Specifically, we will assume that the true $|0\rangle$ CDF is greater than the true $|1\rangle$ CDF for all x. Since the sample CDF is a weighted convex sum of the true $|0\rangle$ and $|1\rangle$ CDF, it lies between both curves. As shown in Figure 5, the maximum α occurs when $\hat{F}_{|0\rangle}(x)$ is maximized and $\hat{F}_{|1\rangle}(x)$ is minimized while staying as close as possible to $\hat{F}_{\text{sample}}(x)$.

5 EXPERIMENTS

This section demonstrates the effectiveness of our proposed quantum measurement methods compared to the state-of-the-art approaches. We first outline our experimental setup for single qubit architectures. Next, we show how our method can be implemented on machines with multiple qubits. Lastly, we show how our method can be implemented on machines with multiple qudits.



(a) Curve selection to minimize α

(b) Curve selection to maximize α

Fig. 5. Curve selection examples. In this scenario, we have a 99% confidence interval for α given as $(0.186, 0.316)$ with a point estimate $\hat{\alpha} = 0.249$.

5.1 Experimental Setup

We use ibm_quito, a 5-qubit machine, to initialize states and perform measurements. Quantum circuits and measurements are performed to output $|0\rangle$ and $|1\rangle$ basis states for each qubit. Each measurement is performed 20,000 times (shots), thereby obtaining 20,000 samples of IQ measurements in each of the $|0\rangle$ and $|1\rangle$ quantum states. Data was partitioned into a training and testing set. We evaluate the effectiveness of our proposed methods compared to each state-of-the-art approach implemented in Qiskit [38]. We use Scipy's optimizations library to perform all necessary minimization using the "Nelder - Mead" method.

5.2 Classification Results for Single Qubits

For a single test, we first shuffle the entirety of the experiment dataset. The set is then partitioned into training and testing data. We then generate 1,000 random values for α, each of which lies between 0 and 1. For each value of α, we then build a mixed testing dataset of size 5,000 composed of $100 \cdot \alpha\%$ randomly selected values from the testing data of $|0\rangle$ and $(1-\alpha)\%$ randomly selected values from the testing data of $|1\rangle$. The model is then evaluated on how well it can reconstruct the value of α. We evaluate the mean absolute error (MAE) as the absolute difference between the measurement and truth, given as $|\hat{\alpha} - \alpha|$.

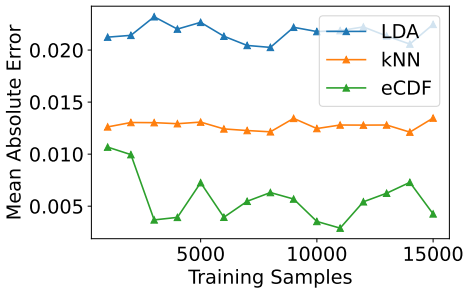
Figure 6a shows the resulting Mean Absolute Error (MAE) of using two traditional discriminators (LDA and kNN) versus our eCDF-based method. The eCDF model attained a lower error at all training sizes and greatly diminished variance as the sample size enlarged. This demonstrates that our proposed approach (eCDF) outperforms traditional discriminators.

Figure 6b shows that our approach attains much lower variance than its machine learning based counterparts, and thus performs more consistently than other methods.

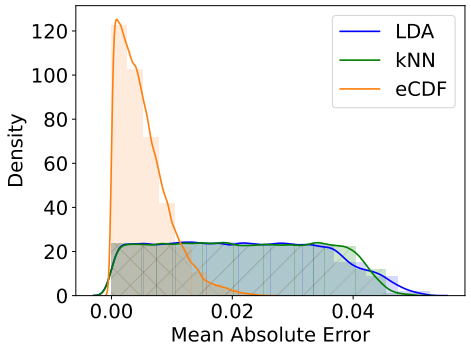
5.3 Classification Results for Single Qutrits

As qutrits are not part of the native device specification, we manually calibrated qutrits on IBM Quito using OpenPulse. This first involved a frequency sweep to find the excitation response for $ket2$. Next, we conduct an amplitude sweep to search for the π-pulse that induces a transition from $|1\rangle \rightarrow |2\rangle$. To reach each of the qutrit states, we perform the following: initialize the qutrit to $|0\rangle$ using a reset; use IBMQ's calibrated X-gate to rotate the qutrit from $|0\rangle \rightarrow |1\rangle$; and finally use our custom π-pulse to transition from $|1\rangle \rightarrow |2\rangle$. We repeat this procedure for each qutrit on the device.

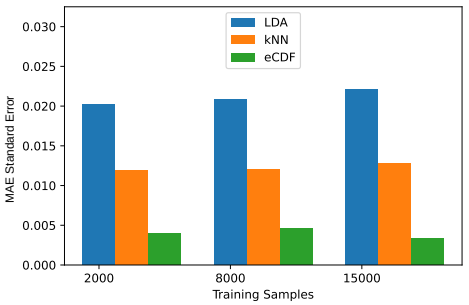
Using the generated data, we constructed a training and test set for a single qutrit, with basis states $|0\rangle, |1\rangle$, and $|2\rangle$. We compare our eCDF qutrit state discriminator with a k-nearest neighbors



(a) Single Qubit Comparison of Accuracy



(b) Single Qubit MAE distributions



(c) Single Qubit Comparison of Standard Errors

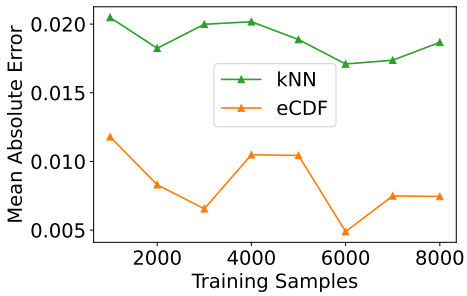
Fig. 6. Mean Absolute Error (MAE) for single qubit reconstruction. (a) Comparison using two traditional models – k-nearest-neighbors (kNN) and linear discriminant analysis (LDA) – and our proposed eCDF discriminator method. The data was shuffled across 120 iterations; and for each iteration, 1000 random distributions were generated, reconstructed, and evaluated. (b) MAE distributions for single qubit state readout of proposed eCDF compared with kNN and LDA. (c) Standard error comparisons showing the tendencies for each model's errors to vary around the means shown in (a)

(kNN) discriminator. Figure 7a shows that the eCDF discriminator achieves a lower mean absolute error (MAE) in all training samples. Moreover, Figure 6c shows that the eCDF discriminator attains a lower variance, and thus performs more consistently than its machine learning counterpart on qutrit measurement data.

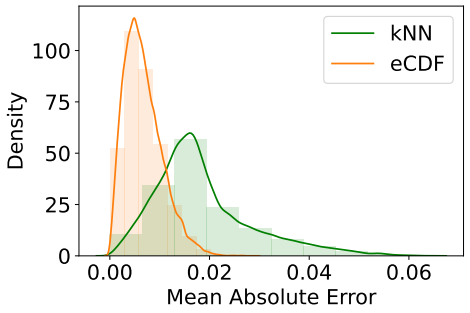
5.4 Classification Results for Multiple Qubits

We begin by configuring the quantum computer to output each of 2^5 basis states for the 5 qubit machine. Similarly, each measurement is performed $k = 20,000$ times. It should be noted, however, that for the purpose of evaluating our method, we consider all 2^n states as candidates for output of the quantum computer. In practice, our method does not require enumeration of all 2^n states, and only considers a maximum of min $(2^n, k)$ states. For a single test, we shuffle the dataset and partition the experimental data into training and testing datasets. We generate a test vector by producing a random convex vector $\boldsymbol{\alpha}$ of size 2^n. We construct a test dataset of size 5,000 by randomly sampling $\alpha_i\%$ values from the i^{th} basis state.

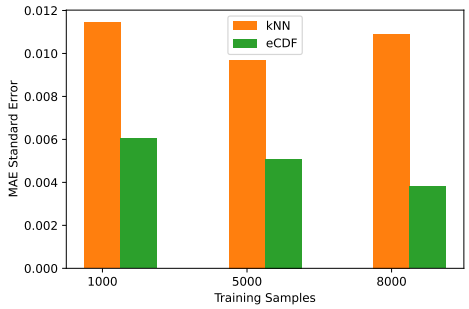
The estimation pipeline is evaluated on how well it can reconstruct the value of α. We compute the mean absolute error as the value $\frac{\|\boldsymbol{\alpha}-\hat{\boldsymbol{\alpha}}_2\|^2}{2^n}$, which we compare to the baseline mean absolute error



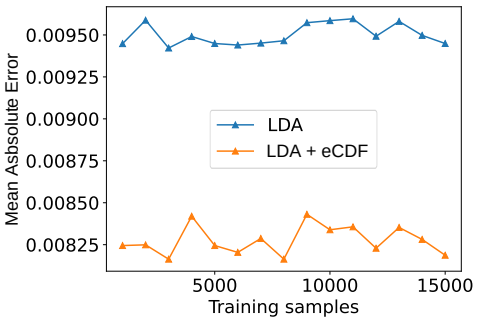
(a) Single Qutrit Comparison of Accuracy



(b) Single Qutrit MAE distributions



(c) Single Qutrit Comparison of Standard Errors



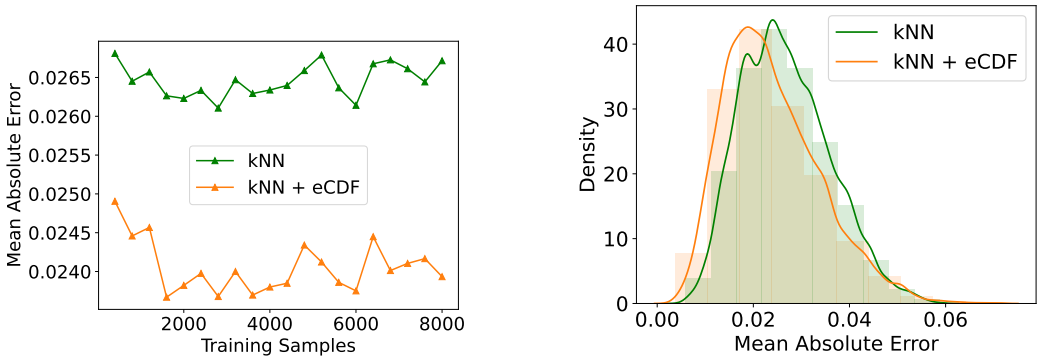
(d) Multi-Qubit Comparison of Accuracy

Fig. 7. Mean Absolute Error (MAE) for single qutrit reconstruction. (a) Comparison between k-nearest-neighbors (kNN) and our proposed eCDF discriminator method. The data was shuffled across 60 iterations; and for each iteration, 1000 random distributions were generated, reconstructed, and evaluated as described above. (b) MAE distributions for single qutrit readouts using proposed eCDF method compared with kNN. (c) Standard error comparisons showing the tendencies for each model's errors to vary around the means shown in (a). (d) MAE improvement for multi-qubit reconstruction of LDA with the proposed eCDF correction. The data was shuffled, and for each iteration 100 random distributions were generated, reconstructed, and evaluated.

given by $\frac{\|\boldsymbol{\alpha}-\hat{\boldsymbol{\alpha}}_1\|^2}{2^n}$. As shown in Figure 7d, the qubit discriminator pipeline with our proposed eCDF model outperforms the traditional discriminator for all training sizes. Therefore, it is beneficial to combine eCDF with traditional models.

5.5 Classification Results for Multiple Qutrits

We extend the dimensionality of our quantum system to demonstrate the robustness of our discriminator. Using the same IBM Quito computer, we constructed a training and test set for a pair of two qutrits with basis states $|00\rangle$, $|01\rangle$, $|02\rangle$, $|10\rangle$, $|11\rangle$, $|12\rangle$, $|20\rangle$, $|21\rangle$, and $|22\rangle$. We select a random sample consisting of a known proportion from each combination of basis states. We compare the MAE of the standalone k-nearest neighbors (kNN) discriminator before and after adding the eCDF correction. Figure 8 shows that with the same inputs, the eCDF method attains significant reductions in MAE at all training levels. These findings highlight the robustness of the eCDF discriminator pipeline and its applications into high-dimensional spaces.



(a) Multi-Qutrit Comparison of Accuracy

(b) Multi-Qutrit MAE distributions

Fig. 8. Mean Absolute Error (MAE) for multi-qutrit reconstruction. (a) Multi-qutrit reconstruction using kNN and our proposed estimation pipeline with eCDF correction. The data was shuffled across 300 iterations, and for each iteration 100 random distributions were generated, reconstructed, and evaluated. (b) MAE distribution for a set of two qutrits using kNN and our eCDF correction using a training set of 800 samples. The figure shows that the usage of the pipeline shifts discriminator outcomes towards lower MAE.

6 CONCLUSION

Quantum measurement classification is fundamental to a successful execution of any quantum algorithm. Measurement classification includes several nuances, such as inherent physical error as well as randomness associated with measured data. In this work, we have introduced a new qubit classifier model that is able to outperform current state-of-the-art machine learning discriminators. The model's performance improvement is achieved through its novel statistical distribution viewpoint, which enables the model to capture important features while diminishing effects of noise and bias associated with individual measurement. Specifically, our proposed eCDF technique significantly reduces mean absolute error (up to 88%) compared with the state-of-the-art in single qubit classification accuracy, with similar improvements in higher dimensional architectures (up to 70% for single qutrits). We demonstrated that this method offers a fundamental improvement (up to 12%) to state-of-the-art multi-qubit classification methods by building a qubit discriminator pipeline that first applies existing qubit discrimination, followed by an eCDF qubit correction stage. In addition to superior performance, our eCDF method can generate confidence intervals for qubit state without requiring a holdout set.

As demand for quantum computing increases, classification techniques that can attain convergence with fewer measurements enable quantum providers to trade off valuable quantum computer resources with processing performed on classical computers. This work opens a path to building robust, yet simple, measurement classifiers based on fundamental statistical principles. It invites quantum engineers to engage with quantum data at the distribution level and provides a framework to add independent qubit distribution insights into existing quantum classification workflows.

ACKNOWLEDGMENTS

This work was partially supported by the National Science Foundation (NSF) grant CCF-1908131.

REFERENCES

[1] E. R. Anschuetz, J. P. Olson, A. Aspuru-Guzik, and Y. Cao, "Variational Quantum Factoring," Aug. 2018.
[2] Y.-H. Oh, H. Mohammadbagherpoor, P. Dreher, A. Singh, X. Yu, and A. J. Rindos, "Solving Multi-Coloring Combinatorial Optimization Problems Using Hybrid Quantum Algorithms," Dec. 2019.
[3] S. Yan, H. Qi, and W. Cui, "Nonlinear quantum neuron: A fundamental building block for quantum neural networks," *Phys. Rev. A*, vol. 102, no. 5, p. 052421, Nov. 2020.

[4] Y. Du, M.-H. Hsieh, T. Liu, and D. Tao, "Expressive power of parametrized quantum circuits," *Phys. Rev. Res.*, vol. 2, no. 3, p. 033125, Jul. 2020.

[5] J.-M. Liang, S.-Q. Shen, M. Li, and L. Li, "Variational quantum algorithms for dimensionality reduction and classification," *Phys. Rev. A*, vol. 101, no. 3, p. 032323, Mar. 2020.

[6] A. Zulehner and R. Wille, *Introducing Design Automation for Quantum Computing*. Springer, 2020.

[7] D. Volya and P. Mishra, "Quantum spectral clustering of mixed graphs," in *ACM/IEEE Design Automation Conference (DAC)*, 2021, pp. 463–468.

[8] A. Valenti, E. van Nieuwenburg, S. Huber, and E. Greplova, "Hamiltonian learning for quantum error correction," *Phys. Rev. Res.*, vol. 1, no. 3, p. 033092, Nov. 2019.

[9] S.-W. Lee, J. Kim, and W. Son, "Noise-adaptive test of quantum correlations with quasiprobability functions," *Phys. Rev. A*, vol. 102, no. 1, p. 012408, Jul. 2020.

[10] D. Volya and P. Mishra, "Impact of noise on quantum algorithms in noisy intermediate-scale quantum systems," in *IEEE International Conference on Computer Design (ICCD)*, 2020.

[11] ——, "Modeling of noisy quantum circuits using random matrix theory," in *IEEE International Conference on Computer Design (ICCD)*, 2022, pp. 132–138.

[12] ——, "Quantum data compression for efficient generation of control pulses," in *Asia and South Pacific Design Automation Conference (ASPDAC)*, 2023.

[13] J. Preskill, "Quantum Computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, Aug. 2018.

[14] S. Tannu and M. Qureshi, "Mitigating measurement errors in quantum computers by exploiting state-dependent bias," in *MICRO*, 2019.

[15] Y. Chen, M. Farahzad, S. Yoo, and T.-C. Wei, "Detector tomography on IBM quantum computers and mitigation of an imperfect measurement," *Phys. Rev. A*, vol. 100, no. 5, p. 052315, Nov. 2019.

[16] R. Bianchetti, S. Filipp, M. Baur, J. M. Fink, C. Lang, L. Steffen, M. Boissonneault, A. Blais, and A. Wallraff, "Control and tomography of a three level superconducting artificial atom," *Physical review letters*, vol. 105, no. 22, p. 223601, 2010.

[17] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "A quantum engineer's guide to superconducting qubits," *Applied Physics Reviews*, vol. 6, no. 2, p. 021318, Jun. 2019.

[18] P. Gokhale, J. M. Baker, C. Duckering, N. C. Brown, K. R. Brown, and F. T. Chong, "Asymptotic improvements to quantum circuits via qutrits," in *International Symposium on Computer Architecture*, 2019, pp. 554–566.

[19] M. Ringbauer, M. Meth, L. Postler, R. Stricker, R. Blatt, P. Schindler, and T. Monz, "A universal qudit quantum processor with trapped ions," *Nat. Phys.*, vol. 18, no. 9, pp. 1053–1057, Sep. 2022.

[20] L. K. Grover, "A fast quantum mechanical algorithm for database search," Nov. 1996.

[21] P. W. Shor, "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer," *SIAM J. Comput.*, vol. 26, no. 5, pp. 1484–1509, Oct. 1997.

[22] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, "Quantum supremacy using a programmable superconducting processor," *Nature*, vol. 574, no. 7779, pp. 505–510, Oct. 2019.

[23] O. Hosten, N. J. Engelsen, R. Krishnakumar, and M. A. Kasevich, "Measurement noise 100 times lower than the quantum-projection limit using entangled atoms," *Nature*, vol. 529, no. 7587, pp. 505–508, Jan. 2016.

[24] L. A. Martinez, Y. J. Rosen, and J. L. DuBois, "Improving qubit readout with hidden markov models," *Phys. Rev. A*, vol. 102, p. 062426, Dec 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.102.062426

[25] U. Azad and H. Zhang, "Machine learning based discrimination for excited state promoted readout," in *2022 IEEE/ACM 7th Symposium on Edge Computing (SEC)*, 2022, pp. 362–367.

[26] D. Quiroga, P. Date, and R. C. Pooser, "Discriminating Quantum States with Quantum Machine Learning," *ICRC*, pp. 56–63, 2021.

[27] U. Azad and H. Zhang, "Machine learning based discrimination for excited state promoted readout," Tech. Rep., 02 2022.

[28] B. Lienhard, A. Vepsäläinen, L. C. Govia, C. R. Hoffer, J. Y. Qiu, D. Ristè, M. Ware, D. Kim, R. Winik, A. Melville *et al.*, "Deep-neural-network discrimination of multiplexed superconducting-qubit states," *Physical Review Applied*, vol. 17, no. 1, p. 014024, 2022.

[29] P. Duan, Z.-F. Chen, Q. Zhou, W.-C. Kong, H.-F. Zhang, and G.-P. Guo, "Mitigating crosstalk-induced qubit readout error with shallow-neural-network discrimination," *Physical Review Applied*, vol. 16, no. 2, p. 024063, 2021.

[30] J. Kim, B. Oh, Y. Chong, E. Hwang, and D. K. Park, "Quantum readout error mitigation via deep learning," *New Journal of Physics*, vol. 24, no. 7, p. 073009, jul 2022.
[31] É. Genois, J. A. Gross, A. Di Paolo, N. J. Stevenson, G. Koolstra, A. Hashim, I. Siddiqi, and A. Blais, "Quantum-tailored machine-learning characterization of a superconducting qubit," *PRX Quantum*, vol. 2, no. 4, p. 040355, 2021.
[32] P. Duan, Z.-F. Chen, Q. Zhou, W.-C. Kong, H.-F. Zhang, and G.-P. Guo, "Mitigating crosstalk-induced qubit readout error with shallow-neural-network discrimination," *Physical Review Applied*, vol. 16, no. 2, p. 024063, 2021.
[33] R. Navarathna, T. Jones, T. Moghaddam, A. Kulikov, R. Beriwal, M. Jerger, P. Pakkiam, and A. Fedorov, "Neural networks for on-the-fly single-shot state classification," *Applied Physics Letters*, vol. 119, no. 11, 2021.
[34] E. Magesan, J. M. Gambetta, A. Córcoles, and J. M. Chow, "Machine learning for discriminating quantum measurement trajectories and improving readout," *Physical Review Letters*, vol. 114, no. 20, 2015.
[35] D. Gavinsky, J. Kempe, O. Regev, and R. De Wolf, "Bounded-error quantum state identification and exponential separations in communication complexity," in *Proceedings of the thirty-eighth annual ACM symposium on Theory of Computing*, 2006, pp. 594–603.
[36] A. Dvoretzky, J. Kiefer, and J. Wolfowitz, "Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator," *The Annals of Mathematical Statistics*, pp. 642–669, 1956.
[37] P. Massart, "The Tight Constant in the Dvoretzky-Kiefer-Wolfowitz Inequality," *The Annals of Probability*, vol. 18, no. 3, pp. 1269–1283, Jul. 1990.
[38] Qiskit contributors, "Qiskit: An open-source framework for quantum computing," 2019.