

# Quantum state discrimination enhanced by path signature

Shuxiang Cao<sup>1,\*</sup>, Zhen Shao<sup>2</sup>, Jian-Qing Zheng<sup>3</sup>, Mohammed Alghadeer<sup>1</sup>, Simone D Fasciati<sup>1</sup>,

Michele Piscitelli<sup>1</sup>, Sajjad Taravati<sup>1</sup>, Mustafa Bakr<sup>1</sup>, Terry Lyons<sup>2</sup>, and Peter Leek<sup>1</sup>

<sup>1</sup> Clarendon Laboratory, Department of Physics, University of Oxford, Oxford, OX1 3PU, UK

<sup>2</sup> Mathematical Institute, University of Oxford, Oxford, OX2 6GG, UK and

<sup>3</sup> The Kennedy Institute of Rheumatology, University of Oxford, Oxford, OX3 7FY, UK

Quantum state discrimination plays an essential role in quantum technology, crucial for quantum error correction, metrology, and sensing. While conventional methods rely on integrating readout signals or classifying raw signals, we developed a method to extract information about state transitions during readout, based on the path signature method, a tool for analyzing stochastic time series. The hardware experiments demonstrate an improvement in transmon qubit state readout fidelity from  $85.9 \pm 1.0\%$  to  $91.0 \pm 0.5\%$ , without the need for additional hardware. This method has the potential to become a foundational tool for quantum technology.

Rapid high-fidelity single-shot readout is essential for engineering quantum processors. The recent demonstration of quantum error correction highlights that a considerable amount of the error budget is related to the readout process, therefore enhancement in readout can benefit quantum error correction [1]. Dispersive readout stands as one of the most widely-used readout mechanisms [2, 3]. The resonator coupled to the qubit exhibits a frequency shift that depends on the state of the qubit, therefore, the qubit state can be inferred by probing the resonance mode. This mechanism has been demonstrated across various solid-state qubits architectures, including spin-qubits [4, 5], quantum dots [6–8], superconducting qubits [9, 10].

The conventional state discrimination method for dispersive readout integrates the readout signal and produces a single data point that reflects the resonator response at the probing frequency [11]. This method assumes the qubit state remains steady during readout, which is not always the case. When the measurement pulse length is not negligible compared to the qubit relaxation time, the qubit may decay during the measurement [12, 13]. When the probing signal is driven too strongly, it may induce state transitions due to the Stark-shift effect [14, 15]. On the other hand, the dispersive readout signals capture a time series of resonator responses, enabling the potential for continuous tracking of the qubit state. To take this advantage, various standard statistical learning techniques have been employed to analyze the readout signal, such as Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), and Support Vector Machines (SVMs) [16]. In addition, machine learning models such as the Hidden Markov Model (HMM) [17], Feed-forward Neural Networks (FFNN) [18], and Variational Autoencoders (VAE) [19] have been applied to improve the readout fidelity by taking the time-series data as inputs. Although they show improvements, the models are often treated as black boxes and do not emphasize their physical significance.

This study focuses on feature engineering for capturing information about during measurement state transitions. Consider a cumulative sum of the collected signal that forms a trajectory that can be modeled as a stochastic controlled differential equation. The qubit’s state and noise are random variables influencing the trajectory. Motivated by this structure, we employ the stochastic time series tool “path signature” to extract information from the readout signal [20–22]. Our findings show that the signature feature captures information about state transition events during the readout, and by applying this advantage, its application considerably boosts readout fidelity.

Consider a resonator coupled to a qubit, where the resonator has resonance frequency  $\omega_r = \omega_r^{(0)}$  when the qubit is at the ground state. It exhibits a dispersive shift  $\chi_i$  when the qubit is in state  $|i\rangle$ , resulting in the resonator frequency becoming  $\omega_r^{(i)} = \omega_r + \chi_i$ . See Fig.1(a). Consider the probing signal sent to the resonator has frequency  $\omega_p$  with a pulse envelope  $f_t$ , where  $t$  denotes the time. The response signal is a complex number time series, denoted as  $R_t$ . The in-phase channel signal  $I_t = \text{Re}(R_t)$  and the quadrature channel  $Q_t = \text{Im}(R_t)$  are the real and imaginary part of the response signal, respectively. The conventional way of implementing state discrimination is to find the integrated signal  $\tilde{R} = \int R_t f(t) dt$ , and classify  $\tilde{R}$  for different states [9, 10, 23]. See to Fig.1(b). Instead of doing simple integration of the signal, we define a two-dimensional variable  $X_t = \{I_t, Q_t\}$ , and construct a trajectory  $F_t = \int_0^t dX_t$  which is the cumulated sum of  $X_t$ . Now consider the composition of the increment  $dX_t$  which corresponds to the experimentally received signal  $R_t$ . Referring to Fig.1(c), we model the signal to contain a deterministic term  $F_k$  which is dependent on the qubit state, and a noise term as a random variable for Gaussian noise  $d\epsilon_t$ , therefore

$$dX_t^k = dF_t^k + d\epsilon_t \quad (1)$$

where  $X_t^k$  is a random variable that denotes the received signal at time  $t$ , and  $F_t^k$  is the resonator noiseless response at time  $t$ . The signal contains two chan-

\* shuxiang.cao@physics.ox.ac.uk

nels, with  $k \in \{I, Q\}$  representing the in-phase and quadrature-phase channels, respectively.

The signature of a multi-dimensional time-series path is a graded, infinite collection of iterated integrals, where the signature of a path  $X$  with dimension  $d$  up to degree  $N$  is the collection

$$\text{Sig}^N(X) := \left( \int_{0 < t_1 < \dots < t_k < 1} \frac{dX_{i_1}}{dt}(t_1) \cdot \frac{dX_{i_2}}{dt}(t_2) \cdots \frac{dX_{i_k}}{dt}(t_k) dt_1 \cdots dt_k \right)_{\substack{k \leq i_1, \dots, i_k \leq d \\ k=0,1,2,\dots,N}} \quad (2)$$

The signature describes the time series in a global, geometric, and interacting way. For example, the degree 1 signatures are the displacement between the endpoint and the startpoint in different dimensions, and the degree 2 signatures describe the area enclosed by the loop formed by building a curve from the time series of increments and closing it with the chord from the end-point to the start-point. See [20–22] for more information about the signature method.

Signatures of variable  $F_t$  have physical significance. The degree 1 signature is given by  $S^{(1)}(F) = \{F_t - F_0\}$ . Since  $F_0$  is always zero in practice, and  $F_t$  is the integral of the received signal,  $S^{(1)}(F)$  falls back to the traditional approach of integrating the signal. The degree 2 signature corresponds to the Lévy area  $A$  of a 2D path  $\{X_i(t), X_j(t)\}$ , given by

$$A = \frac{1}{2} (S(X)^{i,j} - S(X)^{j,i}), \quad (3)$$

where  $S(X)^{i,j}$  corresponds to the terms in Eq.(2) with  $k = 2$ ,  $i_1 = i, i_2 = j$ . The higher-order signatures generalize the definition of Lévy area to higher dimension volumes.

Consider a state transition that occurred at time  $t_s$  during the measurement. For  $t < t_s$ ,  $dF_t$  remains constant, resulting in a linear trajectory under the assumption of a noiseless signal and negligible ramp-up and ramp-down duration of the signal envelope. For  $t > t_s$ ,  $dX_t$  is a different value and forms discontinuity of the signal, creating a new linear segment starting from the end of the previous trajectory. Referring to Fig.1(d), the shaded area, which appears only if a transition occurs, is captured by the degree 2 signatures. An exception arises when the trajectories before and after  $t_s$  are exactly opposite in directions. To address this, we employ the time-augmentation method during the evaluation of signatures, adding time as an additional dimension to the data, which ensures the detection of state transitions.

To benchmark the performance of the signature approach to characterize the state transition during the measurement of real hardware, we first developed a simulator to generate synthetic traces according to Eq.(1) from a dataset collected on the hardware. The detailed characterization of this device has been reported in a previous work [24, 25]. For the details of the data collection, please refer to App.A. The mean values of the synthetic traces are evaluated from the average values of the experimental dataset and then Gaussian noise is added to the mean value to generate synthetic traces. The Gaussian

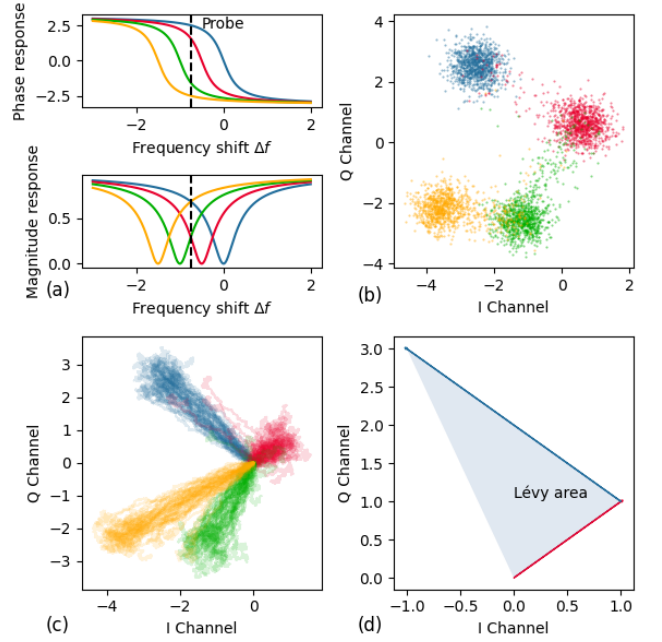


Figure 1: A simulated model for dispersive readout of a transmon coupled to a resonator. The blue, red, green, and yellow colors denote the transmon in the  $|0\rangle$ ,  $|1\rangle$ ,  $|2\rangle$ , and  $|3\rangle$  states, respectively. (a) The phase (top) and amplitude (bottom) response of a resonator. (b) The simulated distribution by integrating the signal on the IQ plane. (c) The constructed trajectory is based on Eq.(1). (d) A generic example of Lévy area of the noiseless signal. The trajectory moves towards the direction where the qubit is in  $|1\rangle$  initially; then it shifts direction to the  $|0\rangle$  state when the qubit decays.

noise’s standard deviation corresponds to the experimental dataset’s standard deviation at each time point. We defined the simulated transition time  $t_s$  for each trace. Before  $t_s$ , traces follows the experimental traces of state  $|1\rangle$ , and after  $t_s$ , they follow the traces of state  $|0\rangle$ . We generated a synthetic dataset of 10,000 traces with a uniform distribution of  $t_s$ . This dataset was randomly divided into training and testing sets and then used to evaluate the performance of linear regression across various depth signatures [26].

The correlation coefficient ( $r^2$ ) values to signature depth are shown in Fig.2(b). The depth- $n$  signature contains all signatures up to degree  $n$ . An upward trend in the  $r^2$  value as the signature depth increases suggests that higher-order signatures contribute additional information for estimating transition time. The  $r^2$  value saturates when depth equals 5. The correlation between simulated and predicted decay times for the depth-5 signature configuration is depicted in Fig.2(a). The clear correlation in the plot demonstrates the effectiveness of using path signatures in modeling transition times.

Fig.2(c) illustrates the model’s application to the experimental dataset, predicting the time that state tran-

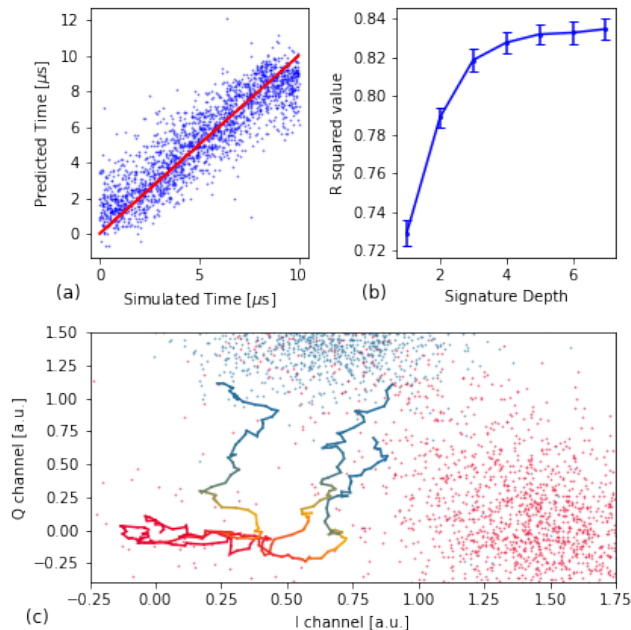


Figure 2: (a) Comparison of simulated transition times with those predicted by an XGBoost regressor [27] trained on a depth-5 signature. This analysis uses a dataset synthesized from experimental data, and the simulated transition times are considered the ground truth for this prediction task. (b) The coefficient of determination ( $r^2$  score) of the linear regressor, trained across various signature depths. (c) Prediction of transition points in a trajectory on the experimental data. This plot contains blue and red point blobs, representing the  $|0\rangle$  and  $|1\rangle$  states, respectively, in a traditional approach for integrating demodulated signals into a single data point. The yellow areas indicate equal probabilities for the  $|0\rangle$  and  $|1\rangle$  states predicted by the regression, with gradients toward red and blue indicating the likelihood of being in the  $|0\rangle$  or  $|1\rangle$  state, respectively.

sition occurs. The presented traces exhibit clear directional changes during measurement, with the yellow color—indicating the state transition point predicted by the model, which aligns closely with these changes. The model, having been trained on a synthetic dataset, is based on the assumption of the existence of exactly one-state transition. The model’s prediction of a state transition at a specific time does not necessarily confirm that such a transition occurred.

Given the above evidence that signature captures additional information indicating the state transition time, we use signatures as a feature set to train the machine learning models for state classifications on the hardware experimental dataset. The traditional integration approach provides the distribution shown in Fig.4(a). We evaluate the depth-5 path signature using LDA to reduce the number of dimensions to 2, and visualize its distri-

bution in Fig.4(b). We have shown that the  $|0\rangle$ ,  $|1\rangle$  and  $|2\rangle$  states are more located in a region, especially since the leakage distributions have disappeared. We plot the log-scale histogram of  $|0\rangle$  and  $|1\rangle$  states in Fig.4(c) and (d) for the integration approach signal and the path signature, respectively.

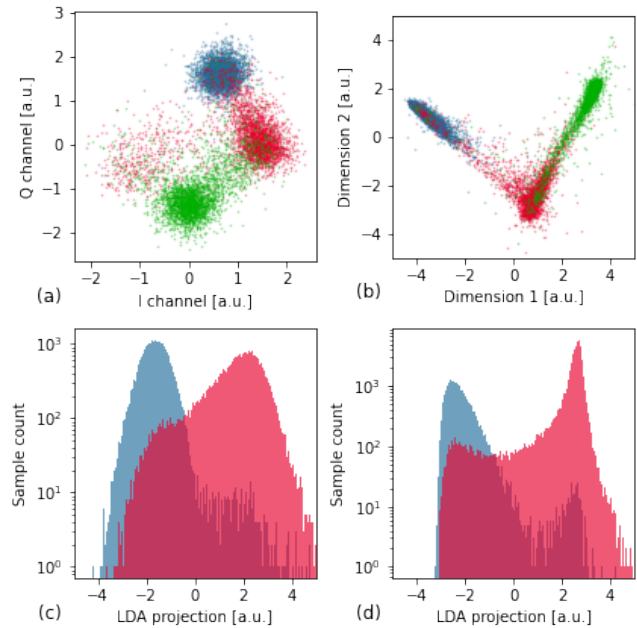


Figure 3: (a) Distribution of experimental dataset on the IQ plane using conventional integration methods. The blue, red and green denotes the state is prepared to  $|0\rangle$ ,  $|1\rangle$ , and  $|2\rangle$  respectively. (b) Projection of a depth-5 signature calculated from the same dataset. The projection direction is evaluated using LDA. (c) Histogram of the integration method projected linearly along the most distinguishable direction using LDA, acting on the data of  $|0\rangle$  and  $|1\rangle$  state only. (d) Histogram of signature features (depth=5) from the same dataset, projected similarly.

The performance of the classifier is measured by assignment accuracy, which is the percentage of correctly identified states compared to the intended prepared states. Each method is benchmarked with variations in the cut-off point of the time series provided to the classifier. See Fig.4(e). Our study demonstrates that combining the random forest (RF) algorithm with the signature method surpasses the performance of the conventional approach of integration followed by the Gaussian Mixture Model (GMM). We selected a linear support vector classifier (SVC) and RF for implementing the classification of the signature features. Notably, the RF performs better than linear SVC, which indicates that the signature features between classes are likely not linearly separable. The performance of the random forest model is enhanced when applied to the path signatures, as opposed to directly using the raw trace signals. The above results indicate

the effectiveness and added value of incorporating path signatures in the analysis.

Finally, we incorporate an extra post-selection in the dataset, to remove the traces where state transition likely occurred during the measurement. This is done by conducting another measurement immediately after the measurement traces are collected for classification analysis, then applying the traditional integration and the GMM method for readout on the collected signals. See Appendix A for more details. We kept only those traces where the last measurement agreed on the state we intended to prepare. When we applied the signature method to this post-selected dataset, we noted an improvement in accuracy at shorter measurement times. However, this enhancement diminished with longer measurement durations. See Fig.4(f). This outcome leads further evidence to the claim that the signature approach is effective by capturing state transitions occurring during the measurement process.

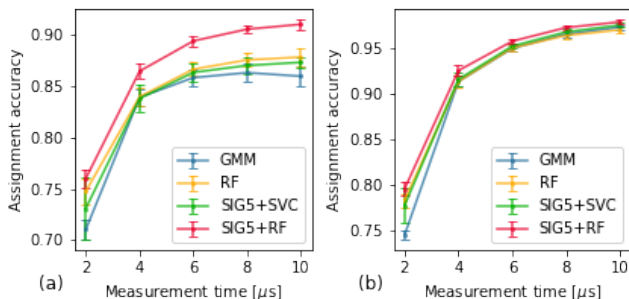


Figure 4: (a) Classification accuracy as a function of measurement length, compared across various classification methods. The tested classification approaches are Gaussian Mixture Model (GMM), Random Forest (RF), Linear support vector classifier on path signature of depth  $x$  (SIG $x$ +SVC), Random forest on path signature of depth  $x$  (SIG $x$ +RF) (b) Classification accuracy, excluding state transitions, which is achieved by implementing an extra measurement for post-selection. See supplementary material for more details.

The signature-based features for dispersive readout provide the following benefits: First, the signature approach offers superior accuracy in state discrimination compared to the standard integration and GMM approach. In addition, compared to other machine learning models that directly process the input signal, signatures can be evaluated during the data integration time, which allows the signature method to be efficiently implemented on Field-Programmable Gate Arrays (FPGAs) [20–22]. This efficiency is essential for enabling fast feedback con-

trol.

This study has certain limitations. Firstly, the system used for this study does not equip a quantum amplifier, which causes long measurement times. While signatures may be helpful in other aspects than state transition tracking, they are not observable here because the state transition dominates the error. Secondly, our approach employs relatively basic machine learning models. The efficacy of the signature method could potentially be enhanced by integrating more advanced models. However, our focus was on identifying simple models that are easily implementable on FPGAs, thus not delving into more complex machine-learning methods.

In the future, there are a few more directions we can continue research using the path signature approach for improvements of dispersive readout. The first is to combine the signature method with multi-frequency probing, which could provide additional dimensional information [28], potentially improving readout fidelity further. In addition, this method shows potential in quantum trajectory studies [29, 30] and weak-measurement experiments [31] for accurately analyzing data traces and tracking state changes.

In summary, the signature-based approach offers considerable improvements for dispersive readout. Combined with its potential for tracking the state transition, it can be marked as a valuable tool for quantum technology.

## ACKNOWLEDGMENTS

This project is supported by the Eric and Wendy Schmidt AI in Science Postdoctoral Fellowship, a Schmidt Future program. S. C. was supported by Schmidt Futures. Z. S. was supported by the EPSRC [EP/S026347/1]. J.-Q. Z. was supported by the Kennedy Trust Prize Studentship [AZT00050-AZ04]. P.L. acknowledges support from the EPSRC b. [EP/T001062/1, EP/N015118/1, EP/M013243/1]. M.B. acknowledges support from the EPSRC QT Fellowship grant [EP/W027992/1]. T. L. was funded in part by the EPSRC [EP/S026347/1], in part by The Alan Turing Institute under the EPSRC [EP/N510129/1], the Data Centric Engineering Programme (under the Lloyd’s Register Foundation grant G0095), the Defence and Security Programme (funded by the UK Government) and in part by the Hong Kong Innovation and Technology Commission (InnoHK Project CIMDA). We thank Youpeng Zhong for the insightful discussion and feedback. The authors would like to acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility in carrying out this work [32].

[1] R. Acharya, I. Aleiner, R. Allen, T. I. Andersen, M. Anshmann, F. Arute, K. Arya, A. Asfaw, J. Atalaya, R. Bab-

bush, D. Bacon, J. C. Bardin, J. Basso, A. Bengtsson,



- S. Boixo, G. Bortoli, A. Bourassa, J. Bovaird, L. Brill, M. Broughton, B. B. Buckley, D. A. Buell, T. Burger, B. Burkett, N. Bushnell, Y. Chen, Z. Chen, B. Chiaro, J. Cogan, R. Collins, P. Conner, W. Courtney, A. L. Crook, B. Curtin, D. M. Debroy, A. Del Toro Barba, S. Demura, A. Dunsworth, D. Eppens, C. Erickson, L. Faoro, E. Farhi, R. Fatemi, L. Flores Burgos, E. Forati, A. G. Fowler, B. Foxen, W. Giang, C. Gidney, D. Gilboa, M. Giustina, A. Grajales Dau, J. A. Gross, S. Habegger, M. C. Hamilton, M. P. Harrigan, S. D. Harrington, O. Higgott, J. Hilton, M. Hoffmann, S. Hong, T. Huang, A. Huff, W. J. Huggins, L. B. Ioffe, S. V. Isakov, J. Iveland, E. Jeffrey, Z. Jiang, C. Jones, P. Juhas, D. Kafri, K. Kechedzhi, J. Kelly, T. Khattar, M. Khezri, M. Kieferová, S. Kim, A. Kitaev, P. V. Klimov, A. R. Klots, A. N. Korotkov, F. Kostritsa, J. M. Kreikebaum, D. Landhuis, P. Laptev, K.-M. Lau, L. Laws, J. Lee, K. Lee, B. J. Lester, A. Lill, W. Liu, A. Locharla, E. Lucero, F. D. Malone, J. Marshall, O. Martin, J. R. McClean, T. McCourt, M. McEwen, A. Megrant, B. Meurer Costa, X. Mi, K. C. Miao, M. Mohseni, S. Montazeri, A. Morvan, E. Mount, W. Mruczkiewicz, O. Naaman, M. Neeley, C. Neill, A. Nersisyan, H. Neven, M. Newman, J. H. Ng, A. Nguyen, M. Nguyen, M. Y. Niu, T. E. O'Brien, A. Opremcak, J. Platt, A. Petukhov, R. Potter, L. P. Pryadko, C. Quintana, P. Roushan, N. C. Rubin, N. Saei, D. Sank, K. Sankaragomathi, K. J. Satzinger, H. F. Schurkus, C. Schuster, M. J. Shearn, A. Shorter, V. Shvarts, J. Skrzynny, V. Smelyanskiy, W. C. Smith, G. Sterling, D. Strain, M. Szalay, A. Torres, G. Vidal, B. Villalonga, C. Vollgraf Heidweiller, T. White, C. Xing, Z. J. Yao, P. Yeh, J. Yoo, G. Young, A. Zalcman, Y. Zhang, N. Zhu, and G. Q. AI, *Nature* **614**, 676 (2023).
- [2] S. Kohler, *Phys. Rev. A* **98**, 023849 (2018).
- [3] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, *Phys. Rev. A* **69**, 062320 (2004).
- [4] A. Crippa, R. Ezzouch, A. Aprá, A. Amisse, R. Laviéville, L. Hutin, B. Bertrand, M. Vinet, M. Urdampilleta, T. Meunier, M. Sanquer, X. Jehl, R. Maurand, and S. De Franceschi, *Nature Communications* **10**, 2776 (2019).
- [5] B. D'Anjou and G. Burkard, *Phys. Rev. B* **100**, 245427 (2019).
- [6] J. I. Colless, A. C. Mahoney, J. M. Hornibrook, A. C. Doherty, H. Lu, A. C. Gossard, and D. J. Reilly, *Phys. Rev. Lett.* **110**, 046805 (2013).
- [7] A. Rossi, R. Zhao, A. S. Dzurak, and M. F. Gonzalez-Zalba, *Applied Physics Letters* **110**, 212101 (2017), [https://pubs.aip.org/aip/apl/article-pdf/doi/10.1063/1.4984224/13493429/212101\\_1\\_online.pdf](https://pubs.aip.org/aip/apl/article-pdf/doi/10.1063/1.4984224/13493429/212101_1_online.pdf).
- [8] G. Burkard and J. R. Petta, *Phys. Rev. B* **94**, 195305 (2016).
- [9] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, *Phys. Rev. Lett.* **95**, 060501 (2005).
- [10] T. Walter, P. Kurpiers, S. Gasparinetti, P. Magnard, A. Potocnik, Y. Salathé, M. Pechal, M. Mondal, M. Oppliger, C. Eichler, and A. Wallraff, *Phys. Rev. Appl.* **7**, 054020 (2017).
- [11] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, *Applied Physics Reviews* **6**, 021318 (2019), [https://pubs.aip.org/aip/apr/article-pdf/doi/10.1063/1.5089550/16667201/021318\\_1\\_online.pdf](https://pubs.aip.org/aip/apr/article-pdf/doi/10.1063/1.5089550/16667201/021318_1_online.pdf).
- [12] S. S. Elder, C. S. Wang, P. Reinhold, C. T. Hann, K. S. Chou, B. J. Lester, S. Rosenblum, L. Frunzio, L. Jiang, and R. J. Schoelkopf, *Phys. Rev. X* **10**, 011001 (2020).
- [13] T. Thorbeck, Z. Xiao, A. Kamal, and L. C. G. Govia, “Readout-induced suppression and enhancement of superconducting qubit lifetimes,” (2023), [arXiv:2305.10508](https://arxiv.org/abs/2305.10508).
- [14] M. Khezri, A. Opremcak, Z. Chen, K. C. Miao, M. McEwen, A. Bengtsson, T. White, O. Naaman, D. Sank, A. N. Korotkov, Y. Chen, and V. Smelyanskiy, *Phys. Rev. Appl.* **20**, 054008 (2023).
- [15] D. Sank, Z. Chen, M. Khezri, J. Kelly, R. Barends, B. Campbell, Y. Chen, B. Chiaro, A. Dunsworth, A. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Mutus, M. Neeley, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, A. Vainsencher, T. White, J. Wenner, A. N. Korotkov, and J. M. Martinis, *Phys. Rev. Lett.* **117**, 190503 (2016).
- [16] E. Magesan, J. M. Gambetta, A. D. Córcoles, and J. M. Chow, *Phys. Rev. Lett.* **114**, 200501 (2015).
- [17] L. A. Martinez, Y. J. Rosen, and J. L. DuBois, *Phys. Rev. A* **102**, 062426 (2020).
- [18] B. Lienhard, A. Vepsäläinen, L. C. Govia, C. R. Hoffer, J. Y. Qiu, D. Ristè, M. Ware, D. Kim, R. Winik, A. Melville, B. Niedzielski, J. Yoder, G. J. Ribeill, T. A. Ohki, H. K. Krovi, T. P. Orlando, S. Gustavsson, and W. D. Oliver, *Phys. Rev. Appl.* **17**, 014024 (2022).
- [19] P. Luchi, P. E. Trevisanutto, A. Roggero, J. L. DuBois, Y. J. Rosen, F. Turro, V. Amitrano, and F. Pederiva, *Phys. Rev. Appl.* **20**, 014045 (2023).
- [20] T. Lyons, in *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. IV* (Kyung Moon Sa, Seoul, 2014) pp. 163–184.
- [21] I. Chevyrev and A. Kormilitzin, arXiv preprint [arXiv:1603.03788](https://arxiv.org/abs/1603.03788) (2016).
- [22] A. Fermanian, T. Lyons, J. Morrill, and C. Salvi, *IEEE BITS the Information Theory Magazine*, 1 (2023).
- [23] J. Heinsoo, C. K. Andersen, A. Remm, S. Krinner, T. Walter, Y. Salathé, S. Gasparinetti, J.-C. Besse, A. Potocnik, A. Wallraff, and C. Eichler, *Phys. Rev. Appl.* **10**, 034040 (2018).
- [24] S. Cao, D. Lall, M. Bakr, G. Campanaro, S. Fasciati, J. Wills, V. Chidambaram, B. Shteynas, I. Rungger, and P. Leek, “Efficient qutrit gate-set tomography on a transmon,” (2022), [arXiv:2210.04857](https://arxiv.org/abs/2210.04857).
- [25] S. Cao, M. Bakr, G. Campanaro, S. D. Fasciati, J. Wills, D. Lall, B. Shteynas, V. Chidambaram, I. Rungger, and P. Leek, “Emulating two qubits with a four-level transmon qudit for variational quantum algorithms,” (2023), [arXiv:2303.04796](https://arxiv.org/abs/2303.04796).
- [26] T. Lyons, “Rough paths, signatures and the modelling of functions on streams,” (2014), [arXiv:1405.4537](https://arxiv.org/abs/1405.4537).
- [27] T. Chen and C. Guestrin, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16 (ACM, New York, NY, USA, 2016) pp. 785–794.
- [28] L. Chen, H. Li, Y. Lu, C. W. Warren, C. J. Krizan, S. Kosen, M. Rommel, S. Ahmed, A. Osman, J. Biznárová, A. Fadavi Roudsari, B. Lienhard, M. Caputo, K. Grigoras, L. Gronberg, J. Govenius, A. F. Kockum, P. Delsing, J. Bylander, and G. Tancredi, *npj Quantum Information* **9** (2023), [10.1038/s41534-023-00689-6](https://doi.org/10.1038/s41534-023-00689-6).

- [29] E. Flurin, L. S. Martin, S. Hacoheh-Gourgy, and I. Siddiqi, *Phys. Rev. X* **10**, 011006 (2020).  
 [30] F. E. van Dorsselaer and G. Nienhuis, *Journal of Optics B: Quantum and Semiclassical Optics* **2**, R25 (2000).  
 [31] R. Flack and B. J. Hiley, *Journal of Physics: Conference Series* **504**, 012016 (2014).  
 [32] A. Richards, (2015), 10.5281/ZENODO.22558.

### Appendix A: Experiment Scheme and Dataset,

The experimental pulse scheme is depicted in Fig.5. It involves three measurements. The qubit state is initially determined using the conventional GMM method, based on data from the first measurement. The traces are then post-selected to ensure that the initial state is the ground state. Subsequently, the qubit is prepared into the  $|0\rangle$ ,  $|1\rangle$ , and  $|2\rangle$  states by applying  $\pi$  pulses for the transitions  $|0\rangle \rightarrow |1\rangle$  and  $|1\rangle \rightarrow |2\rangle$ . The gate fidelity is approximately 99.7%. A second measurement is conducted afterwards, and the traces are recorded for analysis. Following the second measurement, a third measurement is immediately performed. The third measurement aims to identify any state transitions that occur during the second measurement. Transition events are considered to have occurred if the third measurement yields a state different from the intended preparation state. The state discrimination is implemented with the conventional approach, by applying the GMM model on integrated signals.

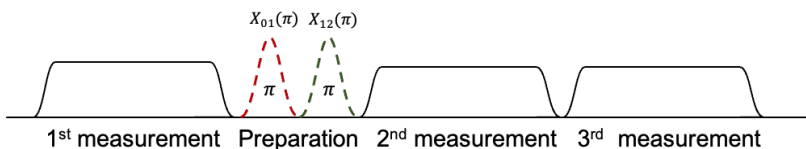


Figure 5: Experimental pulse scheme. There are three measurement pulses involved in the experiment. The first measurement is used to implement post-selection, ensuring the initial state is in the ground state. The second measurement pulse is analyzed using the signature approach, while the third measurement is used to detect if a state transition event occurred during the second measurement.

Each trace was acquired using two analog-digital converters with a sampling rate of 1 Gsps each. The traces have two distinct dimensions (I and Q). The recorded signal data has a carrier frequency of 125 MHz. A short-term Fourier transformation was applied to segments of 256 samples to demodulate the signal at this frequency. Samples of the collected traces are shown in Fig.6.

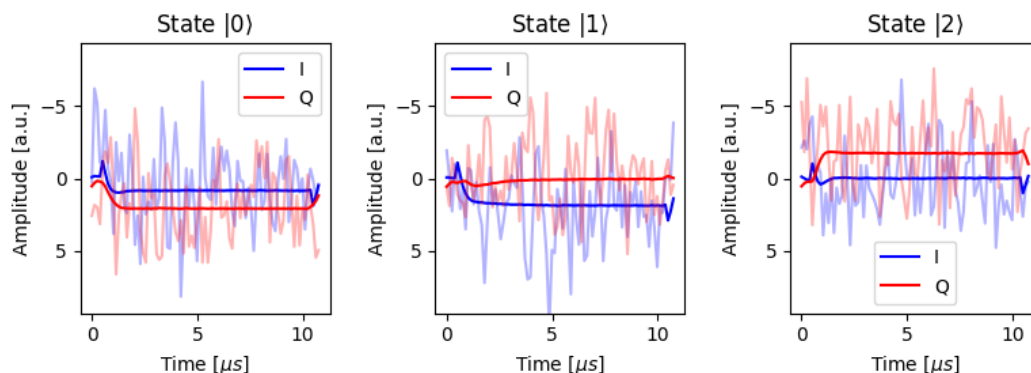


Figure 6: Signal obtained by probing the resonator when the transmon qubit is in states  $|0\rangle$ ,  $|1\rangle$ , and  $|2\rangle$ , respectively. The solid line represents the average result, and the translucent line denotes a single-shot example trace. Blue and red colors correspond to the I and Q channels of the signal, respectively.

Using the described experimental scheme, a database was established containing 70,000 traces for each targeted state, culminating in a total of 210,000 traces. The statistics of the post-selection process are presented in the table below. Here,  $N_i$  represents the number of traces intended for state preparation  $|i\rangle$  that pass the initial post-selection, ensuring the initial state is the ground state.  $\tilde{N}_i$  denotes the number of traces that pass both the initial post-selection and the final post-selection, where the third measurement identifies the state at  $|i\rangle$ . The ratio  $N_i/\tilde{N}_i$  demonstrates a measure of the proportion of the state that remained unchanged during the second measurement.

Prepared state	$N_i$	$\tilde{N}_i$	$\tilde{N}_i/N_i$
$ 0\rangle$	59,768	58,273	97.50%
$ 1\rangle$	52,630	32,414	61.59%
$ 2\rangle$	55,997	32,339	57.75%

For each machine learning classification experiment, 2,000 traces per state were randomly selected from the database, resulting in 6,000 traces per experiment. These traces were then divided into a training set of 4,800 traces and a testing set of 1,200 traces, and the accuracies reported are testing set accuracies. Each hyperparameter configuration was evaluated across 10 repeated experiments with different random seeds for selecting the data from the database and for splitting the training and testing datasets to ensure robust statistical analysis.

## Appendix B: Experiment Setup

The experiment setup of the readout chain is sub-optimal due to the lack of quantum amplifiers. The detailed readout chain is described in Fig. 7.

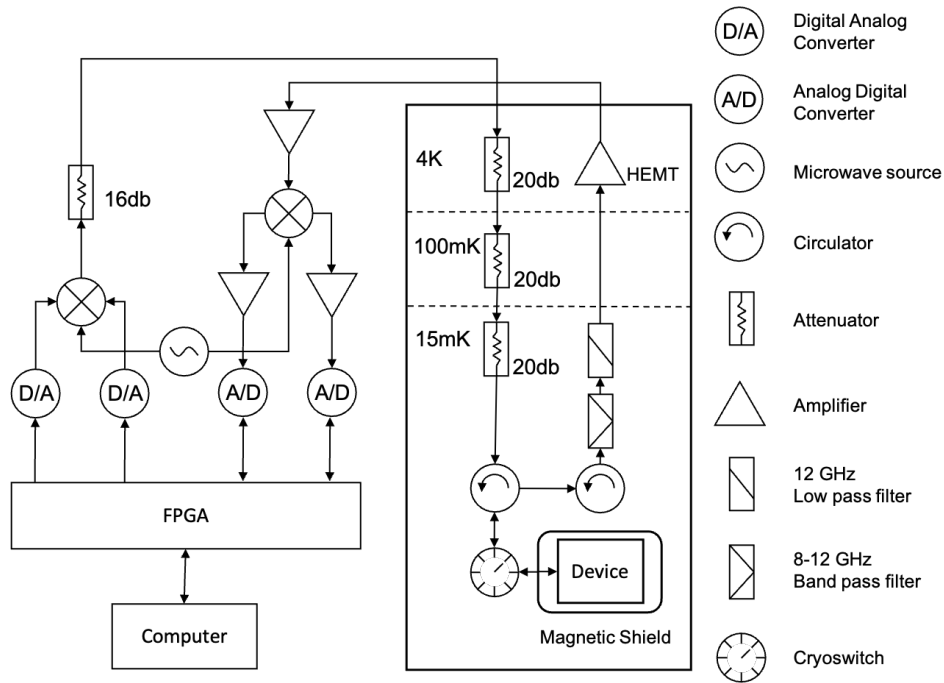


Figure 7: The schematics of the readout chain of the experiment setup.