

# Accelerated characterization of two-level systems in superconducting qubits via machine learning

Avinash Pathapati,<sup>1,2</sup> Olli Mansikkämäki,<sup>1</sup> Alexander Tyner,<sup>1,2</sup> and Alexander Balatsky<sup>1,2</sup>

<sup>1</sup>*Nordita, Stockholm University and KTH Royal Institute of Technology,*

*Hannes Alfvéns väg 12, SE-106 91 Stockholm, Sweden*

<sup>2</sup>*Department of Physics, University of Connecticut, Storrs, Connecticut 06269, USA*

We introduce a data-driven approach for extracting two-level system (TLS) parameters—frequency  $\omega_{\text{TLS}}$ , coupling strength  $g$ , dissipation time  $T_{\text{TLS},1}$ , and the pure dephasing time  $T_{\text{TLS},2}^\varphi$ , labelled as a 4 component vector  $\vec{q}$ , directly from simulated spectroscopy data generated for a single TLS by a form of two-tone spectroscopy. Specifically, we demonstrate that a custom convolutional neural network model(CNN) can simultaneously predict  $\omega_{\text{TLS}}$ ,  $g$ ,  $T_{\text{TLS},1}$  and  $T_{\text{TLS},2}^\varphi$  from the spectroscopy data presented in the form of images. Our results show that the model achieves superior performance to perturbation theory methods in successfully extracting the TLS parameters. Although the model, initially trained on noise-free data, exhibits a decline in accuracy when evaluated on noisy images, retraining it on a noisy dataset leads to a substantial performance improvement, achieving results comparable to those obtained under noise-free conditions. Furthermore, the model exhibits higher predictive accuracy for parameters  $\omega_{\text{TLS}}$  and  $g$  in comparison to  $T_{\text{TLS},1}$  and  $T_{\text{TLS},2}^\varphi$ .

## I. INTRODUCTION

Superconducting qubits offer one of the most promising pathways to build scalable quantum computers. To further advance the performance of qubit based devices, one would need to understand and mitigate the effects that limit the performance. Major limitations for the performance of qubits arise from short coherence times, and this effect is amplified in a quantum chip with large number of qubits. One of the main contributors affecting coherence times has been attributed to the two-level system (TLS) present in the building blocks of the qubits [1–11]. Real-time identification of qubit parameters affected by TLSs, particularly in quantum processors with a large number of qubits, is exceedingly difficult. Streamlining the process of identifying these decoherent qubits will help in producing better quantum chips.

Deep learning [12] has revolutionized various fields, with convolutional neural networks (CNNs) being particularly successful in tasks related to computer vision. CNNs can automatically extract relevant features from images through a sequence of convolution and pooling operations when trained against a required objective. In our case, the input is the spectroscopy data of simulated qubits represented as images and the objective is to predict the four TLS parameters denoted by  $\vec{q} = [\omega_{\text{TLS}}, g, T_{\text{TLS},1}, T_{\text{TLS},2}^\varphi]$ . Here,  $\omega_{\text{TLS}}$  is the TLS frequency,  $g$  is the coupling strength,  $T_{\text{TLS},1}$  is the dissipation time, and  $T_{\text{TLS},2}^\varphi$  is the pure dephasing time. More information on the working of CNNs can be found in e.g., [13]. In our earlier work [14], we developed a two-tone spectroscopy protocol for frequency-fixed qubits and proposed a specific method to extract  $\omega_{\text{TLS}}$  from two-dimensional intensity maps obtained by varying both the drive frequency and pulse duration. As a proof-of-principle, a simple CNN model was used as a benchmark for estimating  $\omega_{\text{TLS}}$ . For more information on this previous work, please see II.

In this work, we extend this approach by proposing a deep learning model that can extract all the TLS parameters  $\vec{q}$  for a single TLS.

The model we put forth in this work is based on ResNet[15]. ResNet is an advanced CNN architecture that was shown to successfully alleviate the problem of vanishing gradients in deep networks [15] during training. ResNet introduced the concept of residual learning, where identity mappings (skip connections) allow gradients to flow more effectively during backpropagation. This trick enables the training of extremely deep networks, significantly improving the performance on image classification benchmarks such as ImageNet [16]. In this paper, we built a custom architecture based on ResNet that can extract  $\vec{q}$  from the spectroscopy images. This method enables swift, automated characterization of faulty qubits and is particularly effective in identifying faulty ones among large qubit arrays, thereby accelerating the cycle of designing better quantum chips.

## II. METHOD

Deep learning models require immense datasets for training. The dataset used for training the model at hand consisted of two-dimensional  $(\omega, t)$  spectroscopy maps from our previous work [14], where  $\omega$  is the drive frequency and  $t$  is the pulse duration. These maps were generated by numerically solving the Lindblad master equation for different

qubit and TLS parameters, with random Gaussian noise added to emulate experimental conditions. The complete dataset construction process is detailed in II A. For the model architecture, ResNet was used as a feature extractor, followed by several fully connected layers (FC) at the end to predict  $\vec{q}$ . Further information on the model architecture and the training procedure is provided in II B and II C respectively.

### A. Dataset

As stated, the model is trained to analyze data generated via two-tone spectroscopy protocol introduced in Ref. [14]. This protocol was put forth to directly determine TLS frequencies and the strength of coupling to the qubit. It is explained in detail in Ref. [14], but is summarized here for clarity.

First, we prepare the system in its ground state,  $|00\rangle$ , by waiting until the population is fully dissipated. The notation  $|ij\rangle$  refers to  $i = 0, 1, 2$  for qubit state and second index refers to  $j = 0, 1$  state of TLS. Second, the transmon is driven with the drive frequency  $\omega_d$  until time  $t_A$ . We label this pulse A. Third, a  $\pi$ -pulse of length  $t_\pi$  is performed at the measured frequency  $\tilde{\omega}_q$ . We label this pulse B. Last, the transmon population is measured. Due to the changing TLS population, the post-pulse B population of the transmon at time  $t_B = t_A + t_\pi$  will vary depending on the time  $t_A$ .

These steps are repeated for a range of drive frequencies  $\omega_d$  and lengths  $t_A$  of the pulse A. The population  $P$  of the first excited level of the transmon after pulse B as a function of the pulse A frequency  $\omega_d$  and length  $t_A$  are then mapped to create the images shown in Fig. (1). The labels mark the state  $|nm\rangle$  of the  $|00\rangle \leftrightarrow |nm\rangle$  transition represented by the feature. The  $|10\rangle$  and  $|20\rangle$  states correspond to the first two excited states of the transmon. The TLS will have three additional transitions: from the ground state to  $|01\rangle$ ,  $|11\rangle$ , and  $|21\rangle$ . The last of these is generally slow compared to typical dissipation times and ignored here. The  $|01\rangle$  corresponds to the two-qubit cross-resonance gate [17, 18]. The  $|11\rangle$  transition corresponds to the bSWAP two-qubit gate [19].

The parameters of the TLSs, i.e., the frequencies  $\tilde{\omega}_k$ , the coupling strengths  $g$ , and the dissipation and pure dephasing times  $T_{1,k} = 1/\gamma_k$  and  $T_{\varphi,k} = 2/\kappa_k$ , all impact the resulting data shown in Fig. (1), creating the opportunity for the machine learning network to extract the TLS parameters and the motivation to use these images as training data.

The training data was generated with the Lindblad master equation

$$\begin{aligned} \frac{d\hat{\rho}}{dt} = & -\frac{i}{\hbar}[\hat{H}, \hat{\rho}] \\ & -\frac{\gamma_q}{2}(\hat{a}^\dagger \hat{a} \hat{\rho} + \hat{\rho} \hat{a}^\dagger \hat{a} - 2\hat{a} \hat{\rho} \hat{a}^\dagger) \\ & -\frac{\kappa_q}{2}(\hat{a}^\dagger \hat{a} \hat{a}^\dagger \hat{a} \hat{\rho} + \hat{\rho} \hat{a}^\dagger \hat{a} \hat{a}^\dagger \hat{a} - 2\hat{a}^\dagger \hat{a} \hat{\rho} \hat{a}^\dagger \hat{a}) \\ & -\frac{\gamma_{\text{TLS}}}{2}(\hat{b}^\dagger \hat{b} \hat{\rho} + \hat{\rho} \hat{b}^\dagger \hat{b} - 2\hat{b} \hat{\rho} \hat{b}^\dagger) \\ & -\frac{\kappa_{\text{TLS}}}{2}(\hat{b}^\dagger \hat{b} \hat{b}^\dagger \hat{b} \hat{\rho} + \hat{\rho} \hat{b}^\dagger \hat{b} \hat{b}^\dagger \hat{b} - 2\hat{b}^\dagger \hat{b} \hat{\rho} \hat{b}^\dagger \hat{b}), \end{aligned} \quad (1)$$

where  $\hat{\rho}$  is the density matrix describing the state of the system,  $\hat{a}$  and  $\hat{b}$  are the bosonic annihilation operators for the qubit and the TLS, respectively.  $\gamma_{q/\text{TLS}} = 1/T_{q/\text{TLS},1}$  and  $\kappa_{q/\text{TLS}} = 1/T_{q/\text{TLS},2}^\varphi$  are the dissipation and pure dephasing rates. The Hamiltonian describing the coherent time evolution is given by

$$\hat{H} = \omega_q \hat{a}^\dagger \hat{a} - \frac{U}{2} \hat{a}^\dagger \hat{a}^\dagger \hat{a} \hat{a} + \omega_{\text{TLS}} \hat{b}^\dagger \hat{b} + g(\hat{a}^\dagger \hat{b} + \hat{b}^\dagger \hat{a}), \quad (2)$$

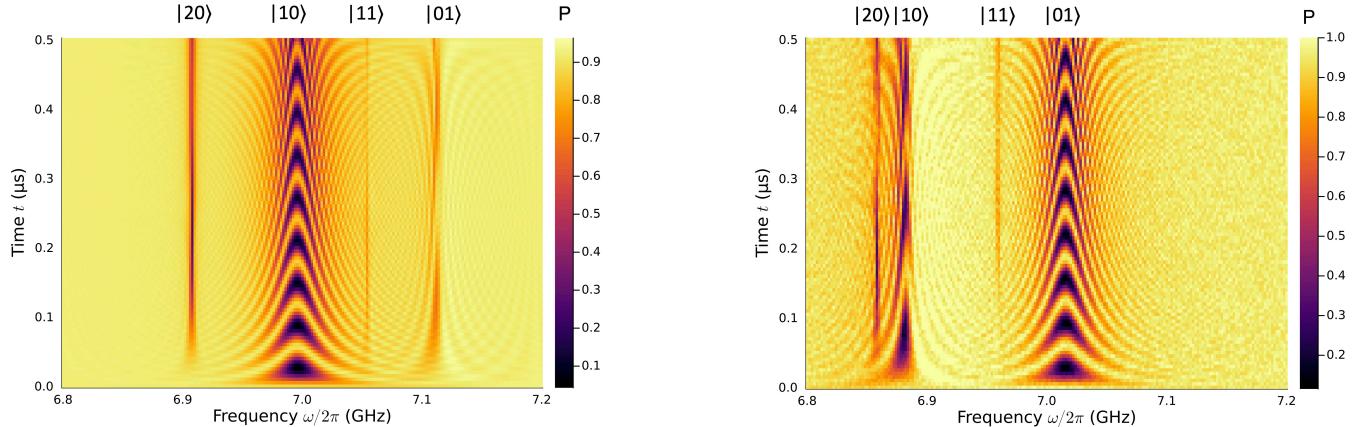
where  $\omega_q$  is the qubit frequency,  $\omega_{\text{TLS}}$  is the TLS frequency,  $U$  is the anharmonicity of the qubit, and  $g$  is the coupling strength. We limit the maximum population of the qubit to two and, as the name suggests, the population of the TLS to one.

Since the detuning  $\Delta = \omega_{\text{TLS}} - \omega_q$  is what matters for the physics of the system, we set  $\omega_q/2\pi = 7\text{GHz}$ . All other parameters are randomized and picked from uniform distributions:  $\omega_{\text{TLS}} \in [6.75\text{GHz}, 7.25\text{GHz}]$ ,  $U/2\pi \in [150\text{MHz}, 300\text{MHz}]$ ,  $g/2\pi \in [5\text{MHz}, 50\text{MHz}]$ ,  $T_{q,1} \in [1\text{ }\mu\text{s}, 10\text{ }\mu\text{s}]$ ,  $T_{q,2}^\varphi \in [0.5\text{ }\mu\text{s}, 5\text{ }\mu\text{s}]$ ,  $T_{\text{TLS},1} \in [0.5\text{ }\mu\text{s}, 10\text{ }\mu\text{s}]$ , and  $T_{\text{TLS},2}^\varphi \in [0.5\text{ }\mu\text{s}, 20\text{ }\mu\text{s}]$ .

The above Lindblad equation models the average time-evolution of a quantum mechanical system under dissipation and dephasing, and therefore predicts a continuous evolution of the qubit population  $P = \langle \hat{a}^\dagger \hat{a} \rangle = \text{Tr}(\hat{\rho} \hat{a}^\dagger \hat{a})$ . In reality, unless the population of the qubit is averaged over a large number of measurements, one would expect to see sharp jumps in the population at essentially random times, resulting in a noisy evolution. Additionally, an experimentally realistic setup would have additional sources of noise, e.g., in the power and frequency of the drive and measurement

errors, that are not captured by our model. Noting that CNNs trained on non-noisy data perform badly on noisy data, we have included a simple noise term on the computed population  $\tilde{P} = P + \delta P$ , where  $\delta P$  is randomly picked from an uniform distribution. The width of the distribution is varied from 0.01 to 0.2 for each sample to represent experimental setups of varying quality.

Figure 1 provides representative examples of spectroscopy images, with and without noise from the dataset. The data discussed in this work always contains only a single strongly coupled TLS to accommodate rapid generation of a large dataset. We briefly outline the steps needed to generate the  $(\omega, t)$  maps from Figure 1 as follows: prepare the system in its ground state  $|00\rangle$ , e.g., by waiting until the population is fully dissipated, then drive the transmon with the drive frequency  $\omega_d$  until time  $t_A$ . We label this pulse A, perform a  $\pi$ -pulse of length  $t_\pi$  at the measured frequency  $\tilde{\omega}_q$ . We label this pulse B and finally measure the transmon population. Due to the changing TLS population, the post-pulse B population of the transmon at time  $t_B = t_A + t_\pi$  will vary depending on the time  $t_A$ ;



**FIG. 1: Synthetic spectroscopic training data.** Examples of synthetic spectroscopy data generated via the Lindblad master equation, eq. (1). The data is represented as an image such that distinct states, labeled  $|nm\rangle$ , are identifiable. This data is generated considering (a) no noise and (b) noise, modeled as a random shift to the qubit population.

## B. CNN model

The architecture of the custom CNN model is illustrated in Figure 2. As discussed, the CNN model employs ResNet as the first layer, extracting the required features from the images. This initial layer is followed by four, independent, fully connected (FC) blocks. Each block consists of multiple FC layers arranged sequentially, with the final output from the block corresponding to one component of  $\vec{q}$ . The FC layers used inside each block transform the output from the ResNet layer into  $\vec{q}$  through a series of linear and non-linear operations. Each FC layer performs this step

$$\mathbf{y} = \sigma(\mathbf{Wx} + \mathbf{b}),$$

where  $\mathbf{x}$  is the input vector,  $\mathbf{W}$  and  $\mathbf{b}$  are the learned weights and biases, respectively, and  $\sigma$  denotes a non-linear activation function. The final output of the network is the parameter vector  $\vec{q}$ . The notation  $\text{FC}(n)$  in the figure 2 represents a stack of  $n$  fully connected layers, applied sequentially. This compact notation is used to indicate depth within each output branch of the network.

## C. Training procedure

The dataset is partitioned into three subsets: a training set (35,000 images), a validation set (15,000 images), and a test set (5,000 images). The output vector  $\vec{q}$  was normalized column-wise to the range [1, 10] using the transformation

$$q_{\text{norm}}^{(i)} = 1 + 9 \cdot \frac{q^{(i)} - q_{\min}^{(i)}}{q_{\max}^{(i)} - q_{\min}^{(i)}}, \quad (3)$$

where  $q^{(i)}$  denotes the values in the  $i$ -th column of the vector  $\vec{q}$ , and  $q_{\min}^{(i)}$  and  $q_{\max}^{(i)}$  are the corresponding minimum and maximum values within that column. This scaling ensures scale-independent learning by preventing parameters with larger magnitudes from disproportionately influencing the loss. Model training is conducted using the training set, while performance on the validation set is monitored to guide the training process. Specifically, we employ an *early stopping* criterion, whereby training is terminated when no further improvement is observed in the validation loss. This approach helps to prevent overfitting and ensures better generalization to unseen data.

The model is trained to minimize the mean squared error (MSE) between the predicted and actual values of the output vector  $\vec{q} = [\omega_{\text{TLS}}, g, T_{\text{TLS},1}, T_{\text{TLS},2}^{\phi}]$

The loss function is defined as:

$$\mathcal{L} = \frac{1}{4 * N} \sum_{i=1}^N \left\| \vec{q}_i - \hat{\vec{q}}_i \right\|^2$$

where  $N$  is the number of training samples,  $\vec{q}_i$  is the ground truth vector for the  $i$ -th sample, and  $\hat{\vec{q}}_i$  is the corresponding prediction by the model.

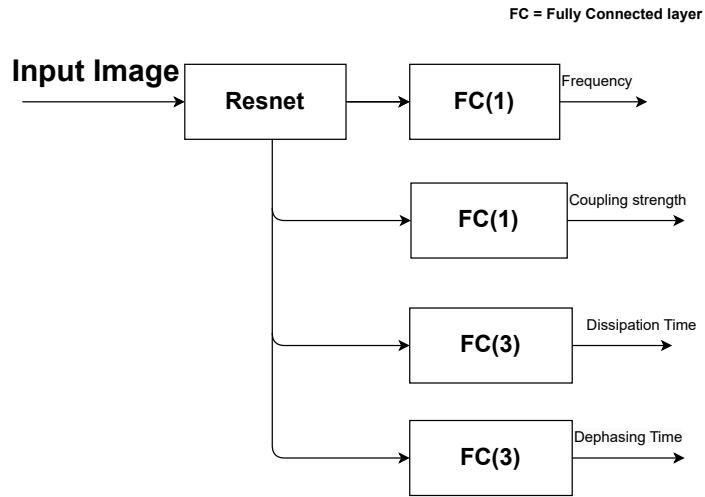


FIG. 2: **Architecture of CNN model.** Schematic architecture of the model used for predicting  $\vec{q}$

### III. RESULTS

Model **M1** is trained under noise-free conditions i.e.  $\delta P = 0$ , Model **M2** is trained on a noise-free dataset ( $\delta P = 0$ ) but tested on the noisy dataset and Model **M3** is trained and tested on the noisy dataset.

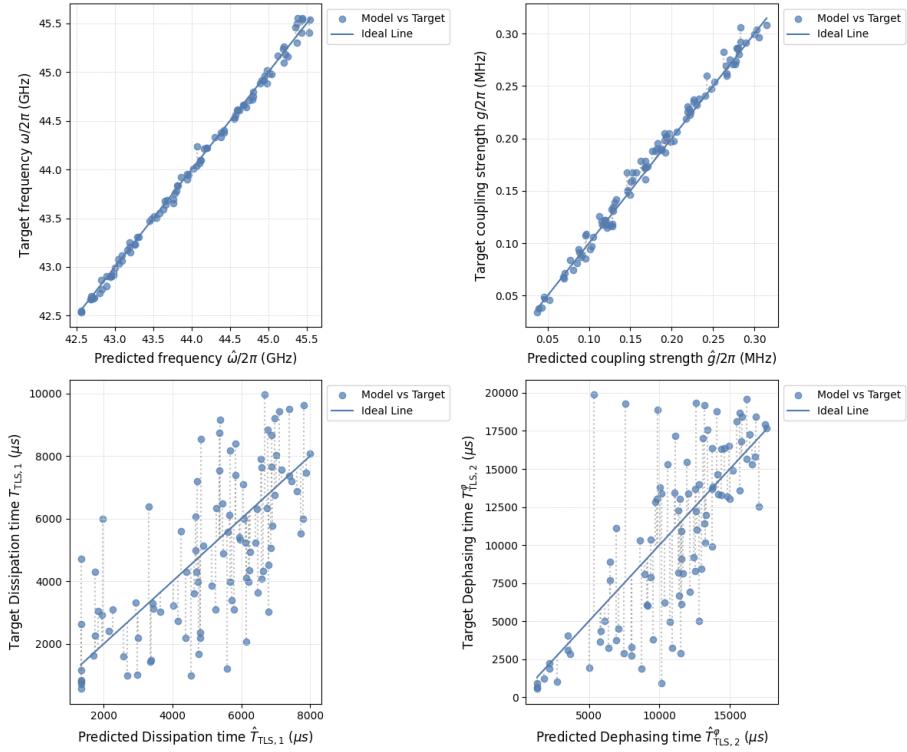
In the case of **M1** model, the model predictions are highly accurate for  $\omega_{\text{TLS}}, g$ . This can be confirmed from the Figure 3 which shows a near-perfect linear correlation between the predicted and expected values. In case of  $T_{\text{TLS},1}$  and  $T_{\text{TLS},2}^{\phi}$ , the predictions show a greater deviation from the ideal line indicating a reduction in performance. This observation is further supported by the MSE loss on the test set, as shown in table I.

TABLE I: Mean squared error (MSE) loss values on the test dataset

Parameter	Model(M1) loss value	Model(M2) loss value	Model(M3) loss value
$\omega$	0.003	8461.5	0.004
$g$	0.00001	4.91	0.00003
$T_{\text{TLS},1}$	4672253.5	3696050176.0	6588405.5
$T_{\text{TLS},2}^{\phi}$	16321636.0	257644380160.0	19207252.0

For model **M2**, there is a significant decline in performance across all four components of  $\vec{q}$ . This is evident from Figure 4 where the predictions ( $\hat{\vec{q}}$ ) deviate significantly from the ideal line( $\vec{q} = \hat{\vec{q}}$ ). To mitigate the impact of noise,

the model **M3** was trained using a dataset that included representative samples of noisy image. This considerably improved the model performance, yielding results(shown in Figure 5) that are comparable to the **M1** model.



**FIG. 3: Model M1: Predictions vs. Targets.** Scatter plots comparing predicted values with ground-truth targets on 100 samples from the test set for a) the TLS frequencies  $\omega_{\text{TLS}}$ , (b) the coupling strengths  $g$ , (c) The dissipation time  $T_{\text{TLS},1}$  and (d) The dephasing time  $T_{\text{TLS},2}^{\varphi}$

#### A. Comparison to the analytic predictions

To establish the usefulness of the CNN approach, we compare the results of the CNN model with those obtained via perturbation theory of the analytical expressions for the TLS frequency  $\omega_{\text{TLS}}$ and coupling strengths  $g$ . We use an automated procedure for estimating the parameters, which works as follows. We take the difference of the maximum and mean populations for each value of the drive frequency  $\omega_d$ , and pick the two highest values near the correct qubit and TLS frequencies  $\omega_q$  and  $\omega_{\text{TLS}}$  taken from the labels. From this, we get the shifted detuning  $\tilde{\Delta} = \tilde{\omega}_{\text{TLS}} - \tilde{\omega}_q$ . Next, we take the vertical slice for the drive frequency at the shifted TLS frequency  $\tilde{\omega}_{\text{TLS}}$ . Then, we fit the function

$$p(t) = c_1 + c_2 e^{-c_3 t} \cos c_4 t + c_5, \quad (4)$$

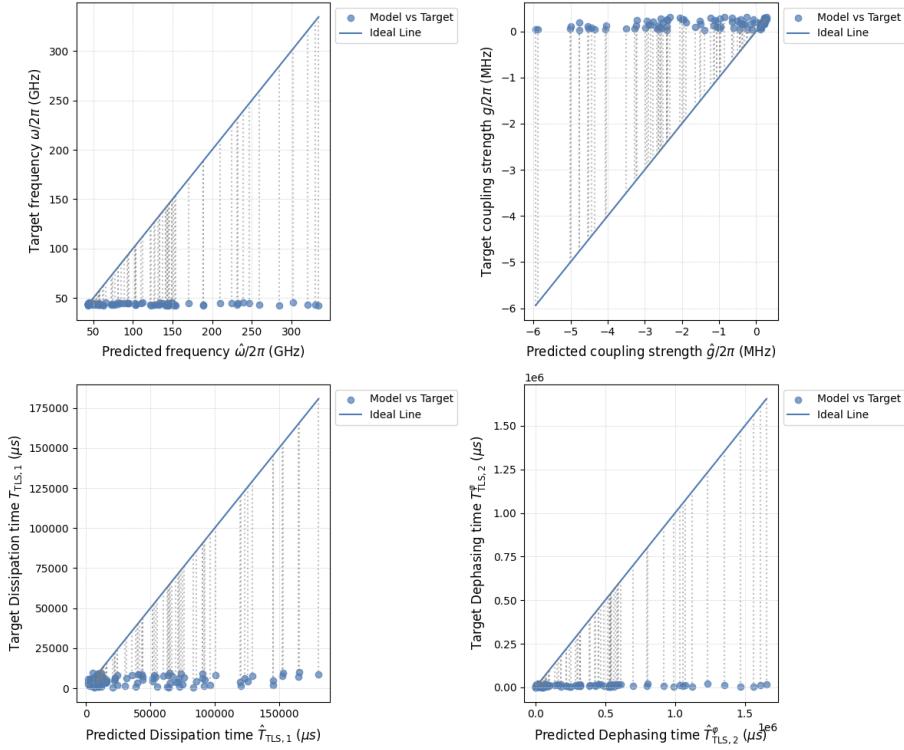
where the fitting parameter  $c_4 = \Omega$  corresponds to the Rabi frequency of the driven transition. With the approximate shifted detuning, Rabi frequency, and the known values of the drive amplitude  $A$  and the anharmonicity  $U$ , we can estimate the coupling strength from

$$g = \frac{\Omega \tilde{\Delta}}{A} \quad (5)$$

and the TLS frequency from

$$\omega_{\text{TLS}} = \tilde{\omega}_{\text{TLS}} + \frac{g^2}{\tilde{\Delta}}. \quad (6)$$

Since the automated procedure does not perfectly pick the peaks corresponding to the correct TLS or qubit frequencies, or fit the function Eq. (4), we pick a set of 200 samples for which the procedure works based on a visual inspection.



**FIG. 4: Model M2: Predictions vs. Targets.** Scatter plots comparing predicted values with ground-truth targets on 100 samples from the test set for a) the TLS frequencies  $\omega_{\text{TLS}}$ , (b) the coupling strengths  $g$ , (c) The dissipation time  $T_{\text{TLS},1}$  and (d) The dephasing time  $T_{\text{TLS},2}^{\varphi}$

This is to avoid comparing the CNN results to estimates where the procedure fails. The mean square errors for the estimates on the 200 samples are  $\text{MSE}_{\omega} = 0.000831$  and  $\text{MSE}_g = 0.0138$ , for the TLS frequency  $\omega_{\text{TLS}}$  and the coupling strength  $g$ , respectively.

Seeing that the analytical expressions are based on perturbation theory, i.e., only valid when  $g \ll \Delta$ , we expect that the analytical expression give the largest errors when the detuning  $\Delta$  is small. This can be seen in Fig. 6 (a). Clearly, this method for estimating the parameters does not work for samples where the TLS frequency is outside the frequency range  $\omega \in [\omega_q - 0.2 \text{ GHz}, \omega_q + 0.2 \text{ GHz}]$ , a limitation not necessarily shared by the CNN model; we omit such samples from the set.

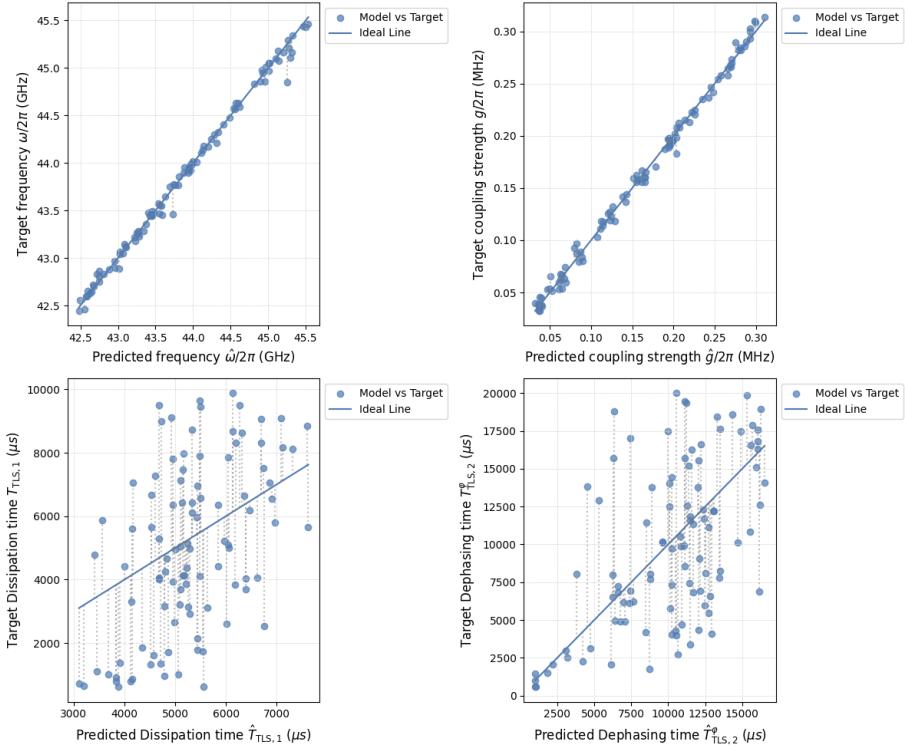
#### IV. DISCUSSION

We have developed a deep learning model that can extract the TLS parameters from spectroscopic data containing a single TLS more efficiently and rapidly than traditional perturbation theory methods. Notably this procedure is successful regardless of whether noise is introduced into the system. In case of noise, training the model on a dataset containing samples with noise significantly improves its performance, enabling it to achieve results comparable to those obtained in the clean limit. The natural application of this protocol would be fast-throughput spectroscopic analysis of qubits for characterization and quality assessment of multiple qubit devices.

For future work, we propose exploring alternative model architectures and conducting extensive hyperparameter tuning to further enhance performance and robustness, as well as extending this approach to enable the extraction of qubit parameters in the presence of multiple strongly coupled TLS.

#### V. ACKNOWLEDGEMENTS

Work of AB, AT and AP was supported by AFOSR program FA9550-25-1-0103, work of OM was supported by Novonordisk Quantum Computing Program. We are grateful to Josu Gomez Beldarrain, Alex Bilmes, Ilya Drozdov,



**FIG. 5: Model M3: Predictions vs. Targets.** Scatter plots comparing predicted values with ground-truth targets on 100 samples from the test set for a) the TLS frequencies  $\omega_{\text{TLS}}$ , (b) the coupling strengths  $g$ , (c) The dissipation time  $T_{\text{TLS},1}$  and (d) The dephasing time  $T_{\text{TLS},2}^{\varphi}$

Vito Iaea, Cameron Kopas, Kasper Grove-Rasmussen, Morten Kjaergaard, Peter Krogstrup, Vincent Michal, Thue Christian Thann Nikoelajs, Dave Pappas, Yaniv Rosen and Mark Svendsen for useful discussions. We thank the Center for Information Technology of the University of Groningen for their support and for providing access to the Hábrók high performance computing cluster.

- [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 (2020).
- [2] C. Wang, C. Axline, Y. Y. Gao, T. Brecht, Y. Chu, L. Frunzio, M. H. Devoret, and R. J. Schoelkopf, Surface participation and dielectric loss in superconducting qubits, *Applied Physics Letters* **107**, 162601 (2015).
- [3] J. M. Gambetta, C. E. Murray, Y.-K.-K. Fung, D. T. McClure, O. Dial, W. Shanks, J. W. Sleight, and M. Steffen, Investigating Surface Loss Effects in Superconducting Transmon Qubits, *IEEE Transactions on Applied Superconductivity* **27**, 1 (2017).
- [4] R. Barends, J. Kelly, A. Megrant, D. Sank, E. Jeffrey, Y. Chen, Y. Yin, B. Chiaro, J. Mutus, C. Neill, P. O’Malley, P. Roushan, J. Wenner, T. C. White, A. N. Cleland, and J. M. Martinis, Coherent Josephson Qubit Suitable for Scalable Quantum Integrated Circuits, *Physical Review Letters* **111**, 080502 (2013).
- [5] C. Müller, J. H. Cole, and J. Lisenfeld, Towards understanding two-level-systems in amorphous solids: insights from quantum circuits, *Reports on Progress in Physics* **82**, 124501 (2019).
- [6] P. Klimov, J. Kelly, Z. Chen, M. Neeley, A. Megrant, B. Burkett, R. Barends, K. Arya, B. Chiaro, Y. Chen, A. Dunsworth, A. Fowler, B. Foxen, C. Gidney, M. Giustina, R. Graff, T. Huang, E. Jeffrey, E. Lucero, J. Mutus, O. Naaman, C. Neill, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. White, S. Boixo, R. Babbush, V. Smelyanskiy, H. Neven, and J. Martinis, Fluctuations of Energy-Relaxation Times in Superconducting Qubits, *Physical Review Letters* **121**, 090502 (2018).
- [7] J. J. Burnett, A. Bengtsson, M. Scigliuzzo, D. Niepce, M. Kudra, P. Delsing, and J. Bylander, Decoherence benchmarking of superconducting qubits, *npj Quantum Information* **5**, 1 (2019).
- [8] S. Schlör, J. Lisenfeld, C. Müller, A. Bilmes, A. Schneider, D. P. Pappas, A. V. Ustinov, and M. Weides, Correlating Decoherence in Transmon Qubits: Low Frequency Noise by Single Fluctuators, *Physical Review Letters* **123**, 190502

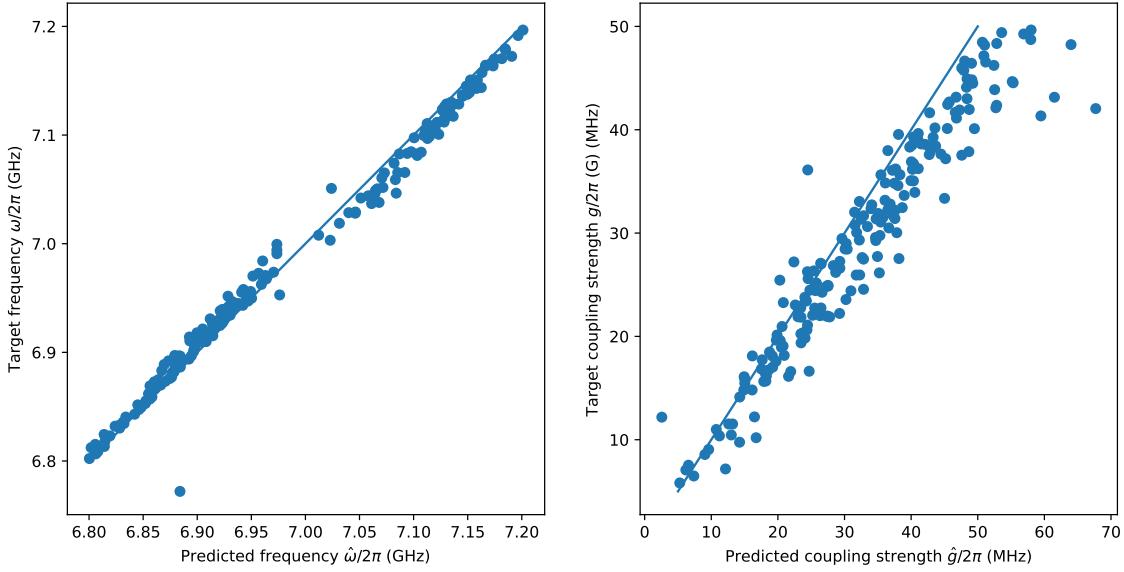


FIG. 6: The analytical estimates vs Target (a) the TLS frequencies  $\omega_{\text{TLS}}$  and (b) the coupling strengths  $g$ .

(2019).

- [9] M. Spiecker, P. Paluch, N. Gosling, N. Drucker, S. Matityahu, D. Gusekova, S. Günzler, D. Rieger, I. Takmakov, F. Valenti, P. Winkel, R. Gebauer, O. Sander, G. Catelani, A. Shnirman, A. V. Ustinov, W. Wernsdorfer, Y. Cohen, and I. M. Pop, Two-level system hyperpolarization using a quantum Szilard engine, *Nature Physics* **19**, 1320 (2023).
- [10] T. Thorbeck, Z. Xiao, A. Kamal, and L. C. Govia, Readout-Induced Suppression and Enhancement of Superconducting Qubit Lifetimes, *Physical Review Letters* **132**, 090602 (2024).
- [11] A. C. Tyner, J. T. Heath, T. C. Thann, V. P. Michal, P. Krogstrup, M. K. Svendsen, and A. V. Balatsky, Identification of soft modes in amorphous  $\text{al}_{2}\text{o}_3$  via first-principles, *Adv. Quantum Technol.*, e2500170.
- [12] Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, *nature* **521**, 436 (2015).
- [13] Y. LeCun, Y. Bengio, *et al.*, Convolutional networks for images, speech, and time series, *The handbook of brain theory and neural networks* **3361**, 1995 (1995).
- [14] O. Mansikkamäki, A. Tyner, A. Bilmes, I. Drozdov, and A. Balatsky, Two-tone spectroscopy for the detection of two-level systems in superconducting qubits, arXiv preprint arXiv:2404.14039 <https://doi.org/10.48550/arXiv.2404.14039> (2024).
- [15] K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition, in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2016) pp. 770–778.
- [16] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, Imagenet: A large-scale hierarchical image database, in *2009 IEEE conference on computer vision and pattern recognition* (Ieee, 2009) pp. 248–255.
- [17] C. Rigetti and M. Devoret, Fully microwave-tunable universal gates in superconducting qubits with linear couplings and fixed transition frequencies, *Physical Review B* **81**, 134507 (2010), publisher: American Physical Society.
- [18] J. M. Chow, A. D. Córcoles, J. M. Gambetta, C. Rigetti, B. R. Johnson, J. A. Smolin, J. R. Rozen, G. A. Keefe, M. B. Rothwell, M. B. Ketchen, and M. Steffen, Simple All-Microwave Entangling Gate for Fixed-Frequency Superconducting Qubits, *Physical Review Letters* **107**, 080502 (2011), publisher: American Physical Society.
- [19] S. Poletto, J. M. Gambetta, S. T. Merkel, J. A. Smolin, J. M. Chow, A. D. Córcoles, G. A. Keefe, M. B. Rothwell, J. R. Rozen, D. W. Abraham, C. Rigetti, and M. Steffen, Entanglement of Two Superconducting Qubits in a Waveguide Cavity via Monochromatic Two-Photon Excitation, *Physical Review Letters* **109**, 240505 (2012).