PAPER • OPEN ACCESS

Noise classification in three-level quantum networks by Machine Learning

IOP Publishing Mach. Learn.: Sci. Technol. 5 (2024) 045049 https://doi.org/10.1088/2632-2153/ad9193

MACHINE LEARNING
Science and Technology

OPEN ACCESS

RECEIVED
9 May 2024

REVISED
17 October 2024

ACCEPTED FOR PUBLICATION
12 November 2024

PUBLISHED
26 November 2024

PAPER

Noise classification in three-level quantum networks by Machine Learning

Shreyasi Mukherjee[1], Dario Penna[2], Fabio Cirinnà[2], Mauro Paternostro[3,4], Elisabetta Paladino[1,5,6], Giuseppe Falci[1,5,6] and Luigi Giannelli[1,5,*]

[1] Dipartimento di Fisica e Astronomia 'Ettore Majorana', Università di Catania, Via S. Sofia 64, 95123 Catania, Italy
[2] Leonardo S.p.A., Cyber & Security Solutions, 95121 Catania, Italy
[3] Università degli Studi di Palermo, Dipartimento di Fisica e Chimica—Emilio Segrè,via Archirafi 36, I-90123 Palermo, Italy
[4] Centre for Theoretical Atomic, Molecular, and Optical Physics, School of Mathematics and Physics, Queens University, Belfast BT7 1NN, United Kingdom
[5] INFN, Sezione di Catania, 95123 Catania, Italy
[6] CNR-IMM, UoS Università, 95123 Catania, Italy
[*] Author to whom any correspondence should be addressed.

E-mail: luigi.giannelli@dfa.unict.it

Keywords: machine learning for quantum, three-level system, noise classification, (non-)Markovianity, noise correlations, quantum network

Abstract

We investigate a machine learning based classification of noise acting on a small quantum network with the aim of detecting spatial or multilevel correlations, and the interplay with Markovianity. We control a three-level system by inducing coherent population transfer exploiting different pulse amplitude combinations as inputs to train a feedforward neural network. We show that supervised learning can classify different types of classical dephasing noise affecting the system. Three non-Markovian (quasi-static correlated, anti-correlated and uncorrelated) and Markovian noises are classified with more than 99% accuracy. On the contrary, correlations of Markovian noise cannot be discriminated with our method. Our approach is robust to statistical measurement errors and retains its effectiveness for physical measurements where only a limited number of samples is available making it very experimental-friendly. Our result paves the way for classifying spatial correlations of noise in quantum architectures.

1. Introduction

Nowadays, quantum systems can be controlled with impressive accuracy [1] to perform new tasks by exploiting their coherence properties, such as superposition and entanglement [2]. However, quantum behavior fades away because of environmental noise that leads to a loss of accuracy of quantum operations and decoherence [3]. The development of strategies to counteract these effects is therefore of crucial importance for progress in quantum technology. In this context, machine learning (ML) is proving to be an innovative and powerful diagnostic tool [4–6]. ML-based approaches have been applied to various quantum control protocols [7–10], to quantum state tomography [11–13], to characterize quantum systems [14–16], to simulate open quantum systems [11, 16–18], to study non-Markovian dynamics [19, 20] and, in a similar spirit similar to this work, to characterize system-environment interactions [21, 22] and perform noise spectroscopy [23–25].

The mitigation of decoherence effects is necessary to achieve quantum advantage. Over the last two decades, several strategies have been developed for both for quantum computation and quantum control in the broader sense. Strategies for error avoidance (or passive stabilization) consist of storing and processing information in suitably designed subspaces of the Hilbert space which are protected from the interaction with the environment [26, 27]. Examples of active stabilization strategies include Quantum Error Correction which relies on encoding nonlocal information [28, 29], and Dynamical decoupling [30–33], which consists in the repeated application of pulsed or switched control, which has been used in solid-state coherent

nanoscience to counteract $1/f$ noise [26, 34–36]. Specific strategies rely on the availability of information about the environment, the characterization of which is therefore a very important step in the development of optimized systems and protocols. For single qubits, this programme has been successfully carried out in various platforms, such as trapped ions [37], photonic qubits [38, 39], nuclear magnetic resonance [40], and nitrogen vacancy centres in diamond [41]. Superconducting systems are an outstanding example in this context [42] whose decoherence times have improved from a few nanoseconds to milliseconds in 15 years [43].

In contrast, the characterization of more complex systems such as multilevel nodes and multi-qubit architectures is still a challenge. In particular, the need to characterize correlations between noises affecting different transitions [44] or different nodes of the quantum architecture [45] is an emerging issue. A strong motivation of principle is that spatially correlated noise between physical qubits can disrupt one of the pillars of digital quantum computation, namely error correction in logical qubits, apart from the fact that correlated errors are indeed observed in noisy intermediate-scale quantum structures [46]. The effects of spatially correlated low-frequency noise between two quantum devices have been observed for several decades [47] and their impact on decoherence has been studied [45] since the first superconducting two-qubit gates were demonstrated. Recently, low-frequency noise correlations between qubit pairs have been characterized by direct measurement of the noise power spectra [48–50]. However, the complete characterization of the noise acting on a controllable quantum system is a difficult task that requires rapidly increasing resources with the upscaling of the systems. In this work, we attempt to take a different route, which is to capture coarse-grained information starting from the categorization into classes of the multivariable noise acting on a quantum architecture. Our main focus is to find an ML-based method that recognizes the presence of noise correlations. We will see that this possibility is intimately related to the (non-)Markovian character of the dynamics. Correlations be further characterized in a second stage using, among others, dynamical decoupling techniques [33] useful in the investigation of the effect that non-Markovianity has in the emergence of many-body phenomena [51, 52].

For simplicity, we illustrate our ML method for identifying correlations in the paradigmatic case of a three-level network affected by diagonal classical noise. Physically, the model describes a three-site quantum network of single-level quantum dots with tunable tunneling amplitudes [53]. The same model describes a qutrit driven by two alternating classical electromagnetic fields. In the former case, we use a control protocol leading to coherent tunneling by adiabatic passage (CTAP) [54, 55], commonly used for population transfer, as a tool to classify noise. We use supervised learning to identify five classes of noise:

(1) Non-Markovian correlated noise;
(2) Non-Markovian anti-correlated noise;
(3) Non-Markovian uncorrelated noise;
(4a) Markovian correlated noise;
(4b) Markovian anti-correlated noise.

We show that by measuring the efficiency of CTAP under three different combinations of pulse amplitudes, we can train a neural network (NN) to classify noise belonging to four of the five classes mentioned above. In particular, our model can accurately classify noise belonging to the first three classes and Markovian noise (classes 4a and 4b together), while it cannot distinguish between correlated and anti-correlated Markovian noise. Referring to real experiments, we analyze in detail how this information can be extracted from a finite number of projective measurements.

The paper is structured as follows: in section 2 we define the models of system and noise that we deal with in our study. In section 3 we present the CTAP protocol and the figures of merit that we use in the training phase. In section 4 we give a brief overview of *supervised learning* and NNs [4, 56–58] and describe how we classify among the noises introduced in section 2. Section 5 presents the results of the classification for a finite number of samples and imperfect measurements. The possible physical implementations are discussed in section 6 together with a physical interpretation of the results. The conclusions are drawn in section 7.

2. Noisy three-level system

To fix the ideas we consider a system of three single-level quantum nodes with eigenbasis $\{|i\rangle,\ i=0,1,2\}$ and on-site energies ϵ_i. This model may describe a system of three quantum dots with an electron tunneling between them [54]. The tunneling rates between the first and the second dot, $\Omega_{\rm p}(t)$, and between the second and the third dot, $\Omega_{\rm s}(t)$, can be controlled by time-dependent external gates. The same model describes a three-level atom driven by two external fields in a multiple rotating frame [59].

The Hamiltonian of the system is $H_{\rm sys}(t) = H_0 + H_{\rm c}(t)$, where H_0 and $H_{\rm c}(t)$ are the free energy and the control term, respectively (throughout the manuscript we use units such that $\hbar = 1$), reading

$$H_0 = \delta_{\rm p}|1\rangle\langle 1| + \delta|2\rangle\langle 2|, \tag{1a}$$

$$H_{\rm c}(t) = \frac{\Omega_{\rm p}(t)}{2}\left(|0\rangle\langle 1| + |1\rangle\langle 0|\right) + \frac{\Omega_{\rm s}(t)}{2}\left(|1\rangle\langle 2| + |2\rangle\langle 1|\right), \tag{1b}$$

where, in the language of the three-dot network, we defined the *detunings* $\delta_{\rm p} = \epsilon_1 - \epsilon_0$ and $\delta = \epsilon_2 - \epsilon_0$.

We consider diagonal noise, i.e. noise affecting the energy levels of the three dots, which may arise from charge noise, for instance [49]. We model it as the following classical stochastic process added to the diagonal entries of the Hamiltonian

$$H_{\rm noise}(t) = \tilde{x}_1(t)\,|1\rangle\langle 1| + \tilde{x}_2(t)\,|2\rangle\langle 2|. \tag{2}$$

Here, $\tilde{x}_i(t)$ $(i = 1, 2)$ is a random variable [60] with normalized probability density $p(x_i, t)$. The total Hamiltonian $H(t)$ in the basis $\{|0\rangle, |1\rangle, |2\rangle\}$ thus reads

$$H(t) = H_{\rm sys}(t) + H_{\rm noise}(t) = \begin{pmatrix} 0 & \Omega_{\rm p}(t)/2 & 0 \\ \Omega_{\rm p}(t)/2 & \delta_{\rm p} + \tilde{x}_1(t) & \Omega_{\rm s}(t)/2 \\ 0 & \Omega_{\rm s}(t)/2 & \delta + \tilde{x}_2(t) \end{pmatrix}. \tag{3}$$

We now link the five classes of noise listed in section 1 to the features of the stochastic mechanism introduced here. We thus consider

- **Non-Markovian types of noise**: When addressing non-Markovian noise, we will make the assumption of *quasistatic processes* where the noise mechanism has a long correlation time and can thus be considered constant over the evolution of the system. The random variables $\tilde{x}_i(t)$ are assumed to be picked from Gaussian distributions with zero mean, thus describing the cumulative effects of independent microscopic sources. Needless to say, while $\tilde{x}_i(t)$ remains constant throughout a single realization of the protocol, it varies between different realizations. We identify the following three classes

  (1) Correlated: $x_2(t) = \eta x_1(t)$ with $\eta > 0$;
  (2) Anti-correlated: $x_2(t) = \eta x_1(t)$ with $\eta < 0$;
  (3) Uncorrelated: $x_2(t)$ and $x_1(t)$, are independent of each other.

- **Markovian types of noise**: The associated dynamics will be ruled by zero-mean, delta-correlated stochastic processes $\tilde{x}_i$ making the dynamics of the system dependent only on its current state rather than its past history. We will thus set

$$\langle \tilde{x}_i(t)\rangle = 0, \quad \langle \tilde{x}_i(t)\tilde{x}_i(t')\rangle = \gamma\delta(t - t') \tag{4}$$

  and consider the two classes

  (4a) Correlated: $x_2(t) = \eta x_1(t)$, with $\eta > 0$;
  (4b) Anti-correlated: $x_2(t) = \eta x_1(t)$, with $\eta < 0$.

3. Dynamics

Our method is based on population transfer by CTAP and the closely related stimulated Raman adiabatic passage (STIRAP) scheme, which has found ample use in atomic physics [59, 61]. We shortly review the CTAP/STIRAP approaches and discuss how we solve the equations and then calculate the classification performances in the presence of the different classes of noise considered.

3.1. CTAP/STIRAP

CTAP [54] (and its spin equivalent spin-CTAP [55]) is a protocol that ideally achieves population transfer from $|0\rangle$ to $|2\rangle$ by adiabatically following a 'trapped state' (often referred to as a dark state) that contains no contribution from the intermediate state $|1\rangle$, which is thus never populated during the protocol. Under ideal conditions, CTAP/STIRAP yields ∼100%-efficiency population transfer with remarkable robustness against parametric fluctuations and state-selectivity.

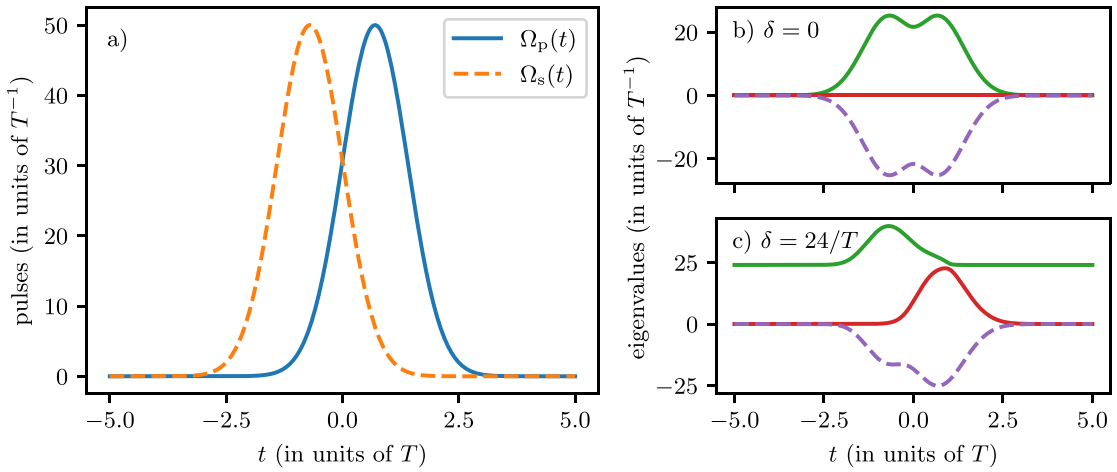

Figure 1. (a) Example of the pump $\Omega_p(t)$ and Stokes $\Omega_s(t)$ pulses, equation (7), as a function of time t. The parameters used are $\Omega_p^{max} = \Omega_s^{max} = 50/T$, $\tau = 0.7T$ and $t \in [-5T, 5T]$. (b) Instantaneous eigenvalues for $\delta = 0$: CTAP/STIRAP achieves 100% population transfer by adiabatically following the trapped state $|\Phi_D(t)\rangle$ (red line). (c) Instantaneous eigenvalues for $\delta \neq 0$: the trapped state is not an instantaneous eigenstate and an adiabatic evolution produces only partial population transfer.

For the transfer process to be successful it is essential to work at small detuning, $\delta \ll \Omega_{p,s}$. In particular, at the so-called two-photon resonance condition $\delta = 0$, the trapped state is an instantaneous eigenstate of the Hamiltonian $H_{sys}(t)$ of the form

$$|\phi_D(t)\rangle = \cos\theta(t)\,|0\rangle - \sin\theta(t)\,|2\rangle \tag{5}$$

with $\theta(t) = \tan^{-1}\left[\Omega_p(t)/\Omega_s(t)\right]$. The protocol is operated in the time interval $[t_i, t_f]$ by applying pulses $\Omega_{p,s}(t)$ in a *counterintuitive* manner [59]: Ω_s is applied *before* Ω_p, while ensuring that the two pulses overlap in a fraction of the duration of the sequence. In this case, the trapped state $|\phi_D(t)\rangle$ at the initial time t_i coincides with $|0\rangle$ and at the final time t_f with the target state $|2\rangle$. If the evolution is adiabatic and the system is prepared in $|0\rangle = |\phi_D(t_i)\rangle$, the system evolves following the trapped state $|\phi_D(t)\rangle$ throughout the evolution as prescribed by the adiabatic theorem [62, 63]. The requested adiabaticity of the process can be cast in the form of the global adiabaticity condition [59]

$$\Omega_{p/s}^{max}\tau \geqslant 10, \tag{6}$$

where τ is the characteristic time scale of the pulses overlap and $\Omega_{p/s}^{max} = \max_t \Omega_{p/s}(t)$.

Under two-photon resonance condition, a counterintuitive sequence satisfying the global adiabaticity request is robust against the actual shape of Ω_p and Ω_s. In fact, several pulse shapes [64] have been used in literature. In this work, we consider Gaussian pulses (see figure 1) of the form

$$\Omega_p(t) = \Omega_p^{max} e^{-\left(\frac{t-\tau}{T}\right)^2}, \qquad \Omega_s(t) = \Omega_s^{max} e^{-\left(\frac{t+\tau}{T}\right)^2}, \tag{7}$$

and we let the system evolve in the time interval $t \in [-5T, 5T]$ with $\tau = 0.7T$. We point out that our aim is not achieving an efficient population transfer, but to exploit the sensitivity of the protocol to get information about the noise affecting the system which deteriorates the efficiency. More details about CTAP/STIRAP can be found in appendix B.

3.2. Figures of merit

The efficiency ξ of the population transfer is defined as the population of the target state $|2\rangle$ at the final time t_f, averaged over all the possible realizations of the noise [65]

$$\xi = \lim_{N\to\infty} \frac{1}{N}\sum_{r=1}^{N} |\langle 2|\psi^{(r)}(t_f)\rangle|^2 = \lim_{N\to\infty} \frac{1}{N}\sum_{r=1}^{N} \xi^{(r)}, \tag{8}$$

with $\xi^{(r)} = |\langle 2|\psi^{(r)}(t_f)\rangle|^2$ the population of $|2\rangle$ for a single trajectory $|\psi^{(r)}(t)\rangle$ identified by an individual realization of noise, i.e. a choice of $\{x_1^{(r)}(t), x_2^{(r)}(t)\}$. The index $r = 1, \ldots, N$ identifies the trajectory being considered.

The quasistatic noise is constant during an individual trajectory thus the index r can be unambiguously mapped to the two real values $x_1^{(r)}$ and $x_2^{(r)}$ of the random processes $\tilde{x}_1$ and $\tilde{x}_2$. Therefore $\xi^{(r)} = \xi(x_1^{(r)}, x_2^{(r)})$. In this case, the efficiency in equation (8) is obtained by averaging over the random processes

$$\xi = \int \xi(x_1, x_2)\, p(x_1, x_2)\, \mathrm{d}x_1 \mathrm{d}x_2. \tag{9}$$

If the noise is (anti-)correlated, then $x_2^{(r)} = \eta x_1^{(r)}$ and equation (9) simplifies to

$$\xi = \int \xi(x_1, \eta x_1)\, p_1(x_1)\, \mathrm{d}x_1. \tag{10}$$

For Markovian noise, the average is calculated via the density matrix $\rho(t)$ of the system which solves a master equation in the Lindblad form (cf appendix A for the derivation)

$$\dot{\rho}(t) = -\mathrm{i}\left[H_0 + H_\mathrm{c}(t), \rho(t)\right] - \frac{\gamma}{2}\left(O^2\rho(t) + \rho(t)\, O^2 - 2O\rho(t)\, O\right), \tag{11}$$

where $O = |1\rangle\langle 1| + \eta|2\rangle\langle 2|$, and γ is the dephasing rate given by $\langle \tilde{x}_1(t)\tilde{x}_1(t')\rangle = \gamma\delta(t-t')$. The efficiency is given by

$$\xi = \langle 2|\rho(t_\mathrm{f})|2\rangle. \tag{12}$$

4. Classification with neural networks

4.1. Neural networks

Supervised learning requires a labeled dataset $\{(\boldsymbol{x}_i, \hat{\boldsymbol{y}}_i)\}_{i=1,\ldots,N}$, where each input feature vector $\boldsymbol{x}_i$ is paired with a target output $\hat{\boldsymbol{y}}_i$. To perform classification, we employ a feedforward neural network, specifically a multilayer perceptron (MLP).

An MLP consists of an input layer, one or more hidden layers, and an output layer. Each neuron computes its output by applying a nonlinear activation function to a weighted sum of the outputs from the previous layer:

$$\boldsymbol{y}^{(l)} = f^{(l)}\left(\boldsymbol{z}^{(l)}\right), \tag{13a}$$

where $f^{(l)}$ is a nonlinear funtion called the *activation function*, and

$$\boldsymbol{z}^{(l)} = \boldsymbol{w}^{(l)}\boldsymbol{y}^{(l-1)} + \boldsymbol{b}^{(l)}. \tag{13b}$$

Here $\boldsymbol{w}^{(l)} \in \mathbb{R}^{D^{(l)} \times D^{(l-1)}}$ is the weight matrix, $\boldsymbol{b}^{(l)} \in \mathbb{R}^{D^{(l)}}$ is the bias vector, and $f^{(l)}$ is the activation function for layer l. The input layer values are set to $\boldsymbol{y}^{(0)} = \boldsymbol{x}$, and the final output is $\boldsymbol{y} = \boldsymbol{y}^{(L)}$.

Training the network involves finding weights $\boldsymbol{W} = \{\boldsymbol{w}^{(l)}\}_l$ and biases $\boldsymbol{B} = \{\boldsymbol{b}^{(l)}\}_l$ that minimize a cost function $C(\{\boldsymbol{y}_i, \hat{\boldsymbol{y}}_i\}_i | \boldsymbol{W}, \boldsymbol{B})$, which quantifies the error between the predicted outputs $\boldsymbol{y}_i$ and the target outputs $\hat{\boldsymbol{y}}_i$. Typically, this minimization process is executed using algorithms such as *stochastic gradient descent* or its variants. The gradients needed for these optimization algorithms are computed through backpropagation, a highly efficient algorithm that leverages the chain rule to propagate errors backward through the network.

We use a neural network consisting of two hidden layers with 128 and 100 neurons, respectively, as detailed in table 1. For the hidden layers, we employ the *leaky rectified linear unit* (LeakyReLU) [66] as the activation function, defined as:

$$f_\mathrm{LReLU}(z) = \begin{cases} z & \text{if} \quad z \geqslant 0, \\ \alpha z & \text{if} \quad z < 0, \end{cases} \tag{14}$$

where α is set to 0.01. This activation function helps to mitigate the 'dying ReLU' problem [67], where neurons become inactive and only output zero.

For the output layer, we use the *softmax* activation function to obtain probabilities over the classes:

$$\boldsymbol{y} = f_\mathrm{softmax}\left(\boldsymbol{z}^{(L)}\right) = \frac{\mathrm{e}^{\boldsymbol{z}^{(L)}}}{\sum_{k=1}^{D^{(L)}} \mathrm{e}^{z_k^{(L)}}}, \tag{15}$$

where the exponentiation and division are performed element-wise. This ensures that the output vector $\boldsymbol{y}$ sums to one, allowing interpretation as class probabilities. We use one-hot encoding for the target outputs $\hat{\boldsymbol{y}}_i$, where each class is represented by a binary vector with a single one corresponding to the correct class.

Table 1. Structure of the neural network used for classification of the 4 or 5 classes of noise.

Layer	# neurons	Activation function
Input	3	
Hidden 1	128	LeakyReLU
Hidden 2	100	LeakyReLU
Output	4 or 5	Softmax

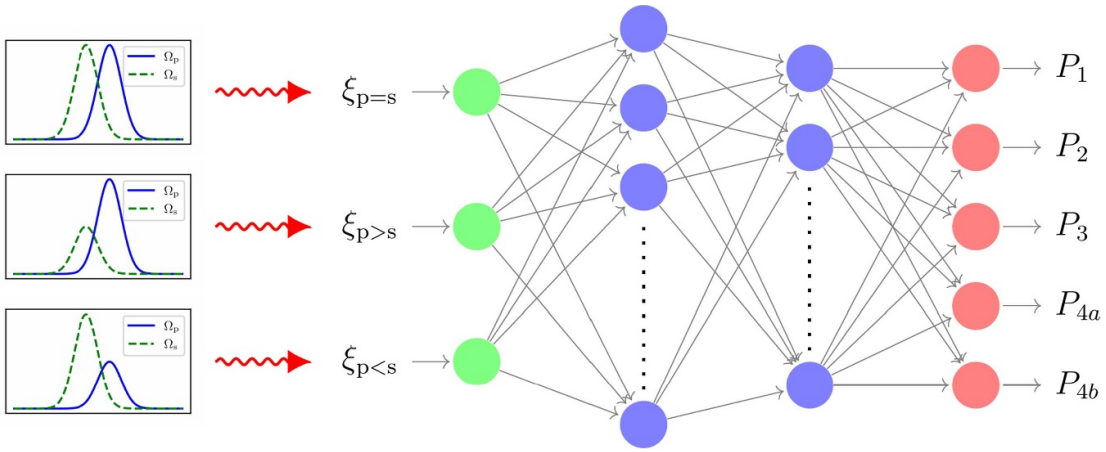

Figure 2. Schematic representation of our model. The efficiencies obtained with the three driving conditions $\Omega_p^{max} = \Omega_s^{max}$, $\Omega_p^{max} > \Omega_s^{max}$, and $\Omega_p^{max} < \Omega_s^{max}$ are given as input to the neural network. The Neural Network is composed of 2 hidden layers with 128 and 100 neurons respectively, and activation function LeakyReLU, equation (14). The output layer composed of 4 or 5 neurons represent the probability of assigning each data to the relative noise class. This is achieved with the softmax activation function, equation (15), on the output layer.

The network is trained by minimizing the *categorical cross-entropy* loss function:

$$C\left(\{\boldsymbol{y}_i, \hat{\boldsymbol{y}}_i\}_i\right) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{D^{(L)}} \hat{y}_{ij} \ln\left(y_{ij}\right), \tag{16}$$

where y_{ij} is the predicted probability for class j of sample i, and $\hat{y}_{ij}$ is the corresponding element of the one-hot encoded target vector.

4.2. Data generation

In order to efficiently classify between the different noise types it is imperative to identify features which are sensitive to the distinct characteristics of the classes introduced in section 2. To this end, we choose as features the efficiency of the CTAP/STIRAP protocol under three different driving conditions. Notice that in this approach, in contrast to other works where the time series of expectation values of several operators are used as input [23, 68, 69], we only need the expectation value of the operator $|2\rangle\langle 2|$ at the final time of the evolution. This makes our method more efficient and appealing for experimental applications. Referring to equation (7) we consider: (i) $\Omega_p^{max} = \Omega_s^{max}$, (ii) $\Omega_p^{max} > \Omega_s^{max}$, (iii) $\Omega_p^{max} < \Omega_s^{max}$. For each type of noise, we solve numerically the dynamics of the system under the three driving conditions and compute the efficiency as described in section 3. We denote the efficiencies as $\xi_{p=s}$, $\xi_{p>s}$, and $\xi_{p<s}$, for the driving conditions (i), (ii), and (iii) respectively, and we use $\boldsymbol{x} = (\xi_{p=s}, \xi_{p>s}, \xi_{p<s})$ as input feature vector to the NN, see figure 2. The parameters we use are (i) $\Omega_p^{max} = \Omega_s^{max} = 50/T$, (ii) $\Omega_p^{max} = 20\sqrt{10}/T > \Omega_s^{max} = 10\sqrt{10}/T$, and (iii) $\Omega_p^{max} = 10\sqrt{10}/T < \Omega_s^{max} = 20\sqrt{10}/T$ with evolution time $t \in [-5T, 5T]$, $\tau = 0.7T$ and $\sigma_1 \approx 17.6/T$ being the standard deviation of the Gaussian distribution $p_1(x_1, t)$.

For the correlated and anti-correlated noise, data are generated by randomly sampling the correlation parameter η in the intervals $[0.1, 5]$ and $[-5, -0.1]$, respectively. For each randomly selected η we calculate three efficiencies, one for each of the pulse conditions (i), (ii), and (iii), thus generating the input feature vector $\boldsymbol{x}$. The efficiencies are calculated using equation (10) for non-Markovian quasistatic noise (with $p_1(x_1, t)$ being a Gaussian distribution) and equation (12) for Markovian noise.

For uncorrelated, non-Markovian quasistatic noise, the values of $\tilde{x}_1$ and $\tilde{x}_2$ are independently sampled from two Gaussian probability distributions $p_1(x_1, t)$ and $p_2(x_2, t)$ with mean $\mu_1 = \mu_2 = 0$. The standard deviation σ_1 of the first distribution is kept fixed, while the standard deviation σ_2 of the distribution $p_2(x_2, t)$

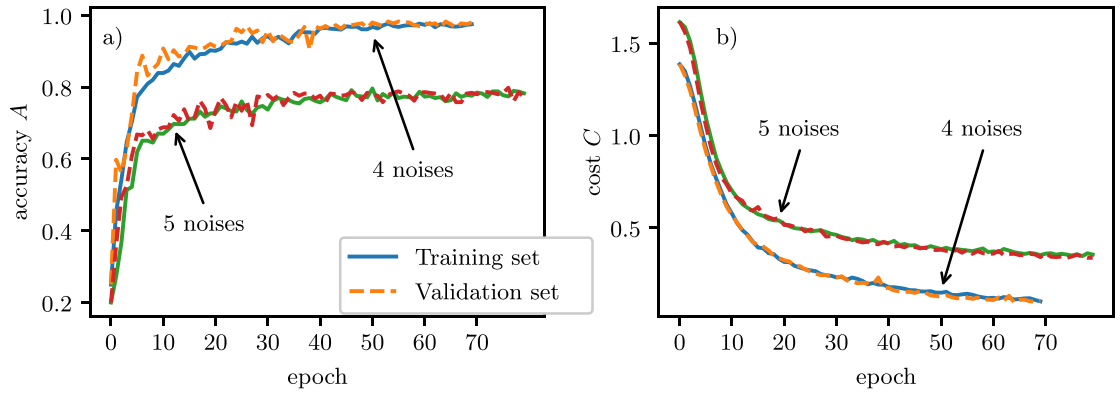

Figure 3. (a) Accuracy, equation (17), and (b) value of the cost function, equation (16), for the training (solid blue for four noises and solid green for five noises) and validation (dashed orange for four noises and dashed red for five noises) sets versus the number of epochs of training. The accuracy on the test set is $A \approx 1$ for four noise classes and $A \approx 0.81$ for five noise classes.

is randomly sampled in the interval $[-5\sigma_1, 5\sigma_1]$. As mentioned in section 4.1, the classes are one-hot encoded, thus the output of the neural network is a layer with four or five neurons. For each noise class, we generate 500 samples such that the total data consists of 2000 or 2500 data points $\{\boldsymbol{x}, \hat{\boldsymbol{y}}\}$. We then split the data in a training, validation and test set with a ratio of 0.6 : 0.2 : 0.2, respectively.

In order to evaluate the efficacy of the model we use the accuracy A defined as the number of correct predictions divided by the number of total predictions N

$$A = \frac{1}{N}\sum_{i=1}^{N} \delta\left(\arg\max \boldsymbol{y}_i, \arg\max \hat{\boldsymbol{y}}_i\right), \tag{17}$$

where the arg max function yields the index j of the maximal component of $\boldsymbol{y}$ and $\delta(\cdot)$ is the Kronecker delta.

5. Results

The ML model and the training are performed with TensorFlow [70] and the results are shown in figure 3. The accuracy A for the training and validation sets is shown in figure 3(a). After training on the test set the accuracy of the model is $A \approx 0.81$ and varies within the values of approx 0.79 and 0.81 depending on the random initialization of the NN weights and the random shuffling of the data for the splitting in the three sets. For the same test set, we report in figure 3(b) how the value of the cost function C changes during training. In figure 4, we show the confusion matrix for the trained model, a tabular representation used to evaluate the performance of a classification model. It cross-tabulates the actual class labels with the model's predictions, providing insight into the correct and incorrect classifications made by the model. Thus it helps to identify which classes are not easily distinguishable from each other posing a challenge to the classification. Each row of the matrix represents the instances of the actual classes, while each column represents the instances of the predicted classes. It is apparent from figure 4(a) that classes 4a and 4b are not distinguishable from each other with the input features of our choice and almost all the samples collapse in class 4a.

We then repeat the same analysis considering the Markovian noise as a single class, i.e. by grouping together correlated and anti-correlated noise and taking $\eta \in [-5, 5]$. As expected, the effectiveness of the model increases reaching an accuracy $A \approx 1$ that varies between $\simeq$0.97 and 1 depending on the random initialization. The accuracy A and the value of the cost function C during the training are shown in figure 3. The confusion matrix for this model is reported in figure 4(b) showing that now the four classes are clearly distinguishable with the chosen input features.

5.1. Results for a finite number of measurement

The analysis performed in the previous section describes the ideal physical situation that the data are obtained from an infinite number of 100%-efficiency of projective quantum measurements. While very large efficiency can be achieved by quantum non-demolition measurements [71], in this section we analyze how the training and the accuracy of the model are affected by the number of measurements M being finite. To this end, we employ the same method described in section 4.2 but compute the efficiency ξ_M as the one resulting from a set of individual 'quantum' trajectories. The result of unravelling each trajectory is simulated by extracting each time the value 1 with probability equal to the population of the target state $|2\rangle$. The output

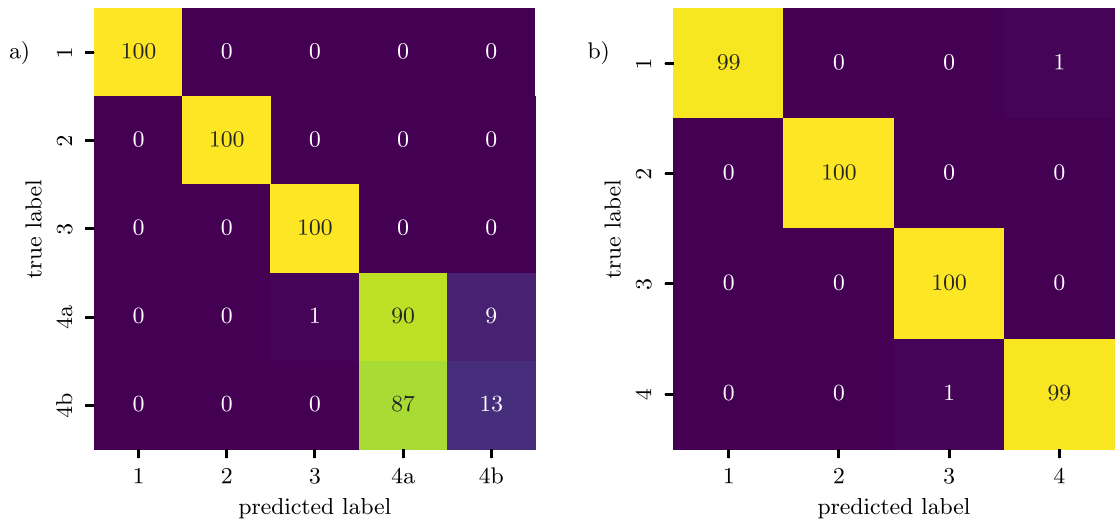

Figure 4. Confusion matrix for the MLP model for the classification of the classes of noise considered in this work. Each row of the matrix represents the instances of the true classes, while each column represents the instances of the predicted classes. It is apparent that classes 4a and 4b, i.e. Markovian correlated and anti-correlated, are not easily distinguishable.

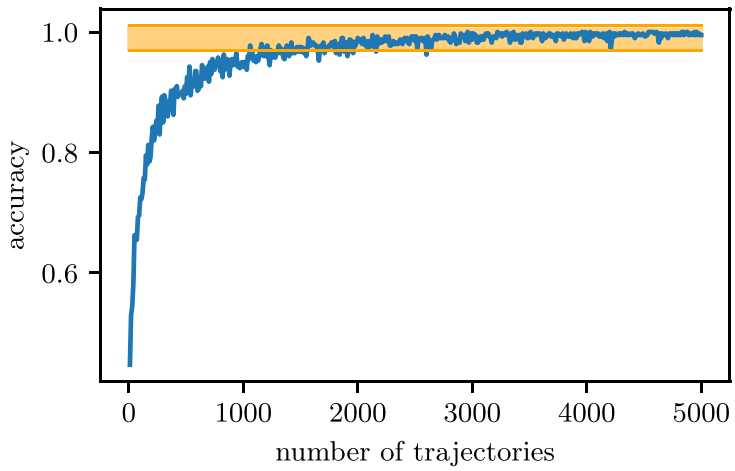

Figure 5. Accuracy of the classification task for four classes of noise versus the number M of projective measurements on quantum trajectories. The shaded orange region corresponds to the accuracy obtained with ideal measurements.

$\boldsymbol{y}$ is produced by averaging over the M quantum trajectories. We produce 500 distinct datasets, each corresponding to a different measurement count $M = 10, \ldots, 5000$ (in steps of 10). For each M, we train the model and evaluate its accuracy using the same methodology as previously described.

The results are presented in figure 5, where we report the accuracy versus the number of measurements. As expected, the accuracy improves for increasing M, approaching the value obtained in the ideal case, which lies in the range $A \approx [0.97, 1]$.

6. Physical implementation and interpretation of the results

6.1. Physical implementations of the model

Spatially correlated noise in solid-state quantum information processing is a topic of current experimental investigation in superconducting [48] and semiconducting [50] quantum devices. Charge- and spin-CTAP have been studied extensively [55, 72] and recently a similar protocol has been implemented in semiconducting quantum dots [73] while CTAP of electromagnetic excitations has been proposed as the building block of quantum operations in the ultrastrong coupling regime [74, 75] circuit-QED architectures. In superconducting devices, STIRAP has been demonstrated in Vee configuration [76–78] whereas the Lambda configuration we study in this work could be implemented either directly [79–81] or by a detuning-modulated protocol [82] which generalizes hyper-STIRAP [59] and bypasses parity selection rules at noise-protected operating points.

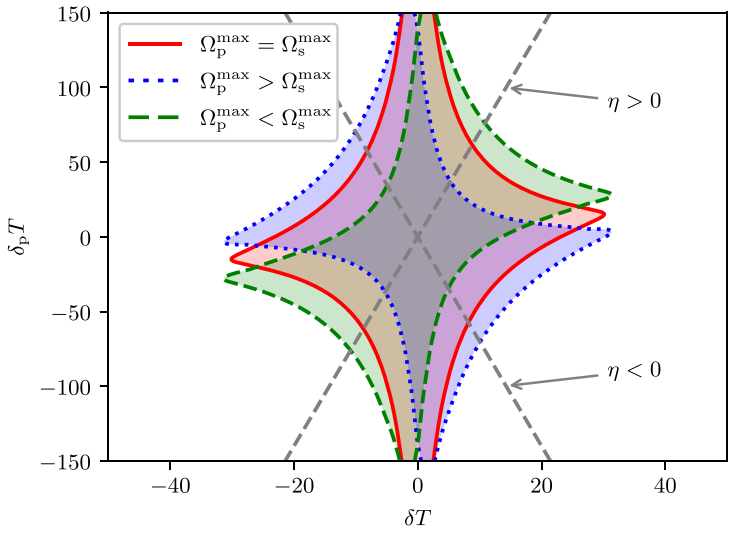

Figure 6. Efficiency of CTAP for the three driving conditions (i) $\Omega_p^{max} = \Omega_s^{max} = 50/T$ (red, solid line), (ii) $\Omega_p^{max} = 20\sqrt{10}/T > \Omega_s^{max} = 10\sqrt{10}/T$ (blue, dotted line), and (iii) $\Omega_p^{max} = 10\sqrt{10}/T < \Omega_s^{max} = 20\sqrt{10}/T$ (green dashed line). The shaded areas within the boundary lines highlight the regions, in the parameter space, where the efficiency is >0.7 (the efficiency being exactly 0.7 along the boundary curves). The efficiencies are calculated with equations (9) and (10). The dashed gray lines represent correlation ($\eta > 0$) or anti-correlation ($\eta < 0$) between the detunings δ and δ_p. The other parameter is $\tau = 0.7T$.

The model of noise studied in this paper accounts for the main effect of sources inducing a fluctuating electric or magnetic polarization of the device. In particular, noise coupled by an operator which commutes with the uncoupled (for CTAP) or undriven (for STIRAP) Hamiltonian is called 'longitudinal' and determines stochastic fluctuations of the bare energy splittings of the devices. In principle, the noise also produces 'transverse' fluctuations of the off-diagonal entries of the Hamiltonian which affect the splittings only at second-order. Moreover, CTAP and STIRAP have the remarkable property that the efficiency of population transfer is almost insensitive to transverse parametric fluctuations [61] which therefore can be ignored. As for the correlations, in CTAP they originate from microscopic noise sources acting on the network in a spatially correlated manner. For STIRAP in a multilevel artificial atom, since noise couples to the device through a single operator it induces fluctuations of the energy level related to each other. In this latter case, the η is given by the ratio of the first derivatives of the energy spectrum at the given bias.

6.2. Interpretation of the results

We now seek an explanation of the ability to classify among differently correlated non-Markovian quasistatic noises. To this end, we analyze the stability plots (efficiency versus the detunings δ and δ_p) for the three driving conditions, shown in figure 6

Ideally, the system is operated at zero detunings, $\delta = \delta_p = 0$. The effect of quasistatic noise is to move, at each individual repetition of the protocol, the point associated with the operations of the system to somewhere else in the space of parameters (cf figure 6). The efficiency for correlated and anti-correlated noise types, is calculated as a weighed average over the line $\delta = \eta\delta_p$, see equation (10). The corresponding efficiency depends on the ratio between Ω_p^{max} and Ω_s^{max}, this dependence being different for $\eta > 0$ or $\eta < 0$. Figure 6 reports as an example the lines $\delta = \eta\delta_p$ for $\eta \gtrless 0$: it is clear that for $\eta > 0$ the efficiency of CTAP with $\Omega_p^{max} < \Omega_s^{max}$ is higher than the one with $\Omega_p^{max} > \Omega_s^{max}$, and vice versa for $\eta < 0$. For a more extensive analysis of this behavior we refer to [83]. For uncorrelated noise, instead, the average is over the whole $\delta - \delta_p$ plane—cf equation (9)—and its dependence on the three different driving conditions is not as explicit.

Such analysis cannot be replicated for Markovian noise, as individual trajectories cannot be represented as single points in figure 6. Instead, the noise assumes all values along the line $\delta = \eta\delta_p$ during each repetition of the protocol.

7. Conclusions

In this work, a ML-based method for categorizing correlated classical noise affecting a three-level quantum system has been presented. We have considered different cases of non-Markovian and Markovian noise acting on a multilevel system focusing on recognizing correlations in the induced fluctuations of the energy splittings.

We have shown that the sensitivity of a CTAP/STIRAP protocol to correlation and Markovianity allow to classify the noise main characteristics in great detail. In particular, data on the efficiency of population transfer under three different driving conditions is sufficient to distinguish between the correlations of non-Markovian quasi-static noise and to discriminate between Markovian and non-Markovian noise. In contrast, this approach does not recognize correlations of Markovian noises. As input to the ML analysis, we have used the numerical solution of the stochastic Schrödinger equation and of the Markovian Lindblad-Master equation, which provide the averages over an infinite number of measurements. Our study was complemented by simulating real projective quantum measurements showing that our approach is robust against errors due to the finite statistics. We stress that the implications of our work go beyond the classification of noise. The methodology developed here provides a new perspective on the use of quantum protocols such as CTAP/STIRAP as a tool for environmental diagnostics. The code developed and the data produced for this work are available in the following github repository [84].

Future directions of this research include exploring the potential of incorporating alternative features of the input to extend the approach to discriminate other classes of noise and to unsupervised learning. In addition, we will investigate the interplay with relaxation [80, 85] and extend the diagnostics to more complex probes, such as interacting multiqubit systems [49], quantum dot chains [86], strongly [87] and ultrastrongly coupled [74, 75] circuit-QED architectures. In these cases control could be implemented by unconventional CTAP/STIRAP protocols operated by detunings [7, 88]. This work and the foreseen advances could provide important elements for developing new integrated robust quantum control and error correction strategies.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://github.com/Shreyasi31/SmallQNetNoiseML.

Acknowledgments

The authors acknowledge Dario Allegra, Lorenzo Catania, Vincenzo Minissale and Giuliano Chiriacò for useful discussions. SM acknowledges support from the project PON Ricerca e Innovazione 2014-2020, React EU, Axis IV, Action IV.4 (CUP: E69J21011430006), within the PhD course in Physics of the Catania University. LG and EP acknowledge support from the PNRR MUR Project PE0000023-NQSTI; GF is supported by the ICSC—Centro Nazionale di Ricerca in High-Performance Computing, Big Data and Quantum Computing. GF and EP acknowledge support from the University of Catania, Piano Incentivi Ricerca di Ateneo 2020-22, Project Q-ICT. EP acknowledges the COST Action SUPERQUMAP (CA 21144). MP acknowledges support from the European Union's Horizon Europe EIC-Pathfinder project QuCoM (Grant No. 101046973), the Royal Society Wolfson Fellowship (Grant No. RSWF/R3/183013), the UK EPSRC (Grant No. EP/T028424/1), and the Department for the Economy Northern Ireland under the US-Ireland R&D Partnership Programme.

Appendix A. Derivation of the Lindblad master equation for Markovian diagonal classical noise

The Lindblad master equation (11) describing of the dynamics of the system subject to Markovian noise is derived as follows. We start from the total Hamiltonian of the system $H = H_{\rm sys} + H_{\rm noise}$, where $H_{\rm noise}$ is given by equation (2). Using the correlation $x_2(t) = \eta x_1(t)$, we write

$$H_{\rm noise} = \tilde{x}_1(t)\left(|1\rangle\langle 1| + \eta|2\rangle\langle 2|\right) = \tilde{x}_1(t)\, O, \tag{A.1}$$

where $O = |1\rangle\langle 1| + \eta|2\rangle\langle 2|$ is the noise operator. We assume that noise has zero mean $\langle \tilde{x}_1(t)\rangle = 0$ and it is Markovian, i.e. $\langle \tilde{x}_1(t)\tilde{x}_1(t')\rangle = \hbar^2\gamma\delta(t-t')$. Since the dynamics of the system is adiabatic and the noise has vanishing correlation time, we can write a master equation for the time-independent system dynamics for each t. This means that we can derive the master equation in an 'instantaneous interaction picture'.

$$\tilde{\rho}(t) = \mathrm{e}^{\frac{\mathrm{i}}{\hbar}H_{\rm sys}t}\rho(t)\,\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}H_{\rm sys}t}, \tag{A.2}$$

where $\rho(t)$ is the density operator in the Schrödinger picture. Physically this approximation amounts in neglecting the backaction of the environment due to the non-adiabatic transitions of the system. The time

evolution of the system is thus given by

$$\dot{\tilde{\rho}}(t) = -\frac{\mathrm{i}}{\hbar}\left[\tilde{H}_{\mathrm{noise}}(t), \tilde{\rho}(t)\right], \tag{A.3}$$

with $\tilde{H}_{\mathrm{noise}}(t) = \mathrm{e}^{\frac{\mathrm{i}}{\hbar}H_{\mathrm{sys}}t}H_{\mathrm{noise}}(t)\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}H_{\mathrm{sys}}t}$. Integrating over the time interval $[0, t]$, we get

$$\tilde{\rho}(t) = \tilde{\rho}(0) - \frac{\mathrm{i}}{\hbar}\int_0^t \left[\tilde{H}_{\mathrm{noise}}(t'), \tilde{\rho}(t')\right] \mathrm{d}t'. \tag{A.4}$$

Substituting $\tilde{\rho}(t)$ in equation (A.3) we obtain

$$\dot{\tilde{\rho}}(t) = -\frac{\mathrm{i}}{\hbar}\mathrm{e}^{\frac{\mathrm{i}}{\hbar}H_{\mathrm{sys}}t}\left[\tilde{x}_1(t)O\rho(0) - \rho(0)O\tilde{x}_1(t)\right]\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}H_{\mathrm{sys}}t} - \frac{1}{\hbar^2}\int_0^t \left[\tilde{H}_{\mathrm{noise}}(t), \left[H_{\mathrm{noise}}(t'), \tilde{\rho}(t')\right]\right] \mathrm{d}t'. \tag{A.5}$$

Averaging over the stochastic variable at time t, we obtain

$$\langle\dot{\tilde{\rho}}(t)\rangle = -\frac{1}{\hbar^2}\left\langle\int_0^t \left[\tilde{x}(t)\tilde{O}(t), \left[\tilde{x}(t')\tilde{O}(t'), \tilde{\rho}(t')\right]\right] \mathrm{d}t'\right\rangle, \tag{A.6}$$

where, $\tilde{O}(t) = \mathrm{e}^{\frac{\mathrm{i}}{\hbar}H_{\mathrm{sys}}t}O\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}H_{\mathrm{sys}}t}$. Taking into account that $t' < t$ and $\tilde{\rho}(t')$ cannot depend on t except for $t' = t$, we can write $\langle\tilde{x}(t)\tilde{x}(t')\tilde{\rho}(t')\rangle = \langle\tilde{x}(t)\tilde{x}(t')\rangle\tilde{\rho}(t')$. We thus obtain

$$\langle\dot{\tilde{\rho}}(t)\rangle = -\frac{\gamma}{2}\left(\tilde{O}^2(t)\tilde{\rho}(t) + \tilde{\rho}(t)\tilde{O}^2(t) - 2\tilde{O}(t)\tilde{\rho}(t)\tilde{O}(t)\right). \tag{A.7}$$

As expected, this equation can be recast in the Lindblad form. In the Schrödinger picture we obtain

$$\langle\dot{\rho}(t)\rangle = -\frac{\mathrm{i}}{\hbar}\left[H_{\mathrm{sys}}, \rho(t)\right] - \frac{\gamma}{2}\left(O^2\rho(t) + \rho(t)O^2 - 2O\rho(t)O\right). \tag{A.8}$$

Appendix B. Adiabatic passage in a three-node network

We consider the Hamiltonian of the three-node network $H_{\mathrm{sys}}(t) = H_0 + H_{\mathrm{c}}(t)$, where H_0 and $H_{\mathrm{c}}(t)$ are the uncoupled and the control parts, respectively ($\hbar = 1$):

$$H_0 = \delta_{\mathrm{p}}|1\rangle\langle 1| + \delta|2\rangle\langle 2|, \tag{B.1a}$$

$$H_{\mathrm{c}} = \frac{\Omega_{\mathrm{p}}(t)}{2}\left(|0\rangle\langle 1| + |1\rangle\langle 0|\right) + \frac{\Omega_{\mathrm{s}}(t)}{2}\left(|1\rangle\langle 2| + |2\rangle\langle 1|\right), \tag{B.1b}$$

and where $\delta_{\mathrm{p}} = \epsilon_1 - \epsilon_0$ and $\delta = \epsilon_2 - \epsilon_0$ define the detunings. CTAP [54] is a protocol that achieves population transfer from $|0\rangle$ to $|2\rangle$ by adiabatic following a trapped state and never populating the intermediate state $|1\rangle$. The atomic analogue is STIRAP [59, 61] where high-fidelity population transfer occurs via a dark state. These protocols have remarkable robustness against parametric fluctuations.

CTAP is based on the counter-intuitive ordering of the pulses $\Omega_{\mathrm{p}}(t)$ and $\Omega_{\mathrm{s}}(t)$. The condition $\delta \approx 0$ is important for the successful population transfer [61]. For $\delta = 0$, the instantaneous eigenstates of $H_{\mathrm{sys}}(t)$ are

$$|\phi_{\mathrm{D}}(t)\rangle = \cos\theta(t)|0\rangle - \sin\theta(t)|2\rangle, \tag{B.2a}$$

$$|\phi_{-}(t)\rangle = \sin\theta(t)\cos\phi(t)|0\rangle - \sin\phi(t)|1\rangle + \cos\theta(t)\cos\phi(t)|2\rangle, \tag{B.2b}$$

$$|\phi_{+}(t)\rangle = \sin\theta(t)\sin\phi(t)|0\rangle - \cos\phi(t)|1\rangle + \cos\theta(t)\sin\phi(t)|2\rangle, \tag{B.2c}$$

with eigenvalues

$$\lambda_{\mathrm{D}}(t) = 0, \tag{B.3a}$$

$$\lambda_{-}(t) = -\frac{\hbar}{2}\sqrt{\Omega_{\mathrm{p}}(t)^2 + \Omega_{\mathrm{s}}(t)^2}\tan\phi(t), \tag{B.3b}$$

$$\lambda_{+}(t) = \frac{\hbar}{2}\sqrt{\Omega_{\mathrm{p}}(t)^2 + \Omega_{\mathrm{s}}(t)^2}\cot\phi(t), \tag{B.3c}$$

where

$$\theta(t) = \tan^{-1}\left(\frac{\Omega_{\mathrm{p}}(t)}{\Omega_{\mathrm{s}}(t)}\right), \tag{B.4}$$

$$\phi(t) = \tan^{-1}\left(\frac{\sqrt{\Omega_{\rm p}(t)^2 + \Omega_{\rm s}(t)^2}}{\delta_{\rm p} + \sqrt{\delta_{\rm p}^2 + \Omega_{\rm p}(t)^2 + \Omega_{\rm s}(t)^2}}\right). \tag{B.5}$$

The eigenstate $|\phi_{\rm D}(t)\rangle$ corresponding to zero eigenvalue is called the *trapped state* since it is trapped in the subspace $\{|0\rangle, |2\rangle\}$. For STIRAP in multilevel atoms it is called *dark state* since it cannot absorb or emit photons despite the external driving. The states $|\phi_-(t)\rangle$ and $|\phi_+(t)\rangle$ are called Autler–Townes states and $|\lambda_+ - \lambda_-|$ the Autler–Townes splitting.

If the pulses are counter-intuitively ordered, i.e. if $\Omega_{\rm s}$ is applied before $\Omega_{\rm p}$:

$$\lim_{t\to t_{\rm i}} \frac{\Omega_{\rm p}(t)}{\Omega_{\rm s}(t)} = \lim_{t\to t_{\rm f}} \frac{\Omega_{\rm s}(t)}{\Omega_{\rm p}(t)} = 0, \tag{B.6}$$

then the dark state $\phi_{\rm D}(t)$ at the initial time $t_{\rm i}$ will coincide with the initial state $|0\rangle$ and at the final time $t_{\rm f}$ will coincide with the target state $|2\rangle$. If the evolution is adiabatic and the system is initially prepared in $|0\rangle = |\phi_{\rm D}(t_{\rm i})\rangle$, by the adiabatic theorem [62, 63], the system evolves following the dark state $|\phi_{\rm D}(t)\rangle$ throughout the time evolution. Thus, the population is transferred from state $|0\rangle$ to state $|2\rangle$ never populating the intermediate state $|1\rangle$. The adiabaticity condition is given by [54, 59]

$$\hbar|\langle\phi_\pm|\dot{\phi}_{\rm D}\rangle| \ll |\lambda_0 - \lambda_\pm|. \tag{B.7}$$

Using the expressions for the dressed basis, equations (B.2) and (B.3), equation (B.7) is recast in terms of the pulses [8, 59, 61]

$$|\dot{\theta}(t)| \ll \frac{1}{2}\left|\delta_{\rm p} \pm \sqrt{\delta_{\rm p}^2 + \Omega_{\rm p}(t)^2 + \Omega_{\rm s}(t)^2}\right|, \tag{B.8}$$

yielding the so-called 'local adiabaticity condition' which has to hold for all times $t \in [t_{\rm i}, t_{\rm f}]$. By time averaging equation (B.8) over τ and assuming $\delta_{\rm p} \ll \Omega_{\rm p}^{\max}, \Omega_{\rm s}^{\max}$ one can obtain the weaker 'global adiabaticity condition' which is often reported as [59]

$$\Omega_{\rm p/s}^{\max}\tau \geqslant 10, \tag{B.9}$$

where τ is the characteristic time scale of the pulses overlap and $\Omega_{\rm p/s}^{\max} = \max_t \Omega_{\rm p/s}(t)$.

In this work, we perform the population transfer using Gaussian pulses of the form

$$\Omega_{\rm p}(t) = \Omega_{\rm p}^{\max} e^{-\left(\frac{t-\tau}{T}\right)^2}, \tag{B.10a}$$

$$\Omega_{\rm s}(t) = \Omega_{\rm s}^{\max} e^{-\left(\frac{t+\tau}{T}\right)^2}, \tag{B.10b}$$

and we let the system evolve in the time interval $t \in [-5T, 5T]$ with $\tau = 0.7T$, see figure 1 for an example of pulses.

ORCID iDs

Mauro Paternostro https://orcid.org/0000-0001-8870-9134
Elisabetta Paladino https://orcid.org/0000-0002-9929-3768
Giuseppe Falci https://orcid.org/0000-0001-5842-2677
Luigi Giannelli https://orcid.org/0000-0001-9704-7304

References

[1] Koch C P *et al* 2022 Quantum optimal control in quantum technologies. Strategic report on current status, visions and goals for research in Europe *EPJ Quantum Technol.* **9** 19
[2] Acín A *et al* 2018 The quantum technologies roadmap: a European community view *New J. Phys.* **20** 080201
[3] Zurek W H 2003 Decoherence, einselection and the quantum origins of the classical *Rev. Mod. Phys.* **75** 715–75
[4] Marquardt F 2021 Machine learning and quantum devices *SciPost Phys. Lect. Notes* **29**
[5] Krenn M, Landgraf J, Foesel T and Marquardt F 2023 Artificial intelligence and machine learning for quantum technologies *Phys. Rev.* A **107** 010101
[6] Gebhart V, Santagati R, Gentile A A, Gauger E M, Craig D, Ares N, Banchi L, Marquardt F, Pezzè L and Bonato C 2023 Learning quantum systems *Nat. Rev. Phys.* **5** 1–16
[7] Brown J, Sgroi S, Giannelli L, Paraoanu G S, Paladino E, Falci G, Paternostro M and Ferraro A 2021 Reinforcement learning-enhanced protocols for coherent population-transfer in three-level quantum systems *New J. Phys.* **23** 093035

[8] Giannelli L, Sgroi S, Brown J, Paraoanu G S, Paternostro M, Paladino E and Falci G 2022 A tutorial on optimal control and reinforcement learning methods for quantum technologies *Phys. Lett.* A **434** 128054
[9] Niu M Y, Boixo S, Smelyanskiy V N and Neven H 2019 universal quantum control through deep reinforcement learning *npj Quantum Inf.* **5** 1–8
[10] Sgroi S, Palma G M and Paternostro M 2021 Reinforcement learning approach to nonequilibrium quantum thermodynamics *Phys. Rev. Lett.* **126** 020601
[11] Banchi L, Grant E, Rocchetto A and Severini S 2018 Modelling non-Markovian quantum processes with recurrent neural networks *New J. Phys.* **20** 123030
[12] Torlai G, Mazzola G, Carrasquilla J, Troyer M, Melko R and Carleo G 2018 Neural-network quantum state tomography *Nat. Phys.* **14** 447–50
[13] Palmieri A M, Kovlakov E, Bianchi F, Yudin D, Straupe S, Biamonte J D and Kulik S 2020 Experimental neural network enhanced quantum tomography *npj Quantum Inf.* **6** 1–5
[14] Couturier R, Dionis E, Guérin S, Guyeux C and Sugny D 2023 Characterization of a driven two-level quantum system by supervised learning *Entropy* **25** 446
[15] Genois É, Gross J A, Di Paolo A, Stevenson N J, Koolstra G, Hashim A, Siddiqi I and Blais A 2021 Quantum-tailored machine-learning characterization of a superconducting qubit *PRX Quantum* **2** 040355
[16] Youssry A, Paz-Silva G A and Ferrie C 2020 Characterization and control of open quantum systems beyond quantum noise spectroscopy *npj Quantum Inf.* **6** 95
[17] Bandyopadhyay S, Huang Z, Sun K and Zhao Y 2018 Applications of neural networks to the simulation of dynamics of open quantum systems *Chem. Phys.* **515** 272–8
[18] Luo D, Chen Z, Carrasquilla J and Clark B K 2022 Autoregressive neural network for simulating open quantum systems via a probabilistic formulation *Phys. Rev. Lett.* **128** 090501
[19] Luchnikov I A, Vintskevich S V, Grigoriev D A and Filippov S N 2020 Machine learning non-Markovian quantum dynamics *Phys. Rev. Lett.* **124** 140502
[20] Fanchini F F, Karpat G, Rossatto D Z, Norambuena A and Coto R 2021 Estimating the degree of non-Markovianity using machine learning *Phys. Rev.* A **103** 022425
[21] Papič M and de Vega I 2022 Neural-network-based qubit-environment characterization *Phys. Rev.* A **105** 022605
[22] Wise D F, Morton J J L and Dhomkar S 2021 Using deep learning to understand and mitigate the qubit noise environment *PRX Quantum* **2** 010316
[23] Barr J, Zicari G, Ferraro A and Paternostro M 2024 Spectral density classification for environment spectroscopy *Mach. Learn.: Sci. Technol.* **5** 015043
[24] Martina S, Buffoni L, Gherardini S and Caruso F 2021 Learning the noise fingerprint of quantum devices (arXiv:2109.11405 [quant-ph])
[25] Martina S, Hernández-Gómez S, Gherardini S, Caruso F and Fabbri N 2023 Deep learning enhanced noise spectroscopy of a spin qubit environment *Mach. Learn.: Sci. Technol.* **4** 02LT01
[26] Paladino E, Galperin Y M, Falci G and Altshuler B L 2014 1/F noise: implications for solid-state quantum information *Rev. Mod. Phys.* **86** 361–418
[27] Vion D, Aassime A, Cottet A, Joyez P, Pothier H, Urbina C, Esteve D and Devoret M H 2002 Manipulating the quantum state of an electrical circuit *Science* **296** 886–9
[28] Roffe J 2019 Quantum error correction: an introductory guide *Contemp. Phys.* **60** 226
[29] Campbell E 2024 A series of fast-paced advances in quantum error correction *Nat. Rev. Phys.* **6** 160
[30] Viola L, Knill E and Lloyd S 1999 Dynamical decoupling of open quantum systems *Phys. Rev. Lett.* **82** 2417–21
[31] Falci G, D'Arrigo A, Mastellone A and Paladino E 2004 Dynamical suppression of telegraph and $1/f$ noise due to quantum bistable fluctuators *Phys. Rev.* A **70** 040101
[32] Suter D and Álvarez G A 2016 Colloquium: Protecting quantum information against environmental noise *Rev. Mod. Phys.* **88** 041001
[33] D'Arrigo A, Piccitto G, Falci G and Paladino E 2024 Open-loop quantum control of small-size networks for high-order cumulants and cross-correlations sensing (arXiv:2401.05766 [cond-mat.other])
[34] Falci G, Hakonen P J and Paladino E 2024 1/f noise in quantum nanoscience *Encyclopedia of Condensed Matter Physics* 2nd edn, ed T Chakraborty (Academic) pp 1003–17
[35] Balandin A A, Paladino E and Hakonen P J 2024 Electronic noise—from advanced materials to quantum technologies *Appl. Phys. Lett.* **124** 050401
[36] Pellegrino F M D, Falci G and Paladino E 2020 1/f critical current noise in short ballistic graphene Josephson junctions *Commun. Phys.* **3** 1–8
[37] Biercuk M J, Uys H, VanDevender A P, Shiga N, Itano W M and Bollinger J J 2009 Optimized dynamical decoupling in a model quantum memory *Nature* **458** 996
[38] Damodarakurup S, Lucamarini M, Di Giuseppe G, Vitali D and Tombesi P 2009 Experimental inhibition of decoherence on flying qubits via "bang-bang" control *Phys. Rev. Lett.* **103** 040502
[39] Orieux A, D'Arrigo A, Ferranti G, Franco R L, Benenti G, Paladino E, Falci G, Sciarrino F and Mataloni P 2015 Experimental on-demand recovery of entanglement by local operations within non-Markovian dynamics *Sci. Rep.* **5** 8575
[40] Jenista E R, Stokes A M, Branca R T and Warren W S 2009 Optimized, unequal pulse spacing in multiple echo sequences improves refocusing in magnetic resonance *J. Chem. Phys.* **131** 204510
[41] Naydenov B, Dolde F, Hall L T, Shin C, Fedder H, Hollenberg L C L, Jelezko F and Wrachtrup J 2011 Dynamical decoupling of a single-electron spin at room temperature *Phys. Rev.* B **83** 081201
[42] Bylander J, Gustavsson S, Yan F, Yoshihara F, Harrabi K, Fitch G, Cory D G, Nakamura Y, Tsai J-S and Oliver W D 2011 Noise spectroscopy through dynamical decoupling with a superconducting flux qubit *Nat. Phys.* **7** 565–70
[43] Kjaergaard M, Schwartz M E, Braumüller J, Krantz P, Wang J I-J, Gustavsson S and Oliver W D 2020 Superconducting qubits: current state of play *Annu. Rev. Condens. Matter Phys.* **11** 369–95
[44] Sung Y *et al* 2021 Multi-level quantum noise spectroscopy *Nat. Commun.* **12** 967
[45] D'Arrigo A, Mastellone A, Paladino E and Falci G 2008 Effects of low-frequency noise cross-correlations in coupled superconducting qubits *New J. Phys.* **10** 115006
[46] Vepsäläinen A P *et al* 2020 Impact of ionizing radiation on superconducting qubit coherence *Nature* **584** 551–6

[47] Zorin A B, Ahlers F-J, Niemeyer J, Weimann T, Wolf H, Krupenin V A and Lotkhov S V 1996 Background charge noise in metallic single-electron tunneling devices *Phys. Rev.* B **53** 13682–7
[48] von Lüpke U *et al* 2020 Two-qubit spectroscopy of spatiotemporally correlated quantum noise in superconducting qubits *PRX Quantum* **1** 010305
[49] Yoneda J, Rojas-Arias J S, Stano P, Takeda K, Noiri A, Nakajima T, Loss D and Tarucha S 2023 Noise-correlation spectrum for a pair of spin qubits in silicon *Nat. Phys.* **19** 1793–8
[50] Zou J, Bosco S and Loss D 2023 Spatially correlated classical and quantum noise in driven qubits: the good, the bad, and the ugly (arXiv:2308.03054 [quant-ph])
[51] Chiriacò G, Tsitsishvili M, Poletti D, Fazio R and Dalmonte M 2023 Diagrammatic method for many-body non-Markovian dynamics: memory effects and entanglement transitions *Phys. Rev.* B **108** 075151
[52] Tsitsishvili M, Poletti D, Dalmonte M and Chiriacò G 2024 Measurement induced transitions in non-Markovian free fermion ladders *SciPost Phys. Core* **7** 011
[53] Pope T J, Rajendran J, Ridolfo A, Paladino E, Pellegrino F M D and Falci G 2019 Coherent trapping in small quantum networks *J. Stat. Mech.* 124024
[54] Greentree A D, Cole J H, Hamilton A R and Hollenberg L C L 2004 Coherent electronic transfer in quantum dot systems using adiabatic passage *Phys. Rev.* B **70** 235317
[55] Gullans M J and Petta J R 2020 Coherent transport of spin by adiabatic passage in quantum dot arrays *Phys. Rev.* B **102** 155404
[56] Burkov A 2019 *The Hundred-Page Machine Learning Book* vol 1 (Andriy Burkov)
[57] Géron A 2023 *Hands-On Machine Learning With Scikit-Learn, Keras and Tensorflow: Concepts, Tools and Techniques to Build Intelligent Systems* (O'Reilly)
[58] Goodfellow I, Bengio Y and Courville A 2016 *Deep Learning* (*Adaptive Computation and Machine Learning*) (The MIT Press)
[59] Bergmann K, Theuer H and Shore B W 1998 Coherent population transfer among quantum states of atoms and molecules *Rev. Mod. Phys.* **70** 1003–25
[60] Mandel L and Wolf E 1995 *Optical Coherence and Quantum Optics* (Cambridge University Press)
[61] Vitanov N V, Rangelov A A, Shore B W and Bergmann K 2017 Stimulated Raman adiabatic passage in physics, chemistry and beyond *Rev. Mod. Phys.* **89** 015006
[62] Born M and Fock V 1928 Beweis des adiabatensatzes *Z. Phys.* **51** 165–80
[63] Messiah A 1961 *Quantum Mechanics* vol 1/2 (North-Holland)
[64] Giannelli L and Arimondo E 2014 Three-level superadiabatic quantum driving *Phys. Rev.* A **89** 033419
[65] Gardiner C W and Zoller P 2010 *Quantum Noise: A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods With Applications to Quantum Optics* (*Springer Series in Synergetics*) (Springer)
[66] Maas A L, Hannun A Y and Ng A Y 2013 Rectifier nonlinearities improve neural network acoustic models *ICML Workshop on Deep Learning for Audio, Speech and Language Processing (WDLASL 2013)*
[67] Douglas S C and Yu J 2018 Why RELU units sometimes die: analysis of single-unit error backpropagation in neural networks *2018 52nd Asilomar Conf. on Signals, Systems and Computers* (IEEE) pp 864–8
[68] Guo Y, Hu Z, Zhang J, Zhu C and Guo X 2023 High-speed photon correlation monitoring of amplified quantum noise by chaos using deep-learning balanced homodyne detection *Appl. Phys. Lett.* **123** 051101
[69] Zeng Y-X, Gebremariam T, Shen J, Xiong B and Li C 2021 Application of machine learning for predicting strong phonon blockade *Appl. Phys. Lett.* **118** 164003
[70] Abadi M *et al* (Google Research) 2015 TensorFlow: Large-scale Machine Learning on Heterogeneous Systems (arXiv:1603.04467)
[71] Braginsky V B, Khalili F Y and Thorne K S 1992 *Quantum Measurement* (Cambridge University Press)
[72] Menchon-Enrich R, Benseny A, Ahufinger V, Greentree A D, Busch T and Mompart J 2016 Spatial adiabatic passage: a review of recent progress *Rep. Prog. Phys.* **79** 074401
[73] Kandel Y P, Qiao H, Fallahi S, Gardner G C, Manfra M J and Nichol J M 2021 Adiabatic quantum state transfer in a semiconductor quantum-dot spin chain *Nat. Commun.* **12** 2156
[74] Falci G, Ridolfo A, Di Stefano P G and Paladino E 2019 Ultrastrong coupling probed by coherent population transfer *Sci. Rep.* **9** 9249
[75] Giannelli L, Paladino E, Grajcar M, Paraoanu G S and Falci G 2024 Detecting virtual photons in ultrastrongly coupled superconducting quantum circuits *Phys. Rev. Res.* **6** 013008
[76] Kumar K S, Vepsäläinen A, Danilin S and Paraoanu G S 2016 Stimulated Raman adiabatic passage in a three-level superconducting circuit *Nat. Commun.* **7** 10628
[77] Xu H K *et al* 2016 Coherent population transfer between uncoupled or weakly coupled states in ladder-type superconducting qutrits *Nat. Commun.* **7** 11018
[78] Gong M, Yu M, Chu Y, Chen W, Cao Q, Wang N, Cai J, Betzholz R and Giannelli L 2024 Two-photon-transition superadiabatic passage in a nitrogen-vacancy center in diamond *Phys. Rev.* A **109** 032626
[79] Siewert J, Brandes T and Falci G 2009 Advanced control with a Cooper-pair box: stimulated Raman adiabatic passage and Fock-state generation in a nanomechanical resonator *Phys. Rev.* B **79** 024504
[80] Falci G, La Cognata A, Berritta M, D'Arrigo A, Paladino E and Spagnolo B 2013 Design of a Lambda system for population transfer in superconducting nanocircuits *Phys. Rev.* B **87** 214515
[81] Earnest N *et al* 2018 Realization of a Λ system with metastable states of a capacitively shunted fluxonium *Phys. Rev. Lett.* **120** 150504
[82] Di Stefano P G, Paladino E, Pope T J and Falci G 2016 Coherent manipulation of noise-protected superconducting artificial atoms in the Lambda scheme *Phys. Rev.* A **93** 051801
[83] Falci G, Berritta M, Russo A, D'Arrigo A and Paladino E 2012 Effects of low-frequency noise in driven coherent nanodevices *Phys. Scr.* **T151** 014020
[84] Mukherjee S and Giannelli L 2024 SmallQNetNoiseML (available at: https://github.com/Shreyasi31/SmallQNetNoiseML)
[85] Spagnolo B, Caldara P, La Cognata A, Augello G and Valenti D 2012 Relaxation phenomena in classical and quantum systems *Acta Phys. Pol.* B **43** 1169
[86] Mills A R, Zajac D M, Gullans M J, Schupp F J, Hazard T M and Petta J R 2019 Shuttling a single charge across a one-dimensional array of silicon quantum dots *Nat. Commun.* **10** 1063
[87] Blais A, Girvin S M and Oliver W D 2020 Quantum information processing and quantum optics with circuit quantum electrodynamics *Nat. Phys.* **16** 247–56
[88] Di Stefano P G, Paladino E, D'Arrigo A and Falci G 2015 Population transfer in a Lambda system induced by detunings *Phys. Rev.* B **91** 224506