UNIVERSITÀ
degli STUDI
di CATANIA

Dipartimento
di Fisica
e Astronomia
"Ettore Majorana"

PhD programme in physics

Quantum sensing of noise based on Machine Learning

Shreyasi Mukherjee

PhD thesis

Supervisor:
Prof. Luigi Giannelli

Co-Supervisors:
Prof. Mauro Paternostro
Dr. Fabio Cirinnà

academic year 2024/2025

Acknowledgements

The years of my PhD have been both challenging and incredibly rewarding. I am deeply grateful to all the people who have influenced me and this work in one way or another.

First and foremost, I would like to thank my supervisor, Prof. Luigi Giannelli. His guidance, inspiring ideas, and constant support have been fundamental to every project I worked on. I am truly thankful for everything he has taught me, and for the immense patience and encouragement he has shown throughout this journey. I am also grateful to our group leader Prof. Giuseppe Falci, who supervised me during the first one and a half years of my PhD. He introduced me to the fascinating world of open quantum systems, and for that, I will always be grateful. His deep knowledge and insightful ideas have enriched the entire group including me, leaving a lasting impact on our way of thinking and approaching research.

I would also like to thank my collaborators with whom I had the pleasure of publishing: Prof. Mauro Paternostro, Prof. Elisabetta Paladino, Fabio Cirinnà, and Dario Penna. A special thank you goes to Dario Fasone, with whom I thoroughly enjoyed working on a project.

Over these years, I've had the good fortune to meet many wonderful people in Prof. Falci's group. I warmly thank my colleagues and friends: Giuliano, Nicola, Alessandro, Vincenzo, Federico, Rosa, and Noemi. Your presence made the workdays lighter and the memories richer.

Now, to my family, thank you from the bottom of my heart. To my mother, Shayamasri Mukherjee, and my father, Dr. Nityananda Mukher-

jee: you are my strong roots, and because of you, I stand tall. Your love, strength, and unwavering belief in me have kept me grounded.

Among my Indian friends, too many to name, I especially want to thank Indranil Maiti, Kuntal Samanta and Amrita Gupta for their friendship and support.

I acknowledge 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.

Contents

Abstract

Noise-induced decoherence is a central challenge in quantum technologies. This thesis explores machine learning (ML)-based methods for characterizing classical noise in quantum systems, with the goal of improving reliability of quantum technologies.

In the first part, we study a three-level quantum networks driven by adiabatic control protocols (CTAP/STIRAP) and show that simple neural networks can classify different spatial correlations in non-Markovian noises and between non-Markovian and Markovian noise classes using only population transfer efficiency.

In the second part, we extend this method to a two ultra-strongly coupled qubit system. Here, the same ML approach successfully distinguishes correlations in Markovian noise as well, due to asymmetries in system response.

In the final part, we analyze the Gaussianity of environmental noise. By modeling qubit dynamics under Gaussian and non-Gaussian noise, we train ML models to estimate key noise parameters and quantify the contributions of both the noise components, even in hybrid environments.

Together, these results demonstrate that ML can serve as a powerful, scalable tool for noise diagnosis in quantum systems.

List of Publications

1. **Noise classification in three-level quantum networks by Machine Learning**
   Shreyasi Mukherjee, et al.,
   Mach. Learn.: Sci. Technol. 5 045049 (2024),
   DOI 10.1088/2632-2153/ad9193

2. **Detection of noise correlations in two qubit systems by Machine Learning**
   Dario Fasone, Shreyasi Mukherjee et al.,
   to be submitted to *Special Topic on "Quantum Dynamics in Theory, Numerics and in Experimental Research" in APL Quantum*.

The following manuscripts have not been included in the thesis.

1. **A machine learning based approach to the identification of spectral densities in quantum open systems**
   Jessica Barr, Shreyasi Mukherjee, Alessandro Ferraro, Mauro Paternostro, Giorgio Zicari
   arXiv:2507.13730v1 (2025)

2. **Machine Learning-aided Optimal Control of a noisy qubit**
   Riccardo Cantone, Shreyasi Mukherjee, Luigi Giannelli, Elisabetta Paladino, Giuseppe Falci
   arXiv:2507.14085 (2025)

Chapter 1

Introduction

1.1 Background and Motivation

We are living in the era of quantum technology, where scientists across the globe are striving to achieve quantum supremacy — a breakthrough poised to revolutionize communication, computation, and sensing. This progress is driven by quantum information theory, which provides the theoretical foundation for many quantum technologies, including quantum communication, computation and cryptography [1, 2]. At its core, the idea is to harness the principles of quantum mechanics to store, transmit, and process information in ways that are either impossible [3] or highly inefficient using classical systems [4].

Solid-state devices such as nitrogen-vacancy (NV) centers in diamond [5], electron spin in semiconductor quantum dots [6] and superconducting nanocircuits based on Josephson junctions [7] have emerged as highly promising platforms for the physical realization of 'qubits', the fundamental units of quantum information. These platforms are appealing because they allow precise control of quantum states, can be scaled to larger systems, and have the potential to work with existing microelectronic technologies. In line with these advances, in 2019, Google experimentally demonstrated quantum supremacy for the first time: their 53-qubit superconducting processor Sycamore completed a specific computational task in approximately 200 seconds, a task estimated to take around $10,000$

years on the most powerful classical supercomputers [8]. IBM recently claimed the development of Condor quantum processor with 1121-qubit, marking a significant milestone toward scalable quantum computing [9].

However, these technological advances come with significant challenges. As the number of qubits increases, maintaining their coherence and preserving delicate quantum correlations becomes progressively more difficult. This difficulty stems from the fundamental fact that no quantum system can be perfectly isolated from its external environment. Inevitably, interactions with the surroundings introduce *noise*, which leads to *decoherence* and error, thereby degrading the quantum properties essential for information processing. Understanding and mitigating the effects of environmental noise is thus a central focus in the quest for scalable, reliable quantum systems — an essential step toward realizing the promises of quantum supremacy. The study of *open quantum systems* is fundamentally dedicated to this purpose: developing theoretical frameworks and practical strategies to counteract decoherence and the inevitable loss of quantum information arising from system-environment interactions.

Characterizing noise becomes even more demanding in multi-qubit or multi-level quantum systems, where the noise can exhibit correlations, cross-talk, and nontrivial spectral features. Accurate knowledge of the noise environment is not only crucial for understanding decoherence but also serves as a prerequisite for designing effective quantum control protocols. Techniques such as dynamical decoupling and error correction rely on detailed information about the noise spectrum and its temporal correlations to suppress decoherence and improve the fidelity of quantum operations [10–12]. Thus, advances in noise characterization directly enable the development of robust noise mitigation strategies, which are essential for scaling up quantum technologies.

Traditional noise characterization techniques such as quantum process tomography [13–15], randomized benchmarking [16–18], and noise spectroscopy [19, 20] are foundational tools for diagnosing and mitigating errors in quantum systems. However, these methods face significant challenges in terms of scalability, resource requirements, and sensitivity to system size and complexity especially in multi-level or multi-qubit architectures. Machine learning (ML), with its powerful ability to learn complex patterns from data, has emerged as a promising alternative for

noise characterization in quantum systems [21, 22]. While predicting qubit dynamics is relatively straightforward when the environment is fully known, the inverse problem—inferring properties of the environment from observed qubit behavior is significantly more challenging. In many practical scenarios, the environment is only partially known or cannot be easily described using analytical models. By simulating the evolution of qubits under realistic noisy conditions, ML models, particularly neural networks can be trained to map experimentally accessible measurements of qubit dynamics to features of the underlying noise environment, without requiring complete prior knowledge of the total system.

1.2 Scope and Objectives

When studying the interaction between a quantum system and its environment, the standard approach is to model the environment quantum mechanically. However, as previously discussed, for complex or large environments, an exact quantum mechanical description can become highly challenging or even impractical. In such cases, adopting a classical description of the environment offers a valid and effective alternative that has been shown to yield accurate results. Typically, the environment is modeled as a classical stochastic field, and numerous studies have established the conditions under which environmental influences can be reliably approximated as classical noise [23–26].

In this PhD thesis, we primarily model environmental noise as a classical stochastic process and utilize the quantum system as a probe to sense and measure these fluctuations. Our objective is to combine quantum mechanical principles with data-driven machine learning (ML) approaches to propose effective techniques for noise sensing, classification, and characterization in quantum systems. The analysis focuses on two broad categories of classical noise: *Markovian* and *non-Markovian*, which are well-defined within the framework of Kolmogorov probability theory. We explore different subclasses of Markovian and non-Markovian noise based on their spatial correlations and investigate their impact on decoherence. This is particularly significant, as studies have shown that spatially correlated noise in multi-qubit systems can severely undermine

the effectiveness of quantum error correction protocols [27]. To address this, we apply ML-based methods to classify various noise types in relatively simple quantum systems, such as three-level systems and two-qubit systems.

Beyond simple classification, we aim to quantitatively characterize the noise in terms of key parameters such as its strength, correlation time, and rate. In this context, we model the environment using both *Gaussian* and *non-Gaussian* stochastic processes. While qubit evolution under Gaussian noise is relatively tractable, as it can be fully characterized by second order statistics, the presence of non-Gaussian noise complicates the analysis necessitating a more detailed characterization that incorporates non-trivial higher-order correlations. In fact, the noise that predominantly affects solid-state quantum devices is often non-Gaussian in nature, typically generated by bistable fluctuators (random telegraph noise) or by an ensemble of fluctuators with a broad distribution of switching rates ($1/f^{\alpha}$ noise) [23, 28]. Accurately characterizing such complex noise structures requires sophisticated approaches, which we address through the development and application of ML-assisted techniques. These methods are designed to enable noise diagnostics with minimal measurement overhead, making them particularly relevant for scalable quantum architectures.

1.3 Structure of the Thesis

- In chapter 2, we briefly introduce the fundamental concepts of probability and statistics that form the basis for the main work presented in this thesis. We also discuss the properties of classical stochastic processes, focusing on Markovian, Gaussian (Ornstein-Uhlenbeck), and non-Gaussian (Random Telegraph) processes. In addition, we provide a brief overview of open quantum system dynamics, which is essential for understanding noise-induced effects in quantum systems. Finally, we derive the expressions for qubit decoherence under the influence of the latter two noise models.

- In chapter 3, we introduce fundamental concepts of machine learning, such as activation functions, cost functions, backpropagation,

and gradient descent. Within the broad class of machine learning techniques, we primarily focus on supervised learning, providing a comparative study between classification and regression tasks. Among the various supervised learning approaches, we elaborate on how neural networks can be employed for supervised learning, which is particularly relevant to the objectives of this thesis.

- In chapter 4, we demonstrate that machine learning can classify noise in three single level quantum dots by detecting spatial and multilevel noise correlations and Markovianity. Using the CTAP (*coherent tunneling by adiabatic passage*)/STIRAP (*stimulated Raman adiabatic passage*) protocol, we analyze how population transfer efficiency varies under different noise conditions. The efficiency of population transfer under three driving pulse conditions serves as inputs to train a neural network, which classified five noise types: non-Markovian correlated, anti-correlated, uncorrelated, and Markovian correlated, anti-correlated. The model achieves over 99% accuracy in distinguishing between Markovian and non-Markovian noise, as well as among non-Markovian subclasses. However, it fails to differentiate between the Markovian subclasses, indicating a feature limitation. Our approach, which uses population transfer efficiency instead of complex time-series data, is computationally efficient and robust against statistical errors, making it suitable for experimental applications. These results advance noise characterization and quantum error correction strategies.

- In chapter 5, building on the previous work, we focus on detecting noise correlations and Markovianity in a two ultra-strongly coupled qubit system. Under the rotating wave approximation and in certain parameter regimes, we can reduce the unperturbed four-level system to a three-level one, allowing the use of the STIRAP protocol for efficient population transfer. By applying the technic developed in the previous chapter, we extend our approach to distinguish among the previously studied noise classes in this two-qubit setting. Furthermore, we demonstrate that in this case our approach is capable of distinguishing between Markovian noise subclasses as well. Simple neural network is used to classify the noise based on the

efficiency of population transfer under different conditions. This method remains computationally efficient and can be applied experimentally, advancing our ability to analyze noise correlations in multi-qubit systems.

- In chapter 6, we shift our research focus towards analyzing Gaussianity—another important characteristic of noise and estimating key noise parameters that influence quantum systems. In this work, we specifically study the interaction of a single qubit with both Gaussian and non-Gaussian noise sources. For the Gaussian case, we model the environment using the Ornstein-Uhlenbeck (OU) process, which captures colored noise with finite correlation time. To represent non-Gaussian noise, we employ the Random Telegraph Process (RTP), characterized by stochastic switching between discrete values, leading to non-Gaussian statistical features. The analysis includes, i) environments with numerous weakly coupled bistable fluctuators collectively mimicking Gaussian behavior, and ii) hybrid case with both strongly and weakly coupled fluctuators. We analyze qubit coherence decay under these conditions aiming to find noise parameters like, strength, coherence time and rate of the two types of noises and a quantifier for Gaussianity of noise. The machine learning model is trained on a dataset of simulated qubit evolutions, enabling accurate predictions of noise characteristics, even when parameters vary randomly.

- In chapter 7, we summarize the main findings of the thesis and outline potential directions for future research.

Chapter 2

Preliminaries

In this chapter, we introduce key concepts and fundamental definitions from probability theory and open quantum systems that form the foundation for the later chapters of this thesis. While this chapter provides essential background, it is not intended to be exhaustive; instead, it focuses on the aspects most relevant to the discussions and analyses that follow.

For probability theory, we have mostly reviewed the literature from [29, 30]. Some other recommended references are [31, 32].

2.1 Random variables

A *random* or *stochastic variable* is a rule for assigning every possible outcome ζ of an experiment a numerical value $\tilde{x}(\zeta)$. The domain of $\tilde{x}$ is a set $\mathcal{S}$, which consists of all possible outcomes of an experiment and the set of numerical values that it can assume is often referred to as *range* [30].

Random variables can be categorized as *discrete* or *continuous* based on whether the set of possible values it assumes $\{\tilde{x}(\zeta), \forall \zeta \in \mathcal{S}\}$ is countable or not [29].

For a discrete random variable, we associate each possible value x_i with a probability $p(x_i)$, where $i \in [1, 2, \dots]$. These probabilities satisfy

the normalization condition

$$\sum_i p(x_i) = 1. \tag{2.1}$$

For a continuous random variable, we define a probability density function $p(x)$ such that $p(x)dx$ represents the probability of the variable taking a value within the interval $[x, x + dx]$. The corresponding normalization condition is

$$\int_{-\infty}^{\infty} p(x)dx = 1. \tag{2.2}$$

An important quantity associated with a random variable is its *average* or *mean*, defined as follows:

- For a discrete random variable:

$$\langle \tilde{x} \rangle = \sum_i x_i p(x_i). \tag{2.3}$$

- For a continuous random variable:

$$\langle \tilde{x} \rangle = \int x p(x)dx. \tag{2.4}$$

More generally, if $f(\tilde{x})$ is any function of the random variable, then its expectation value is given by:

$$\langle f(\tilde{x}) \rangle = \sum_i f(x_i) p(x_i). \tag{2.5}$$

Correspondingly for continuous case

$$\langle f(\tilde{x}) \rangle = \int f(x) p(x)dx. \tag{2.6}$$

These definitions form the foundation for understanding and analyzing the behavior of random variables in probabilistic and statistical contexts. From this point onward, we focus primarily on continuous random variables when defining the relevant quantities, as the corresponding expressions for the discrete case can be obtained in a straightforward manner.

2.1.1 Moments, correlation, covariance

Equation 2.4 defines the first *moment* of $\tilde{x}$ relative to the origin at $x = 0$. Generally, the m-th order moment is denoted by $\langle x^m \rangle$ [29]. A related metric is the *central moment*, where the m-th central moment is formulated as

$$\mu^{(m)} = \langle (\tilde{x} - \langle \tilde{x} \rangle)^m \rangle = \int (x - \langle x \rangle)^m p(x) dx. \tag{2.7}$$

According to this definition, the first central moment ($m = 1$) is $\mu^{(1)} = 0$. For $m = 2$, it represents the *variance*, given by

$$\mathrm{Var}\{\tilde{x}\} = \mu^{(2)} = \left\langle (\tilde{x} - \langle \tilde{x} \rangle)^2 \right\rangle. \tag{2.8}$$

The square root of the variance $\sigma = \sqrt{\left\langle (\tilde{x} - \langle \tilde{x} \rangle)^2 \right\rangle}$ represents the *standard deviation*, which indicates how the values of $\tilde{x}$ diverge from the mean $\langle \tilde{x} \rangle$. Moments can also be defined jointly for different random variables $\tilde{x}_i, \tilde{x}_j, \tilde{x}_k, \ldots$. If $p(x_i, x_j, x_k, \ldots)$ is the joint probability density of $\tilde{x}_i, \tilde{x}_j, \tilde{x}_k, \ldots$, then the central moment of the order $(l, m, n, \ldots)$ is given by

$$\mu_{ijk\ldots}^{(lmn\ldots)} = \left\langle (\tilde{x}_i - \langle \tilde{x}_i \rangle)^l (\tilde{x}_j - \langle \tilde{x}_j \rangle)^m (\tilde{x}_k - \langle \tilde{x}_k \rangle)^n \ldots \right\rangle. \tag{2.9}$$

An important special case of central moments is the *covariance*. For two random variables $\tilde{x}_i$ and $\tilde{x}_j$, the covariance is defined as

$$Cov\{\tilde{x}_i, \tilde{x}_j\} = \mu_{ij}^{(11)} = \left\langle (\tilde{x}_i - \langle \tilde{x}_i \rangle)(\tilde{x}_j - \langle \tilde{x}_j \rangle) \right\rangle = \langle \tilde{x}_i \tilde{x}_j \rangle - \langle \tilde{x}_i \rangle \langle \tilde{x}_j \rangle. \tag{2.10}$$

The covariances $\mu_{ij}^{(11)}$ form the entries of a $N \times N$ matrix called *covariance matrix*, whose diagonal elements correspond to the variances of individual variables $\mu_{11}^{(11)}, \mu_{22}^{(11)}, \ldots, \mu_{NN}^{(11)}$. If $\tilde{x}_i$ and $\tilde{x}_j$ are statistically independent, then their covariance is zero meaning that the covariance matrix is diagonal. On the other hand, if $\tilde{x}_i = \tilde{x}_j$, then

$$\mathrm{Cov}(\tilde{x}_i, \tilde{x}_j) = \mathrm{Var}\{\tilde{x}_i\}. \tag{2.11}$$

2.1.2 Characteristic function, cumulants

For a random variable $\tilde{x}$, the characteristic function is defined as the Fourier transform of its probability density function [31].

$$\varphi_{\tilde{x}}(s) = \left\langle e^{isx} \right\rangle = \int p(x) e^{isx} dx. \tag{2.12}$$

where $p(x)$ is the probability density function of $\tilde{x}$. The logarithm of the characteristic function is called *cumulant generating function* [33].

$$\log[\varphi_{\tilde{x}}(s)] = \sum_{j=1}^{\infty} \kappa_j \frac{(\mathrm{i}s)^j}{j!} \tag{2.13}$$

where $\kappa_j = \kappa_j(\tilde{x})$, $j \in \mathbb{N}$ are the cumulants of the random variable $\tilde{x}$. These cumulants are obtained by taking derivatives of the cumulant generating function and evaluating at zero.

$$\kappa_j(\tilde{x}) = (-\mathrm{i})^j \frac{d^j}{ds^j} \log[\varphi_{\tilde{x}}(s)]|_{s=0} \tag{2.14}$$

The first few cumulants have specific interpretations:

- κ_1: Mean of the distribution.
- κ_2: Variance, indicating the dispersion.
- κ_3: Skewness, measuring asymmetry.
- κ_4: Kurtosis, describing the "tailedness" of the distribution.

2.1.3 Some probability distribution functions

Normal or Gaussian Distribution

The *Normal* or *Gaussian distribution* is among the most widely utilized probability distribution functions. Within this thesis, we will frequently use this distribution. A continuous random variable $\tilde{x}$, defined within an interval $[-\infty, \infty]$ is called *Gaussian random variable* if its probability density function has the form

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\langle x \rangle)^2}{2\sigma^2}}. \tag{2.15}$$

It has two main parameters the mean $\langle \tilde{x} \rangle$ and the variance σ^2, and is normalised to 1 such that

$$\int_{-\infty}^{\infty} p(x)dx = 1. \tag{2.16}$$

Poison distribution

A discrete random variable $\tilde{x}$ is called a *Poisson random variable* with parameter λ if it takes non-negative integer values $0, 1, 2, \ldots, \infty$ with probabilities given by

$$p(\tilde{x} = k) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad \forall k = 0, 1, 2, \ldots, \infty. \tag{2.17}$$

where $\lambda > 0$ is the average rate at which events occur. This probability distribution models the number of times an event occurs within a fixed interval of time or space under the following condition [30]:

- The occurrence of one event does not affect the probability of another event occurring.
- Events occur at a constant average rate λ throughout the interval.
- Two events cannot occur exactly at the same instant; each event happens independently.

A notable property of the Poisson distribution is that both its mean and variance are equal to λ.

2.2 Stochastic Process: Properties and Definitions

A *stochastic* or *random process* $\tilde{x}(t, \zeta)$ is a rule that assigns to each outcome ζ of an experiment a time dependent function $\tilde{x}(t)$, $t \in T$. Thus a random process is an esemble of time functions $\{\tilde{x}(t, \zeta_1), \tilde{x}(t, \zeta_2), \ldots\}$ [30]. The following properties hold

- For a fixed outcome $\zeta = \zeta_k$, $\tilde{x}(t, \zeta_k)$ is a time function only. Different outcomes give different time function.

- If we consider a fixed time point i.e. $t = t_k$, then $\tilde{x}(t_k, \zeta)$ is a random variable. A stochastic process can equivalently be defined as a collection of random variables $\{\tilde{x}(t_k)\}_k$ for each time instant t_k.
- If both time and outcome are fixed, then $\tilde{x}(t_k, \zeta_k)$ is a number.

For our purpose we will use the notation $\tilde{x}(t)$ to refer random process, omitting ζ dependence for the rest of the thesis. A random process $\tilde{x}(t)$ can be discrete or continuous time depending on whether the time points are countable i.e. $t \in \{t_1, t_2, \dots\}$ or not i.e. $t \in \mathbb{R}$. In both the cases the state space can be discrete or continuous.

At each instant of time t, the random variable $\tilde{x}(t)$ has a probability density function $p[x(t)]$ or $p(x,t)$ that satisfies the the usual normalization condition:

$$\int p(x,t)dx = 1 \quad \forall t. \tag{2.18}$$

The average or mean value of the random process $\tilde{x}(t)$ at time t can be calculated using the probability density function $p[x(t)]$ as:

$$\langle \tilde{x}(t) \rangle = \int x p(x,t) dx. \tag{2.19}$$

Next, we will define some important quantities essential for random processes.

2.2.1 Joint probability

Consider that $\tilde{x}(t)$ takes values $x_1, x_2, \dots$ at times $t_1, t_2, \dots$, respectively. We can define the n-fold joint probability density as

$$p(x_n, t_n; x_{n-1}, t_{n-1}; \dots; x_1, t_1), \tag{2.20}$$

which gives the probability density that the process takes the value x_1 at t_1, x_2 at t_2 and so on up to x_n at t_n and provides increasingly detailed information about the random process $\tilde{x}(t)$. The joint probability density function must satisfy the consistency condition

$$\int p(x_n, t_n; x_{n-1}, t_{n-1}; \dots; x_1, t_1) dx_{k+1} dx_{k+2} \dots dx_n = p(x_k, t_k; \dots; x_1, t_1), \tag{2.21}$$

for any integer $k < n$. Using $n = 2$, i.e., the 2-fold joint probability density, we can calculate an important quantity

$$\langle \tilde{x}(t_1)\tilde{x}(t_2)\rangle = \int x_1 x_2, p(x_2, t_2; x_1, t_1) dx_2 dx_1. \tag{2.22}$$

$\langle \tilde{x}(t_1)\tilde{x}(t_2)\rangle$ is called the *two-time auto-correlation function* of the random process $\tilde{x}(t)$. This quantity measures how a single random process $\tilde{x}(t)$ is correlated with itself at different time lags $\tau = t_2 - t_1$. Another measure of interest is *auto covariance function* which can be defined in the following manner, demonstrating its relationship with the auto correlation function.

$$C(t_1, t_2) = \langle \tilde{x}(t_1)\tilde{x}(t_2)\rangle - \langle \tilde{x}(t_1)\rangle \langle \tilde{x}(t_2)\rangle . \tag{2.23}$$

In the simplest case of a stochastic process, the values $x_1, x_2, \ldots$ taken by $\tilde{x}(t)$ at different times $t_1, t_2, \ldots$ are completely independent. In this case, the joint probability density can be expressed as

$$p(x_n, t_n; x_{n-1}, t_{n-1}; \ldots; x_1, t_1) = \prod_{i=1}^{n} p(x_i, t_i). \tag{2.24}$$

2.2.2 Conditional Probability

Given the knowledge of the joint probability density, we can now define the *conditional probability*. This represents the probability that $\tilde{x}(t)$ takes values x_{k+1} at time t_{k+1}, x_{k+2} at time t_{k+2}, ..., given that at earlier times it took values x_k at time t_k, x_{k-1} at time t_{k-1}, ..., and x_1 at time t_1. It is formulated as

$$p(x_n, t_n; \ldots; x_{k+1}, t_{k+1} \mid x_k, t_k; \ldots; x_1, t_1) = \frac{p(x_n, t_n; \ldots; x_1, t_1)}{p(x_k, t_k; \ldots; x_1, t_1)}. \tag{2.25}$$

This definition allows for a detailed statistical description of how future values of the stochastic process depend on past values. If the values taken by the process $\tilde{x}(t)$ at different times are completely independent of each other then the conditional probability equals to the *unconditional* or joint probability density, i.e.,

$$p(x_n, t_n; \ldots; x_{k+1}, t_{k+1} \mid x_k, t_k; \ldots; x_1, t_1) = p(x_n, t_n; \ldots; x_{k+1}, t_{k+1}). \tag{2.26}$$

Substituting the above equation in equation 2.25 and putting $k = 1$, we have

$$p(x_n, t_n; \ldots; x_1, t_1) = p(x_n, t_n; \ldots; x_2, t_2)p(x_1, t_1). \tag{2.27}$$

Notice that repeated application of this relation eventually leads to 2.24.

2.2.3 Stationarity

A stationary process refers to a stochastic process where the unconditional joint probability distribution remains invariant under time shifts. Mathematically, stationarity implies that

$$p(x_n, t_n; x_{n-1}, t_{n-1}; \ldots; x_1, t_1) = p(x_n, t_n + T; x_{n-1}, t_{n-1} + T; \ldots; x_1, t_1 + T), \tag{2.28}$$

$\forall$ time shift T. As a consequence, statistical properties such as the mean, variance, and higher-order moments do not depend on the absolute time, but only on time differences.

2.3 Markov Process

We define *Markov process* as the one, where the future of $\tilde{x}(t)$ is influenced by only its present and not by its past history. From the definition of conditional probability used in 2.2.2 we can mathematically define Markov process or more precisely *first order Markov process* in the following way,

$$p(x_n, t_n; \ldots; x_{k+1}, t_{k+1} | x_k, t_k; \ldots; x_1, t_1) = p(x_n, t_n; \ldots; x_{k+1}, t_{k+1} | x_k, t_k). \tag{2.29}$$

Following equation 2.25 we can break the joint probability density as

$$p(x_n, t_n; \ldots; x_1, t_1) = p(x_n, t_n | x_{n-1}, t_{n-1}; \ldots; x_1, t_1)p(x_{n-1}, t_{n-1}; \ldots; x_1, t_1). \tag{2.30}$$

On the other hand using the Markov property, we can write $p(x_n, t_n | x_{n-1}, t_{n-1}; \ldots; x_1, t_1) = p(x_n, t_n | x_{n-1}, t_{n-1})$. Thus we obtain,

$$p(x_n, t_n; \ldots; x_1, t_1) = p(x_n, t_n | x_{n-1}, t_{n-1})p(x_{n-1}, t_{n-1}; \ldots; x_1, t_1). \tag{2.31}$$

Thus repeated application of Markov property in the above eq. results in

$$p(x_n, t_n; \ldots; x_1, t_1) = p(x_n, t_n | x_{n-1}, t_{n-1})p(x_{n-1}, t_{n-1} | x_{n-2}, t_{n-2}) \ldots p(x_1, t_1). \tag{2.32}$$

For Markov process, the conditional probability density $p(x_n, t_n|x_{n-1}, t_{n-1})$ is called *transition probability density*. It can be used to derive the following integral relation for Markov process. For any $t_3 \geq t_2 \geq t_1$, it can be written

$$\begin{aligned} p(x_3, t_3; x_1, t_1) &= \int p(x_3, t_3; x_2, t_2; x_1, t_1) dx_2 \\ \Rightarrow p(x_3, t_3; x_1, t_1) &= \int p(x_3, t_3|x_2, t_2) p(x_2, t_2|x_1, t_1) p(x_1, t_1) dx_2 \\ \Rightarrow p(x_3, t_3|x_1, t_1) &= \int p(x_3, t_3|x_2, t_2) p(x_2, t_2|x_1, t_1) dx_2 \quad (t_3 \geq t_2 \geq t_1). \end{aligned} \tag{2.33}$$

Where we have used 2.32 in the right hand side in the second line of the above equation and in the third line we have divided both sides with $p(x_1, t_1)$. This relation is known as *Chapman-Kolmogorov equation*.

For discrete state space, the Chapman-Kolmogorov equation takes the following form

$$p(x_3, t_3|x_1, t_1) = \sum_{x_2} p(x_3, t_3|x_2, t_2) p(x_2, t_2|x_1, t_1). \tag{2.34}$$

2.4 Gaussian Process

Another category of stochastic processes is *Gaussian* or *non-Gaussian process*. A stochastic process is called Gaussian if it can be characterised fully with 2nd order statistics i.e. the mean and two time auto covariance function [34] (the definition of auto covariance function is illustrated in eq. (2.23)). In other words, all the information of the process is contained in its power spectrum. More formally, a stochastic process $\{\tilde{x}(t) : t \in T\}$ is said to be Gaussian if for any finite set of time instances $t_1, t_2, \ldots, t_n \in T$, the corresponding random variables $(\tilde{x}(t_1), \tilde{x}(t_2), \ldots, \tilde{x}(t_n))$ follow a joint normal distribution. The characteristic function of a single random variable $\tilde{x}$ having Gaussian distribution $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ with mean μ and standard deviation σ is (see the definition eq. (2.12))

$$\varphi_{\tilde{x}}(u) = \left\langle e^{\mathrm{i}u\tilde{x}} \right\rangle = \int_{-\infty}^{\infty} p(x) e^{\mathrm{i}ux} dx = e^{\mathrm{i}\mu u - \frac{1}{2}u^2\sigma^2}. \tag{2.35}$$

For a Gaussian process, the joint characteristic function of any finite collection of n random variables $\{\tilde{x}(t_1), \tilde{x}(t_2), \ldots, \tilde{x}(t_n)\} \subset \{\tilde{x}(t) : t \in T\}$ is thus [35]

$$\begin{aligned} \varphi_{\tilde{x}(t_1),\ldots,\tilde{x}(t_n)}(u_1, \ldots, u_n) &= \left\langle \exp\left(\mathrm{i} \sum_{j=1}^{n} u_j \tilde{x}(t_j) \right) \right\rangle \\ &= \exp\left(\mathrm{i} \sum_{j=1}^{n} u_j \mu(t_j) - \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} u_j u_k C(t_j, t_k) \right). \end{aligned} \tag{2.36}$$

where $C(t_j, t_k)$ is the auto covariance function as defined in (2.23). In case of continuous time Gaussian process, considering t varying continuously within the range $[t_0, t] \in T$ the above equation takes the form

$$\begin{aligned} &\left\langle \exp\left(\mathrm{i} \int_{t_0}^{t} u_j(s) \tilde{x}(s) ds \right) \right\rangle \\ &= \exp\left(i \int_{t_0}^{t} u(s) \mu(s) ds - \frac{1}{2} \int_{t_0}^{t} \int_{t_0}^{t} u(s) u(s') C(s, s') ds ds' \right). \end{aligned} \tag{2.37}$$

As mentioned before, it is evident that the Gaussian process is fully characterised by its mean and auto covariance or auto correlation function (for 0 mean Gaussian process). Gaussian processes play a fundamental role in quantum physics, particularly in modeling environmental noise affecting qubits. In open quantum systems, Gaussian noise is commonly assumed due to its analytical tractability and its prevalence in thermal and electromagnetic fluctuations.

2.4.1 Example of Gaussian process

Ornstein Uhlenbeck (OU) process

OU process is a widely used example for Gaussian process with finite correlation time. In quantum physics it is particularly relevant for describing realistic noise sources. It is a stationary, Gaussian, and Markovian process that obeys the following stochastic differential equation [31]

$$d\tilde{x}(t) = k(-x(t) + \mu)dt + \sqrt{D} dW(t), \tag{2.38}$$

where $k = \frac{1}{\tau_c}$ is the inverse of correlation time τ_c, μ is stationary mean, and $D = 2k\sigma^2$ is known as the diffusion coefficient (σ is the standard deviation of the probability distribution of the OU process $x(t)$), and $W(t)$ represents the Wiener process. By solving the above eq. (2.38) using an integrating factor $e^{\int k dt}$, we can derive the solution,

$$\tilde{x}(t) = \tilde{x}(0)e^{-kt} + \mu(1 - e^{-kt}) + \sqrt{D}\int_0^t e^{-K(t-t')}dW(t'). \tag{2.39}$$

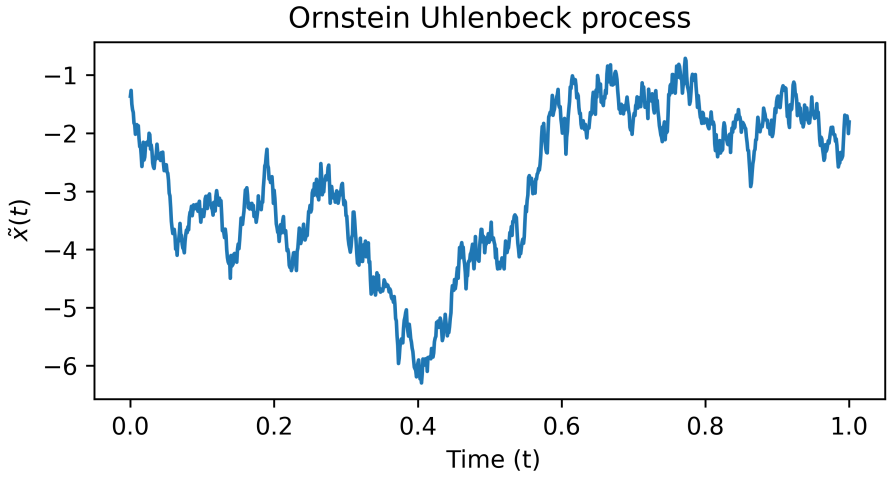

Figure 2.1: Example of a single trajectory of OU process with parameters $k = 1$, $\mu = 0$, $\sigma^2 = 10$.

Now it is easy to obtain the average, variance and correlation function of the OU process.

1. Mean:
$$\langle \tilde{x}(t) \rangle = \langle \tilde{x}(0) \rangle e^{-kt} + \mu(1 - e^{-kt}). \tag{2.40}$$

2. Variance :
$$\mathrm{Var}[\tilde{x}(t)] = \left\langle (\tilde{x}(t) - \langle \tilde{x}(t) \rangle)^2 \right\rangle \tag{2.41}$$
We substitute the expression of $x(t)$ and $\langle x(t) \rangle$ in the above equation. Using the approximation $dW(t') \approx dt'$ and considering $\langle dW(t') \rangle =$

0, the obtained expression of variance is

$$\mathrm{Var}[\tilde{x}(t)] = \left(\mathrm{Var}[\tilde{x}(0)] - \frac{D}{2k}\right) e^{-2kt} + \frac{D}{2k} \tag{2.42}$$

At stationary state $t \to \infty$, $\langle \tilde{x}(t) \rangle = \mu$ and $\mathrm{Var}[\tilde{x}(t)] = \frac{D}{2k}$.

3. Covariance : From the definition of the covariance for a random process as given in equation 2.23, it can be shown for an Ornstein-Uhlenbeck (OU) process that

$$\begin{aligned} C(t,t') &= \left\langle \left(\tilde{x}(t) - \langle \tilde{x}(t) \rangle\right) \left(\tilde{x}(t') - \left\langle \tilde{x}(t') \right\rangle\right) \right\rangle \\ &= \left[\mathrm{Var}[\tilde{x}(0)] - \frac{D}{2k}\right] e^{-k(t+t')} + \frac{D}{2k} e^{-k|t-t'|} \end{aligned} \tag{2.43}$$

When the mean is zero, the covariance is equivalent to the autocorrelation function, given by $C(t_1, t_2) = \langle \tilde{x}(t)\tilde{x}(t') \rangle = \frac{D}{2k} e^{-k|t-t'|}$ (considering stationary state i.e. $t \to \infty$). The power spectrum of the OU process can be calculated by the Fourier transform of the autocorrelation function

$$S_{OU}(\omega) = \frac{D}{(k^2 + \omega^2)}, \tag{2.44}$$

which is a Lorentzian function.

2.5 Non-Gaussian process

Unlike Gaussian processes, non-Gaussian processes cannot be fully characterized by their mean and covariance alone. Instead, all the higher-order moments are required for an exact description. In quantum physics, specially in superconducting qubits and quantum dots it is common to characterize the environment with non-Gaussian processes. A non-Gaussian environment is sometimes approximated as Gaussian for analytical or computational simplifications, such an approximation often breaks down in scenarios where a quantum system is strongly coupled to discrete frequency modes or an environment comprising bistable fluctuators that shows $\frac{1}{f}$ type behaviour [23]. These cases introduce memory effects and

non-trivial correlations that cannot be captured by Gaussian assumptions [28, 36]. Moreover, the power spectrum alone does not provide a complete characterization of environmental noise. Different stochastic processes can share the same power spectrum while inducing different effects on system dynamics due to variations in their higher-order statistical properties, such as non-Gaussian correlations and memory effects. Consequently, a detailed description of the underlying stochastic process is crucial for accurately modeling noise-induced decoherence and relaxation. A more comprehensive understanding of these processes enables precise control and mitigation strategies, which are essential for improving the coherence and performance of quantum devices. A widely studied model for generating non-Gaussian noise involves bistable fluctuators that switch randomly between two discrete levels. The behavior of an individual bistable fluctuator, also known as a two-level system (TLS), can be effectively described by a random telegraph process (RTP). Under certain coupling regimes, the collective effect of multiple bistable fluctuators gives rise to $1/f$ noise, which is another well-known example of non-Gaussian noise. In the following section, we will further elaborate on the RTP.

2.5.1 Example of non Gaussian process

Random telegraph process (RTP)

A RTP is a Markov process in which the random variable $\tilde{x}(t)$ at each instant of time t, toggles between two constant values, a and b, with transition rates Γ^+ (a to b) and Γ^- (b to a) [31]. From the Chapman-Kolmogorov equation 2.34, we express this as follows

$$\begin{aligned} p(a, t+\Delta t|x, t_0) &= \sum_m p(a, t+\Delta t|m, t)p(m, t|x, t_0) \\ &= p(a, t+\Delta t|a, t)p(a, t|x, t_0) + p(a, t+\Delta t|b, t)p(b, t|x, t_0) \\ &= (1-\Gamma^+\Delta t)p(a, t|x, t_0) + \Gamma^-\Delta t p(b, t|x, t_0). \end{aligned} \tag{2.45}$$

The following relations have been applied in the above expression

$$p(b, t + \Delta t|a, t) = \Gamma^+ \Delta t, \tag{2.46}$$
$$p(a, t + \Delta t|b, t) = \Gamma^- \Delta t. \tag{2.47}$$

Thus, we can formulate

$$\begin{aligned} &\lim_{\Delta t \to 0} \frac{p(a, t + \Delta t|x, t_0) - p(a, t|x, t_0)}{\Delta t} = -\Gamma^+ p(a, t|x, t_0) + \Gamma^- p(b, t|x, t_0) \\ &\Rightarrow \dot{p}(a, t|x, t_0) = -\Gamma^+ p(a, t|x, t_0) + \Gamma^- p(b, t|x, t_0). \end{aligned} \tag{2.48}$$

Similarly, it is possible to show that

$$\dot{p}(b, t|x, t_0) = \Gamma^+ p(a, t|x, t_0) - \Gamma^- p(b, t|x, t_0). \tag{2.49}$$

Eq. 2.48 and 2.49 are the master equations for RTP. The two coupled differential equations can be solved using the initial conditions $p(a, t|x, t_0) - p(b, t|x, t_0) = \delta p(t_0)$ and $p(a, t|x, t_0) + p(b, t|x, t_0) = 1$ which give the following solutions

$$p(a, t|x, t_0) = \frac{\Gamma^-}{\Gamma^+ + \Gamma^-} + \left[\delta p(t_0) - \frac{\Gamma^- - \Gamma^+}{\Gamma^+ + \Gamma^-}\right] \frac{e^{-(\Gamma^+ + \Gamma^-)(t - t_0)}}{2}, \tag{2.50}$$

$$p(b, t|x, t_0) = \frac{\Gamma^+}{\Gamma^+ + \Gamma^-} - \left[\delta p(t_0) - \frac{\Gamma^- - \Gamma^+}{\Gamma^+ + \Gamma^-}\right] \frac{e^{-(\Gamma^+ + \Gamma^-)(t - t_0)}}{2}. \tag{2.51}$$

It is now straightforward to calculate the following properties of the random variable $\tilde{x}(t)$.

1. Mean:

$$\begin{aligned} \langle \tilde{x}(t)|x_0, t_0\rangle &= \sum_{x=\{a,b\}} x p(x, t|x_0, t_0) \\ &= \frac{a\Gamma^- + b\Gamma^+}{\Gamma^+ + \Gamma^-} + (a - b)\left[\delta p(t_0) - \frac{\Gamma^- - \Gamma^+}{\Gamma^+ + \Gamma^-}\right] \frac{e^{-(\Gamma^+ + \Gamma^-)(t - t_0)}}{2}. \end{aligned} \tag{2.52}$$

   At stationary sate i.e. for $t_0 \to -\infty$,

$$\langle \tilde{x}(t)\rangle = \frac{a\Gamma^- + b\Gamma^+}{\Gamma^+ + \Gamma^-}. \tag{2.53}$$

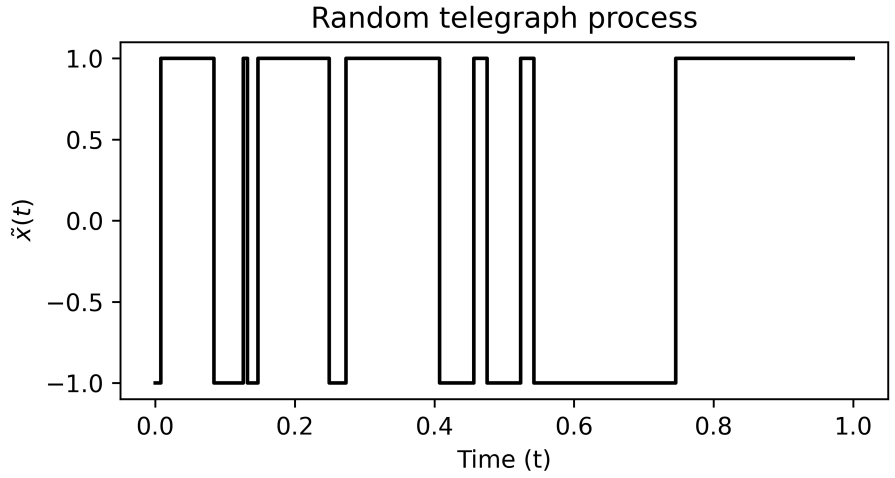

Figure 2.2: Example of a single trajectory of RTP with parameters $\Gamma^+ = \Gamma^- = 10$.

2. Variance:

$$\begin{aligned}\mathrm{Var}[\tilde{x}(t)] &= \left\langle (\tilde{x}(t) - \langle \tilde{x}(t) \rangle)^2 \right\rangle \\ &= \sum_{x=\{a,b\}} x^2 p(x,t|x_0,t_0) - \left(\sum_{x=\{a,b\}} x p(x,t|x_0,t_0) \right)^2 \\ &\approx \frac{(a-b)^2 \Gamma^+ \Gamma^-}{(\Gamma^+ + \Gamma^-)^2} \quad \text{(at stationary state).}\end{aligned} \tag{2.54}$$

3. Auto covariance:

$$\begin{aligned}C(t,t') &= \left\langle \left(\tilde{x}(t) - \langle \tilde{x}(t) \rangle\right) \left(\tilde{x}(t') - \left\langle \tilde{x}(t') \right\rangle\right) \right\rangle \\ &= \sum_{x,x'} x x' p(x,t;x',t') - \sum_{x=\{a,b\}} x p(x,t|x_0,t_0) \sum_{x'=\{a,b\}} x' p(x',t'|x_0,t_0) \\ &= \sum_{x,x'} x x' p(x,t|x',t') p(x',t') - \sum_{x=\{a,b\}} x p(x,t|x_0,t_0) \sum_{x'=\{a,b\}} x' p(x',t'|x_0,t_0) \\ &\approx \frac{(a-b)^2 \Gamma^+ \Gamma^-}{(\Gamma^+ + \Gamma^-)^2} e^{-(\Gamma^+ + \Gamma^-)|t-t'|} \quad \text{(at stationary state).}\end{aligned} \tag{2.55}$$

2.6 Decoherence

In practice, no quantum system is truly closed or completely isolated from its external environment; there is always some degree of unavoidable coupling. This interaction leads to correlations between the system and its environment, causing the environment to retain information about the system's state [37]. When the environmental degrees of freedom are traced out, certain states within the system's Hilbert space remain stable, while superpositions of these states gradually degrade over time. This loss of quantum coherence due to environmental influence is known as decoherence.

Decoherence poses a fundamental challenge to the principle of superposition, a cornerstone of quantum mechanics. Consequently, it presents significant obstacles to quantum information processing, quantum communication, and quantum sensing, where the superposition principle serves as a crucial foundation [25].

For an open quantum system, the system (S) and the environment (E) are typically modeled as components of a bipartite system ($S + E$), where the composite Hilbert space is given by the tensor product $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$. The system and environment are described by their respective density matrices, $\rho_S \in \mathcal{H}_S$ and $\rho_E \in \mathcal{H}_E$. In most cases, we are not interested in the detailed dynamics of the environment but rather in its influence on the system. This influence is captured by computing the reduced density matrix of the system, obtained by tracing out the environmental degrees of freedom. The corresponding equation of motion for the system is given by

$$\dot{\rho}_S(t) = -\mathrm{i}\,\mathrm{Tr}_E\left[H_{SE}(t), \rho_{SE}(t)\right], \tag{2.56}$$

where $\rho_{SE}(t)$ is the density matrix of the total system and $H_{SE}(t)$ is the total Hamiltonian. Although the total system $S + E$ evolves unitarily according to the Von Neumann equation (2.56) (also known as *exact master equation*), it is of not much use. This is because it requires solving a very large total system and the state $\rho_{ES}(t)$ is often unknown due to the lack of knowledge of the environment. To overcome this problem, we seek an effective description of the system dynamics in terms of a quantum

dynamical map that governs the evolution of the system alone

$$\rho_S(t) = \mathrm{Tr}_E[\rho_{SE}(t)] = \mathrm{Tr}_E\left[U(t)\rho_{SE}(0)U^\dagger(t)\right] = \mathcal{V}(t)\rho_S(0), \tag{2.57}$$

where $U(t) = \mathcal{T}e^{-i\int_0^t H(s)ds}$ is the unitary time evolution operator for the full system, with $\mathcal{T}$ denoting time ordering and $\mathcal{V}(t)$ is an operator that maps $\rho_S(0) \to \rho_S(t)$ and is called *quantum dynamical map* (or *quantum map* or *dynamical map*). The map must follow certain conditions to represent a valid dynamics, such as it must be *convex-linear*, *completely positive* (CP) and *trace preserving* (TP) quantum operation [37, 38]:

- *Convex-linear*: $\mathcal{V}(t)\left[\sum_i p_i\rho_{S,i}(0)\right] = \sum_i p_i\mathcal{V}(t)\rho_{S,i}(0)$ where $\rho_{S,i}(0) \in H_S$ and $\sum_i p_i = 1$.
- *Completely positive*: This implies that the map not only maps a positive semi-definite operator to another positive semi-definite operator (positivity), $\mathcal{V}(t)\rho_S(0) \geq 0$; but also maintains its positivity when extended to a larger Hilbert space i.e. $\mathcal{V}(t) \otimes \mathbb{I}_{ext} > 0$.
- *Trace preserving*: $\mathrm{Tr}_S[\mathcal{V}(t)\rho_S(0)] = \mathrm{Tr}_S[\rho_S(0)] = 1$, ensure the state is a legitimate, normalized density matrix.

A quantum map with all the above properties can be obtained starting from equation (2.57). A way of constructing the map is by *Kraus operator representation* (also called *Operator-sum representation*). According to Kraus theorem, any quantum map satisfying the above conditions can be represented as

$$\begin{aligned}
\rho_S(t) &= \mathcal{V}_S(t)\rho_S(0) \\
&= \mathrm{Tr}_E\left[U(t)\rho_{SE}(0)U^\dagger(t)\right] \\
&= \sum_{\alpha\beta}\langle\phi_\beta|U(t)|\phi_\alpha\rangle\sqrt{\lambda_\alpha}\rho_S(0)\sqrt{\lambda_\alpha}\langle\phi_\alpha|U^\dagger(t)|\phi_\beta\rangle \\
&= \sum_{\alpha\beta}K_{\alpha\beta}\rho_S(0)K_{\alpha\beta}^\dagger
\end{aligned} \tag{2.58}$$

where we have considered that initially the system and the environment is in a factorised state $\rho_S(0) \otimes \rho_E(0)$ and the spectral decomposition $\rho_E(0) =$

$\lambda_\alpha|\phi_\alpha\rangle\langle\phi_\alpha|$ holds. Here we defined the operator $K_{\alpha\beta} = \sqrt{\lambda_\alpha}\langle\phi_\beta|U(t)|\phi_\alpha\rangle$ which evidently satisfy the completeness relation $K^\dagger_{\alpha\beta}K_{\alpha\beta} = \mathbb{I}$. The Kraus representation is not unique [25].

In this approach we need to first compute the time evolution of the total system $\rho_{SE}(t)$ before we can trace over the environment which is often a difficult task. To overcome this difficulty some approximations or assumptions are often taken on the map, which allows us to compute approximated evolution of the system even when the global evolution is not possible to calculate. For example, to derive a more tractable form of the master equation one often take the following approximations

1. Born Approximation: Assumes weak coupling between the system and the environment, allowing the total density matrix to be approximated as a product state at all times:

$$\rho_{SE}(t) \approx \rho_S(t) \otimes \rho_E. \tag{2.59}$$

   This approximation neglects correlations between the system and environment during their interaction.

2. Markov Approximation: Assumes that environmental correlations decay rapidly compared to the system's dynamics, leading to memoryless behavior. This allows us to approximate the system's future evolution as depending only on its current state, not its history.

With these assumptions the map can be expressed as

$$\mathcal{V}(t) = \exp(\mathcal{L}t), \tag{2.60}$$

which immediately allows us to write a first-order differential equation of the following form, referred to as *quantum Markovian master equation* or *Lindblad maser equation*

$$\dot{\rho}_S(t) = \mathcal{L}\rho_S(t) = -\mathrm{i}[\mathcal{H}(t), \rho_S(t)] + \mathcal{D}[\rho_S(t)], \tag{2.61}$$

where $\mathcal{L}$ is called Liouville super operator. The first term of the above equation describes the unitary evolution while the second term describes the decoherence and dissipation. The most general form of the dissipator

$\mathcal{D}$ for a Markovian master equation that preserves complete positivity and trace is

$$\mathcal{D}[\rho_S(t)] = \sum_k \gamma_k \left(L_k \rho_S(t) L_k^\dagger - \frac{1}{2}\{L_k^\dagger L_k, \rho_S(t)\} \right), \tag{2.62}$$

where L_k are the Lindblad or jump operators, γ_k are the non-negative quantities having the dimension of inverse of time and $\{.,.\}$ is the anti-commutator. In our work we will treat the environment classically. In some cases we can consider the Markovian nature of the environment while in other cases when this is not true, we will use the general approach 2.57. But for classical environment the partial trace over the environment is replaced by averaging over the stochastic process describing the environment.

2.6.1 Interaction with classical noise

In many physically relevant scenarios, the environment influencing the quantum system can be effectively modeled as a classical stochastic field [20, 22, 23, 26]. In such cases, rather than considering a full quantum treatment of the environment, we describe its effect through a random process that modulates the system's Hamiltonian. This approach is particularly useful when the environment is too complex to model microscopically, or when it exhibits classical-like behavior due to decoherence or macroscopic degrees of freedom.

In this section, we focus on a paradigmatic example of a quantum system (a qubit) interacting with a classical environment. The environmental effect can be represented by a stochastic process $\tilde{x}(t)$ 2.2. The total Hamiltonian can be written as

$$H(t) = H_S + H_I[\tilde{x}(t)] = -\frac{\omega}{2}\sigma_z - \frac{v}{2}\tilde{x}(t)\boldsymbol{\sigma}. \tag{2.63}$$

where H_S is the qubit Hamiltonian and the coupling of the environment is encoded in the interaction Hamiltonian H_I, which is the functional of the stochastic process, ω is the energy difference between the two levels of the qubit and v is the interaction strength of the noise with the qubit. σ_z is the Pauli z matrix, the choice of the z axis is arbitrary. For a given

realization of the stochastic process, the time evolution of the total system within $[0, t]$ can be described with the unitary operator as a solution to the Schrödinger equation.

$$\mathcal{U}(t) = \mathcal{T} e^{-\mathrm{i} \int_0^t H(s) ds}, \tag{2.64}$$

where $\mathcal{T}$ is the time ordering operator. To obtain the dynamical map for classical environments, in eq. (2.57) the trace over the environment has to be replaced by average over all possible realizations of the stochastic process. The resulting map is thus

$$\mathcal{V}(t)\rho(0) = \left\langle \mathcal{U}(t)\rho(0)\mathcal{U}^{\dagger}(t) \right\rangle, \tag{2.65}$$

which is a CPTP map. Generally, obtaining an analytical solution for the map poses challenges since the Hamiltonian specified in eq. (2.63) does not commute with itself at different time instances due to the algebra of the Pauli matrices and time dependence of $H_I(t)$. The analytical solution is possible to obtain only in some special cases. For simplicity we will consider the case of *pure dephasing* which is most relevant to our work.

Pure dephasing

Pure dephasing refers to a type of decoherence where a quantum system loses phase coherence without any energy exchange with its environment. This means that while the off-diagonal elements of the system's density matrix decay over time, the populations (diagonal elements) remain unchanged. Consider the following Hamiltonian

$$H(t) = H_S + H_I[\tilde{x}(t)] = -\frac{\omega}{2}\sigma_z - \frac{v}{2}\tilde{x}(t)\sigma_z, \tag{2.66}$$

here the noise $\tilde{x}(t)$ acts parallelly to the qubit altering its energy ω and thus referred as *longitudinal noise*. In this case, $[\sigma_z, H_S] = 0$ ensuring pure dephasing dynamics. Here the Hamiltonian commutes with itself at each instance of time, thus the unitary operator describing its dynamics can be written as

$$\mathcal{U}(t) = e^{-\mathrm{i} \int_0^t H(s) ds} = e^{\frac{\mathrm{i}}{2}(\omega t + v\varphi(t))\sigma_z}, \tag{2.67}$$

where $\varphi(t) = \int_0^t \tilde{x}(s)ds$ is the noise phase. Consequently we can write the qubit evolution following eq. (2.65)

$$\rho_S(t) = \begin{bmatrix} \rho_{00}(0) & \rho_{01}(0)e^{\mathrm{i}\omega t}\left\langle e^{\mathrm{i}v\varphi(t)}\right\rangle \\ \rho_{10}(0)e^{-\mathrm{i}\omega t}\left\langle e^{-\mathrm{i}v\varphi(t)}\right\rangle & \rho_{11}(0) \end{bmatrix}. \tag{2.68}$$

If we move to the interaction picture with respect to the qubit Hamiltonian H_S, i.e. performing $e^{\mathrm{i}\int_0^t H_S dt'}\rho_s(t)e^{-\mathrm{i}\int_0^t H_S dt'}$ we get rid of the exponential containing the qubit energy, resulting in

$$\rho_S^I(t) = \begin{bmatrix} \rho_{00}(0) & \rho_{01}(0)\left\langle e^{\mathrm{i}v\varphi(t)}\right\rangle \\ \rho_{10}(0)\left\langle e^{-\mathrm{i}v\varphi(t)}\right\rangle & \rho_{11}(0) \end{bmatrix}, \tag{2.69}$$

where $\rho_S^I(t)$ is the density matrix in the interaction picture.

For Gaussian noise, the averaging over the stochastic process appearing in the coherence terms in eq. (2.69) can be calculated easily using properties of Gaussian process 2.4. Since the noise phase $\varphi(t)$ is itself a Gaussian random variable with distribution

$$p(\varphi(t)) = \frac{1}{\sqrt{2\pi\left\langle\varphi^2(t)\right\rangle}}e^{-\frac{(\varphi(t)-\mu)^2}{2\left\langle\varphi^2(t)\right\rangle}}, \tag{2.70}$$

we obtain from eq. (2.12), with mean $\mu = 0$ and $u = -v$

$$\left\langle e^{\mathrm{i}v\varphi(t)}\right\rangle = \int_{-\infty}^{\infty} e^{\mathrm{i}v\varphi(t)}p(\varphi(t))d\varphi(t) = e^{-\frac{1}{2}v^2\left\langle\varphi^2(t)\right\rangle}, \tag{2.71}$$

where $\left\langle\varphi^2(t)\right\rangle = \int_{t_0}^t\int_{t_0}^t \left\langle x(s)x(s')\right\rangle dsds'$.

For non Gaussian noise, the distribution of the noise phase $\varphi(t)$ cannot simply be described by Gaussian distribution. Since we primarily focus on RTN, its probability distribution and characteristic function, as derived in [39, 40], are given as follows

$$\begin{aligned} p(\varphi,t) = {} & \frac{1}{2}e^{-\Gamma_0 t}\left[\delta(\varphi+vt)+\delta(\varphi-vt)\right] \\ & + \frac{\Gamma_0}{2v}e^{-\Gamma_0 t}\left[\Theta(\varphi+vt)-\Theta(\varphi-vt)\right] \\ & \times\left[\frac{\mathbf{I}_1\left(\Gamma_0 t\sqrt{1-(\varphi/vt)^2}\right)}{\sqrt{1-(\varphi/vt)^2}}+\mathbf{I}_0\left(\Gamma_0 t\sqrt{1-(\varphi/vt)^2}\right)\right], \end{aligned} \tag{2.72}$$

where Γ_0 is the switching rate of the RTP, $\delta(.)$ is delta function, $\mathbf{I}(.)$ is the modified Bessel function and $\Theta(.)$ is the Heaviside step function

$$\Theta(x) = \begin{cases} 1, & x > 0 \\ 0, & x < 0 \end{cases} \tag{2.73}$$

The characteristic function is therefore

$$\left\langle e^{\mathrm{i}v\varphi(t)} \right\rangle = \int p(\varphi, t) e^{\mathrm{i}v\varphi(t)} d\varphi = e^{-\Gamma_0 t} \left[\cosh\left(Z\Gamma_0 t\right) + \frac{1}{Z} \sinh\left(Z\Gamma_0 t\right) \right], \tag{2.74}$$

where $Z = \sqrt{1 - \frac{v^2}{(\Gamma_0)^2}}$. When $\Gamma_0 > v$, i.e. in weak coupling regime, the function is monotonically decreasing in time whereas, in the opposite case, i.e. in strong coupling regime $\Gamma_0 < v$ the characteristic function has oscillating behaviour. The above solution is valid when the switching rates are the same ($\Gamma^+ = \Gamma^- = \Gamma_0$) and the initial population difference (δp^0) between the two levels of the TLS is zero. A more general expression of the solution is given in [41]

$$\left\langle e^{\mathrm{i}v\varphi(t)} \right\rangle = A e^{-\Gamma(1-\alpha)t/2} + (1 - A) e^{-\Gamma(1+\alpha)t/2}. \tag{2.75}$$

where the variables of the equation are $\alpha = \sqrt{1 - g^2 - 2\mathrm{i}g\bar{\delta p}}$, $A = (1 + \alpha - \mathrm{i}\delta p^0 g)/(2\alpha)$, $g = 2v/\Gamma$, $\Gamma = \Gamma^+ + \Gamma^-$, $\bar{\delta p} = (\Gamma^- - \Gamma^+)/\Gamma$, and $\delta p^0 = \rho_{00}(0) - \rho_{11}(0)$. Here δp^0 represent the initial difference in populations and $\bar{\delta p}$ the stationary one.

Chapter 3

Machine learning

Machine learning (ML) is a discipline where computers extract patterns or knowledge from data without being directly instructed [42, 43]. It is a branch of artificial intelligence that only focuses on some fundamental aspects of *intelligence*, such as memory, adaptability, and generalization. It is a great tool specifically for problems that require a long list of conditions or rules, huge data set, fluctuating environment and where the traditional analytical input-output relationship is unknown. Depending on the problem and available data, ML can be categorized broadly in the following classes [44]

1. *Supervised learning:* Here, the goal is to train a model from labeled training data that can make accurate predictions on unseen inputs. "Supervised" indicates that each training example comes with a known output (label). The labeled data is used to train a model, which can then predict outcomes for new, unlabeled inputs. In this thesis we primarily focus on supervised learning, as it forms the core of the ML methods applied in our work. A more detailed description is provided later.

2. *Unsupervised learning:* Unlike supervised learning, the algorithm analyzes unlabeled data to discover hidden patterns or intrinsic structures without relying on predefined output labels. This approach allows the model to learn directly from the data distribution, which

eliminates the need for costly and time-consuming human annotation. Typical tasks in unsupervised learning include clustering data points into meaningful groups, identifying lower-dimensional manifolds where data lies, learning generative models to sample from complex distributions, and denoising corrupted data. Some widely used unsupervised learning algorithms are:

- *K-Means Clustering*
- *DBSCAN*
- *Principal Component Analysis (PCA)*
- *t-Distributed Stochastic Neighbor Embedding (t-SNE)*
- *Apriori*
- *Eclat*

Unsupervised learning is particularly useful for exploratory data analysis and feature learning when labeled data is scarce or unavailable.

3. *Reinforcement learning (RL):* It is a type of ML where an agent learns to make decisions by interacting with an environment. Instead of learning from labeled data, the agent receives feedback in the form of rewards or penalties based on its actions. The goal is to learn a policy that maximizes the cumulative reward over time. RL is widely used in robotics, game playing, and autonomous systems where trial-and-error and sequential decision making are essential.

3.1 Supervised learning

Our discussion of this section is primarily based on the reference [45]. In supervised learning, an algorithm is trained on a dataset $\{(\boldsymbol{x}_i, \hat{\boldsymbol{y}}_i)\}_{i=1,\ldots,N}$, known as the *training set*, where each instance $\boldsymbol{x}_i$ (*input*) is associated with a corresponding correct response $\hat{\boldsymbol{y}}_i$ (*label*). The goal of the algorithm is to learn a mapping from inputs to outputs such that it can generalize to unseen data of the same type. Each $\boldsymbol{x}_i$ in the set of N instances is a D-dimensional vector, referred to as a *feature vector*. The components

of this vector, denoted as x_{ij}, where $j = 1, \ldots, D$, represent distinct *features* that characterize the given instance. For example, in the estimation of an apartment's price, x_{i1} may represent the number of rooms, x_{i2} the location, and so on. These features serve as the input variables based on which the model makes predictions. Supervised learning problems can be broadly categorized into two types: *classification* and *regression*. A typical example of classification is email filtering, where an email is classified as either 'spam' or 'not spam'. In contrast, if the goal is to predict a continuous numerical value, such as the price of an apartment based on features like location, number of rooms, and area, the problem falls under regression. Thus, in a classification task, the labels $\hat{\boldsymbol{y}}_i$ characterize a set of categories or a finite number of classes $\{1, \ldots, C\}$, whereas in regression, they take real values. Several key supervised learning algorithms exist, including

- *k-Nearest Neighbors*
- *Linear Regression*
- *Logistic Regression*
- *Support Vector Machines (SVMs)*
- *Decision Trees and Random Forests*
- *Neural Networks (NNs)*.

In this discussion, our primary focus will be on NN-based algorithms.

3.2 Neural Networks (NNs)

Over the past decade, Neural Networks (NNs) have gained immense importance in science and technology [46–48]. In physics, they are extensively used for Hamiltonian construction and wave function prediction in quantum systems, controlling quantum systems, predicting material properties, classifying phases of matter, interpreting astronomical images, and much more. Neural Networks (NNs) are best understood as flexible function approximators [49], capable of learning complex mappings from

inputs to outputs based on training data. In supervised learning, the goal is to approximate an unknown function f that relates an input x to an output $y = f(x)$. Thus, we can approximate the unknown function f with a neural network as

$$y = f(x) \approx NN(x, \theta), \tag{3.1}$$

where θ are parameters. In many problems, we are given a dataset where each data sample consists of an input-output pair (x_i, y_i). The goal is that the NN approximates f as accurately as possible by optimizing the parameters $\theta = (\theta_1, \theta_2, \dots)$, thus "learning" the underlying input-output relationship. In principle, the approximation of $f(x)$ can be constructed in an arbitrary manner and it does not need to be a NN, but the desired representation should satisfy the following properties: (i) *Scalability*: The structure should be easily extendable to accommodate more parameters and handle higher-dimensional inputs and outputs. (ii) *Efficiency*: Not only should the approximation be computationally feasible to evaluate, but also its derivatives with respect to the model parameters should be efficiently computable, as required for training. NNs satisfy both of these conditions due to their characteristic structure. From now on we will call NNs and *Multi-layer Perceptrons* MLP in a synonymous way.

An MLP consists of three primary parts: the *input layer*, the *hidden layer(s)*, and the *output layer*. Each layer is composed of neurons, which are the fundamental units of a neural network. In a fully connected neural network, each neuron in a layer receives an input from each neuron of the preceding layer and calculates a real-valued output which is passed on to the next layer.

To illustrate the concept, consider a MLP consisting of $L + 1$ layers (as depicted in Fig. 3.1), with the input layer identified by the index $l = 0$ and the output layer by the index $l = L$. The value of the $D^{(l)}$ neurons in layer l are calculated from the values of the $D^{(l-1)}$ neurons in layer $l - 1$, with

$$\boldsymbol{y}^{(l)} = f^{(l)}(\boldsymbol{z}^{(l)}), \tag{3.2a}$$

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

$$\boldsymbol{z}^{(l)} = \boldsymbol{w}^{(l)}\boldsymbol{y}^{(l-1)} + \boldsymbol{b}^{(l)}. \tag{3.2b}$$

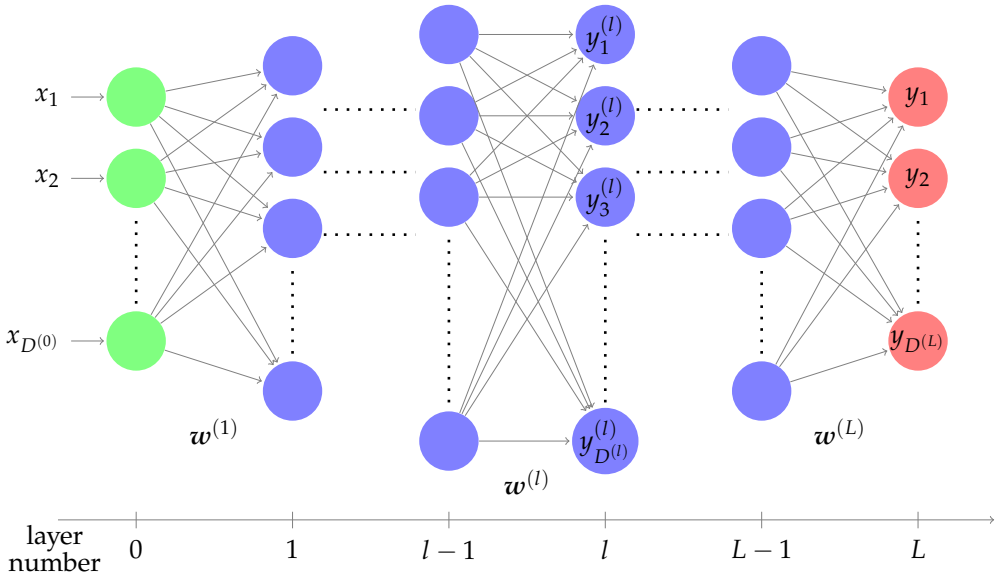

Figure 3.1: Schematic representation of a Multi Layer Perceptron (MLP) with $L-1$ hidden layers. $(x_1, x_2, \ldots, x_{D^{(0)}})$ are the inputs to the MLP, $\boldsymbol{w}^{(l)}$ are the weights connecting layer $l-1$ with layer l and $(y_1, y_2, \ldots, y_{D^{(L)}})$ are the outputs. Details on its working are given in the text.

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 called the *bias* vector and $\boldsymbol{y}^{(l-1)} \in \mathbb{R}^{D^{(l-1)}}$ are the values of the neurons in layer $l-1$. Some well known activation functions are

(i) 'Sigmoid', a smoothened step like function

$$f(x) = \frac{1}{1+e^{-x}}. \tag{3.3}$$

(ii) 'ReLU' (Rectified Linear Unit), a piecewise linear function, defined as

$$f(x) = \begin{cases} 0, & \text{for } x = 0, \\ x, & \text{for } x > 0. \end{cases} \tag{3.4}$$

(iii) 'Leaky ReLU' (Leaky Rectified Linear Unit) is a commonly used activation function in neural networks that addresses the "dying

ReLU" problem, where neurons can become inactive and always output zero. Unlike the standard ReLU, LeakyReLU allows a small, non-zero gradient when the input is negative, which helps keep the network learning.

$$f(x) = \begin{cases} x, & \text{for } x \geq 0, \\ \alpha x, & \text{for } x < 0. \end{cases} \tag{3.5}$$

Here, α is a small positive constant (typically $\alpha \approx 0.01$).

(iv) 'Softmax', converts a vector M real numbers into probability distribution of M possible outcomes, ensuring that the output values are non-negative and sum to 1. Given an input vector $\boldsymbol{x} = \{x_1, x_2, \ldots, x_M\}$ the softmax function defined as

$$f(x_m) = \frac{x_m}{\sum_{n=1}^{M} x_n} \quad \text{for } n = 1, \ldots, M. \tag{3.6}$$

The values of the input layer $l = 0$ are set to be the input feature vector $\boldsymbol{y}^{(0)} = \boldsymbol{x}$, while the values of the output layer $\boldsymbol{y}^{(L)}$ is the output that the model associates to the given input: it is called *predicted output* and we indicate it simply as $\boldsymbol{y} = \boldsymbol{y}^{(L)}$. The weight matrices, together with the biases and the activation functions, govern the collective influence of the neurons from one layer to the next.

The training process consists in finding weights $\boldsymbol{W} = \{\boldsymbol{w}^{(l)}\}_{l=1,2,\ldots,L}$ and biases $\boldsymbol{B} = \{\boldsymbol{b}^{(l)}\}_{l=1,2,\ldots,L}$ such that the model predicts the correct output given an input, and is able to generalize to unseen inputs. This is done by minimizing a cost function $C(\{\boldsymbol{y}_i, \hat{\boldsymbol{y}}_i\}_i | \boldsymbol{W}, \boldsymbol{B})$ which measures the average distance ('error') between the NN predicted output $\boldsymbol{y}_i$ and the target output $\hat{\boldsymbol{y}}_i$. For a simplest case, we can consider quadratic deviation between network's output and true output

$$C(\{\boldsymbol{y}_i, \hat{\boldsymbol{y}}_i\}_i | \boldsymbol{W}, \boldsymbol{B}) = \frac{1}{N} \sum_i (\boldsymbol{y}_i - \hat{\boldsymbol{y}}_i)^2 . \tag{3.7}$$

The minimization of the cost function is typically carried out using gradient descent or one of its variants such as stochastic gradient descent (SGD), Adam, or RMSprop. These optimization algorithms update the

network parameters $\boldsymbol{W}$ and $\boldsymbol{B}$, iteratively in the direction that reduces the cost function the most, i.e., along the negative gradient of the cost function. For l-th layer the update rule can be illustrated as

$$\boldsymbol{w}_{\text{new}}^{(l)} = \boldsymbol{w}^{(l)} - \lambda \frac{\partial C}{\partial \boldsymbol{w}^{(l)}}, \tag{3.8}$$

$$\boldsymbol{b}_{\text{new}}^{(l)} = \boldsymbol{b}^{(l)} - \lambda \frac{\partial C}{\partial \boldsymbol{b}^{(l)}}. \tag{3.9}$$

where λ is the learning rate, a hyperparameter controlling the step size of the update. Notice that, these algorithms require computing gradients of the cost function with respect to all the weights and biases in the network. An efficient way of doing it is using the *backpropagation* algorithm, which computes these gradients using *chain rule* to propagate error derivatives backward through the layers of the network. This process has two main steps:

1. **Forward pass**:

   - Initialization: Once the cost function is chosen, the model training begins with randomly initializing the weights and biases of the network.
   - Compute the output: The next step is for a given input vector $\boldsymbol{x_i}$, to calculate the activations of each layer sequentially as per eq. (3.2b) and eq. (3.2a). Eventually we calculate the cost function at the output layer L according to eq. (3.7).

2. **Backward pass** :

   - Next, we compute the gradient of the cost function with respect to all the weights and biases starting in the final layer L.

$$\frac{\partial C\left(\{\boldsymbol{y}_i, \hat{\boldsymbol{y}}_i\}_i | \boldsymbol{W}, \boldsymbol{B}\right)}{\partial w_{kr}^{(L)}} = \frac{2}{N} \sum_i \sum_j \left(y_{ij} - \hat{y}_{ij}\right) \frac{\partial y_{ij}}{\partial w_{kr}^{(L)}} \tag{3.10}$$

   where, y_{ij} is the j-th element of the vector $\boldsymbol{y}_i^{(L)}$ representing the label of i -th sample. Now, to compute the partial derivative

w.r.t the network parameters, we apply the chain rule.

$$\begin{aligned}\frac{\partial y_{ij}}{\partial w_{kr}^{(L)}} &= f'^{(L)}\left(z_{ij}^{(L)}\right)\frac{\partial z_{ij}^{(L)}}{\partial w_{kr}^{(L)}} \\ &= f'^{(L)}\left(z_{ij}^{(L)}\right) y_{ir}^{(L-1)}\delta_{jk},\end{aligned} \tag{3.11}$$

where δ is the Kronecker delta. Similarly,

$$\begin{aligned}\frac{\partial y_{ij}}{\partial b^{(L)}} &= f'^{(L)}\left(z_{ij}^{(L)}\right)\frac{\partial z_{ij}^{(L)}}{\partial b^{(L)}} \\ &= f'^{(L)}\left(z_{ij}^{(L)}\right).\end{aligned} \tag{3.12}$$

- Recursively compute gradients for each layer's (l) weights $w_k^{(l)}$ and biases $b_k^{(l)}$ by applying the chain rule backwards through the network and store them.

Finally, the network parameters are updated following eq. (3.8) and (3.9). By iteratively performing these steps over many data samples (or batches), the model learns to approximate the desired function $f(\boldsymbol{x})$ and generalize to unseen data.

There are also other kind of cost functions that are commonly used in ML depending on the nature of the learning task. For classification problems especially when the output represents probabilities over multiple classes—a widely used choice is the *categorical cross-entropy* cost function. If the true output $\hat{\boldsymbol{y}}_i$ and the predicted output $\boldsymbol{y}_i$ are probability distributions over M classes (i.e., $\sum_{m=1}^{M}\hat{y}_{i,m} = 1$ and $\sum_{m=1}^{M} y_{i,m} = 1$), the cost function is given by

$$C_{\text{CE}}\left(\{\boldsymbol{y}_i, \hat{\boldsymbol{y}}_i\}_i | \boldsymbol{W}, \boldsymbol{B}\right) = -\frac{1}{N}\sum_i \sum_{m=1}^{M} \hat{y}_{i,m}\log y_{i,m}, \tag{3.13}$$

where $y_{i,m}$ denotes the predicted probability for class m and $\hat{y}_{i,m}$ is 1 if sample i belongs to class m, and 0 otherwise (one-hot encoded labels). For binary classification problems, where the output is a single probability

$y_i \in [0, 1]$ representing the likelihood of class 1, the appropriate loss is the binary cross-entropy

$$C_{\text{BCE}}(\{y_i, \hat{y}_i\}_i | \boldsymbol{W}, \boldsymbol{B}) = -\frac{1}{N} \sum_i \left[\hat{y}_i \log y_i + (1 - \hat{y}_i) \log(1 - y_i)\right]. \quad (3.14)$$

Chapter 4

Noise classification in three-level quantum networks by Machine Learning

4.1 Introduction

Nowadays, quantum systems can be controlled with impressive accuracy [50] to perform new tasks by exploiting their coherence properties, such as superposition and entanglement [51]. However, quantum behavior fades away because of environmental noise that leads to a loss of accuracy of quantum operations and decoherence [52]. 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 [49, 53, 54]. ML-based approaches have been applied to various quantum control protocols [55–58], to quantum state tomography [59–61], to characterize quantum systems [22, 62, 63], to simulate open quantum systems [22, 59, 64, 65], to study non-Markovian dynamics [66, 67] and, in a similar spirit similar to this work, to characterize system-environment interactions [21, 68] and perform noise spectroscopy [69–71].

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 [23, 72]. Examples of active stabilization strategies include Quantum Error Correction which relies on encoding nonlocal information [73, 74], and Dynamical decoupling [75–78], 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 [23, 79–81]. 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 [82], photonic qubits [83, 84], nuclear magnetic resonance [85], and nitrogen vacancy centres in diamond [86]. Superconducting systems are an outstanding example in this context [87] whose decoherence times have improved from a few nanoseconds to milliseconds in 15 years [88].

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 [19] or different nodes of the quantum architecture [89] 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 [27]. The effects of spatially correlated low-frequency noise between two quantum devices have been observed for several decades [90] and their impact on decoherence has been studied [89] 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 [91–93]. 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 infor-

mation 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 [78] useful in the investigation of the effect that non-Markovianity has in the emergence of many-body phenomena [94, 95].

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 [96]. 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) [97, 98], 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 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 chapter is structured as follows: in Sec. 4.2 we define the models of system and noise that we deal with in our study. In Sec. 4.3 we present

the CTAP protocol and the figures of merit that we use in the training phase. In Sec. 4.4 we give a brief overview of *supervised learning* and *neural networks* (NN) [42, 45, 49, 99] and describe how we classify among the noises introduced in Sec. 4.2. Section 4.5 presents the results of the classification for a finite number of samples and imperfect measurements. The possible physical implementations are discussed in Sec. 4.6 together with a physical interpretation of the results. The conclusions are drawn in Sec. 4.7.

4.2 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 [97]. The tunneling rates between the first and the second dot, $\Omega_{\mathrm{p}}(t)$, and between the second and the third dot, $\Omega_{\mathrm{s}}(t)$, 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 [100].

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

$$H_0 = \delta_{\mathrm{p}}|1\rangle\langle 1| + \delta|2\rangle\langle 2|, \tag{4.1a}$$

$$H_{\mathrm{c}}(t) = \frac{\Omega_{\mathrm{p}}(t)}{2}\left(|0\rangle\langle 1| + |1\rangle\langle 0|\right) + \frac{\Omega_{\mathrm{s}}(t)}{2}\left(|1\rangle\langle 2| + |2\rangle\langle 1|\right), \tag{4.1b}$$

where, in the language of the three-dot network, we defined the *detunings* $\delta_{\mathrm{p}} = \epsilon_1 - \epsilon_0$ 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 [92]. We model it as the following classical stochastic process (see Sec. 2.1 and Sec. 2.2) added to the diagonal entries of the Hamiltonian

$$H_{\mathrm{noise}}(t) = \tilde{x}_1(t)|1\rangle\langle 1| + \tilde{x}_2(t)|2\rangle\langle 2|. \tag{4.2}$$

Here, $\tilde{x}_i(t)$ $(i = 1, 2)$ is a random variable [29] with normalized probability density $p(x_i, t)$. The total Hamiltonian $H(t)$ in the basis $\{|0\rangle, |1\rangle, |2\rangle\}$ thus

reads

$$H(t) = H_{\rm sys}(t) + H_{\rm noise}(t) = \begin{pmatrix} 0 & \Omega_{\rm p}(t)/2 & 0 \\ \Omega_{\rm p}(t)/2 & \delta_{\rm p} + \tilde{x}_1(t) & \Omega_{\rm s}(t)/2 \\ 0 & \Omega_{\rm s}(t)/2 & \delta + \tilde{x}_2(t) \end{pmatrix}. \quad (4.3)$$

We now link the five classes of noise listed in Sec. 4.1 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 $\eta < 0$[1];

  (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

$$\langle \tilde{x}_i(t) \rangle = 0, \quad \langle \tilde{x}_i(t)\tilde{x}_i(t') \rangle = \gamma\delta(t - t') \quad (4.4)$$

  where the auto-correlation function is defined in eq. (2.22). We 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 $\eta < 0$.

[1] $\eta < 0$ leads to larger detuning than $\eta > 0$

4.3 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 [100, 101]. 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.

4.3.1 CTAP/STIRAP

CTAP [97] (and its spin equivalent spin-CTAP [98]) 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 $\sim 100\%$-efficiency population transfer with remarkable robustness against parametric fluctuations and state-selectivity.

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 $H_{\rm sys}(t)$ of the form

$$|\phi_{\rm D}(t)\rangle = \cos\theta(t)|0\rangle - \sin\theta(t)|2\rangle \tag{4.5}$$

with $\theta(t) = \tan^{-1}\left[\Omega_{\rm p}(t)/\Omega_{\rm s}(t)\right]$. The protocol is operated in the time interval $[t_{\rm i}, t_{\rm f}]$ by applying pulses $\Omega_{p,s}(t)$ in a *counterintuitive* manner [100]: $\Omega_{\rm s}$ is applied *before* $\Omega_{\rm p}$, while ensuring that the two pulses overlap in a fraction of the duration of the sequence. In this case, the trapped state $|\phi_{\rm D}(t)\rangle$ at the initial time $t_{\rm i}$ coincides with $|0\rangle$ and at the final time $t_{\rm f}$ with the target state $|2\rangle$. If the evolution is adiabatic and the system is prepared in $|0\rangle = |\phi_{\rm D}(t_{\rm i})\rangle$, the system evolves following the trapped state $|\phi_{\rm D}(t)\rangle$ throughout the evolution as prescribed by the adiabatic theorem [102, 103]. The requested adiabaticity of the process can be cast in the form of the global adiabaticity condition [100]

$$\Omega_{\rm p/s}^{\rm max}\tau \geq 10, \tag{4.6}$$

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 [104] have been used in literature. In this work, we consider Gaussian pulses (see Fig. 4.1) of the form

$$\Omega_p(t) = \Omega_p^{max} e^{-(\frac{t-\tau}{T})^2}, \qquad \Omega_s(t) = \Omega_s^{max} e^{-(\frac{t+\tau}{T})^2}, \tag{4.7}$$

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 popula-

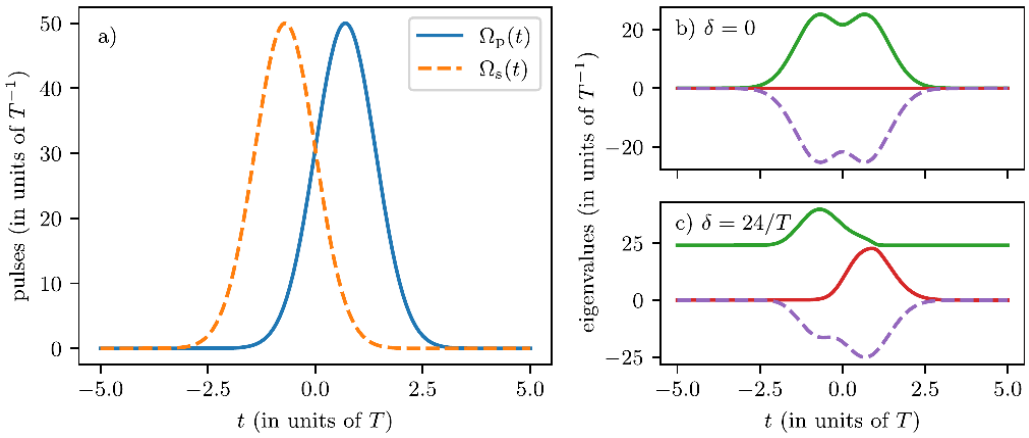

Figure 4.1: (a) Example of the pump $\Omega_p(t)$ and Stokes $\Omega_s(t)$ pulses, Eqs. (4.7), as a function of time t. The parameters used are $\Omega_p^{max} = \Omega_s^{max} = 50/T$, $\tau = 0.7T$ and $t \in [-5T, 5T]$. (b) Instantaneous eigenvalues for $\delta = 0$: CTAP/STIRAP achieves 100% population 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.

tion 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 4.B.

4.3.2 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 [105]

$$\xi = \lim_{N\to\infty} \frac{1}{N} \sum_{r=1}^{N} |\langle 2|\psi^{(r)}(t_\mathrm{f})\rangle|^2 = \lim_{N\to\infty} \frac{1}{N} \sum_{r=1}^{N} \xi^{(r)} \tag{4.8}$$

with $\xi^{(r)} = |\langle 2|\psi^{(r)}(t_\mathrm{f})\rangle|^2$ 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 Eq. (4.8) is obtained by averaging over the random processes

$$\xi = \int \xi(x_1, x_2) p(x_1, x_2) \mathrm{d}x_1 \mathrm{d}x_2. \tag{4.9}$$

If the noise is (anti-)correlated, then $x_2^{(r)} = \eta x_1^{(r)}$ and Eq. 4.9 simplifies to

$$\xi = \int \xi(x_1, \eta x_1) p_1(x_1) \mathrm{d}x_1. \tag{4.10}$$

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. 4.A for the derivation)

$$\dot{\rho}(t) = -\mathrm{i}[H_0 + H_\mathrm{c}(t), \rho(t)] - \frac{\gamma}{2}(O^2\rho(t) + \rho(t)O^2 - 2O\rho(t)O), \tag{4.11}$$

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

$$\xi = \langle 2|\rho(t_\mathrm{f})|2\rangle. \tag{4.12}$$

4.4 Classification with neural networks

4.4.1 Neural Networks

Supervised learning requires a labeled dataset $\{(\boldsymbol{x}_i, \hat{\boldsymbol{y}}_i)\}_{i=1,\dots,N}$, 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). See Sec. 3.1 and 3.2 for more details about supervised learning and MLP.

We use a neural network consisting of two hidden layers with 128 and 100 neurons, respectively, as detailed in Table 4.1. For the hidden layers, we employ the *leaky rectified linear unit* (LeakyReLU) [106] as the activation function, defined as:

$$f_{\text{LReLU}}(z) = \begin{cases} z & \text{if} \quad z \geq 0, \\ \alpha z & \text{if} \quad z < 0, \end{cases} \tag{4.13}$$

where α is set to 0.01. This activation function helps to mitigate the "dying ReLU" problem [107], where neurons become inactive and only output zero.

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

$$\boldsymbol{y} = f_{\text{softmax}}(\boldsymbol{z}^{(L)}) = \frac{e^{\boldsymbol{z}^{(L)}}}{\sum_{k=1}^{D^{(L)}} e^{z_k^{(L)}}}, \tag{4.14}$$

Layer	# neurons	Activation function
input	3	
hidden 1	128	LeakyReLU
hidden 2	100	LeakyReLU
output	4 or 5	Softmax

Table 4.1: Structure of the neural network used for classification of the 4 or 5 classes of noise.

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.

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

$$C(\{\boldsymbol{y}_i, \hat{\boldsymbol{y}}_i\}_i) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{D^{(L)}} \hat{y}_{ij}\ln(y_{ij}), \tag{4.15}$$

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.

4.4.2 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 Sec. 4.2. 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 [69, 108, 109], 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 Eqs. (4.7) we consider: 1. $\Omega_{\mathrm{p}}^{\max} = \Omega_{\mathrm{s}}^{\max}$, 2. $\Omega_{\mathrm{p}}^{\max} > \Omega_{\mathrm{s}}^{\max}$, 3. $\Omega_{\mathrm{p}}^{\max} < \Omega_{\mathrm{s}}^{\max}$. 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 Sec. 4.3. We denote the efficiencies as $\xi_{\mathrm{p=s}}$, $\xi_{\mathrm{p>s}}$, and $\xi_{\mathrm{p<s}}$, for the driving conditions (i), (ii), and (iii) respectively, and we use $\boldsymbol{x} = (\xi_{\mathrm{p=s}}, \xi_{\mathrm{p>s}}, \xi_{\mathrm{p<s}})$ as input feature vector to the NN, see Fig. 4.2. The parameters we use are (i) $\Omega_{\mathrm{p}}^{\max} = \Omega_{\mathrm{s}}^{\max} = 50/T$, (ii) $\Omega_{\mathrm{p}}^{\max} = 20\sqrt{10}/T > \Omega_{\mathrm{s}}^{\max} = 10\sqrt{10}/T$, and (iii) $\Omega_{\mathrm{p}}^{\max} = 10\sqrt{10}/T < \Omega_{\mathrm{s}}^{\max} = 20\sqrt{10}/T$ 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 $[0.1, 5]$

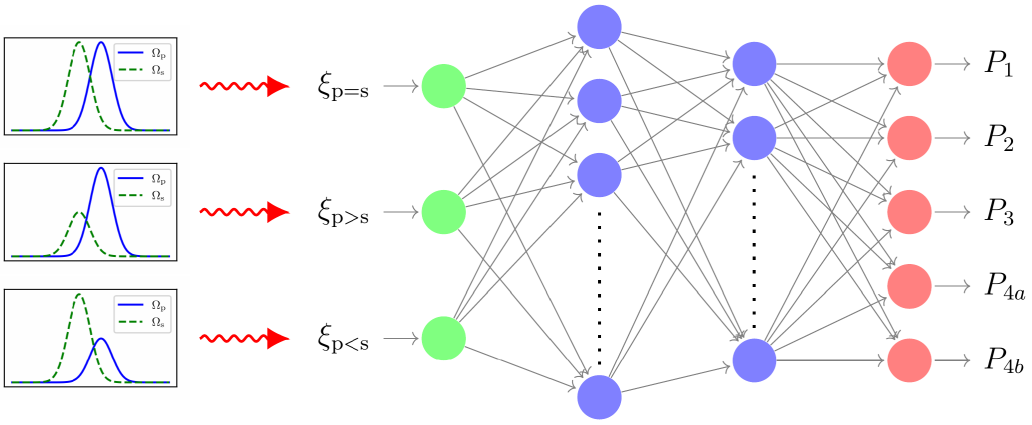

Figure 4.2: Schematic representation of our model. The efficiencies obtained with the three driving conditions $\Omega_p^{max} = \Omega_s^{max}$, $\Omega_p^{max} > \Omega_s^{max}$, and $\Omega_p^{max} < \Omega_s^{max}$ are given as input to the neural network. The Neural Network is composed of 2 hidden layers with 128 and 100 neurons respectively, and activation function LeakyReLU, Eq. (4.13). 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, Eq. (4.14), on the output layer.

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 Eq. (4.10) for non-Markovian quasistatic noise (with $p_1(x_1, t)$ being a Gaussian distribution) and Eq. (4.12) 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 $p_2(x_2, t)$ is randomly sampled in the interval $[-5\sigma_1, 5\sigma_1]$. As mentioned in Sec. 4.4.1, 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

$$A = \frac{1}{N}\sum_{i=1}^{N}\delta(\arg\max \boldsymbol{y}_i, \arg\max \hat{\boldsymbol{y}}_i), \tag{4.16}$$

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

4.5 Results

The ML model and the training are performed with TensorFlow [110] and the results are shown in Fig. 4.3. The accuracy A for the training and validation sets is shown in Fig. 4.3(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 Fig. 4.3(b) how the value of the cost function C changes during training. In Fig. 4.4, 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 Fig. 4.4(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

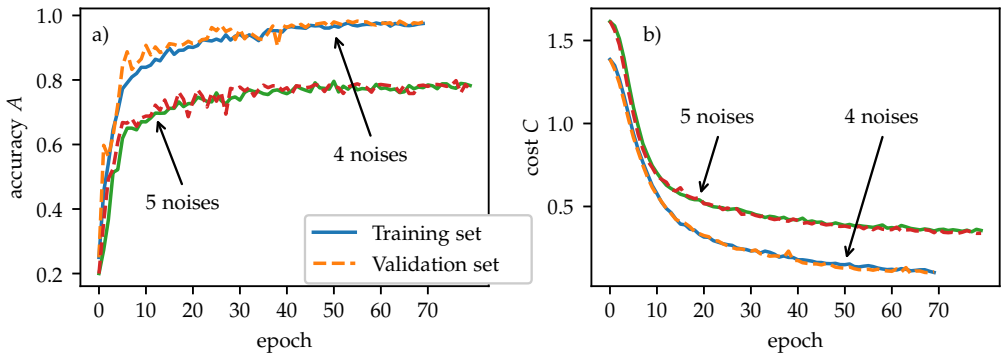

Figure 4.3: (a) Accuracy, Eq. (4.16), and (b) value of the cost function, Eq. (4.15), for the training (solid blue for four noises and solid green for five noises) and validation (dashed orange for four noises and dashed red for five noises) sets versus the number of epochs of training. The accuracy on the test set is $A \approx 1$ for four noise classes and $A \approx 0.81$ for five noise classes.

C during the training are shown in Fig. 4.3. The confusion matrix for this model is reported in Fig. 4.4(b) showing that now the four classes are clearly distinguishable with the chosen input features.

4.5.1 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 [111], 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 Sec. 4.4.2 but compute the efficiency ξ_M as the one resulting from a set of individual "quantum" trajectories. The result of unraveling 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 $\boldsymbol{y}$ is produced by averaging over the M quantum

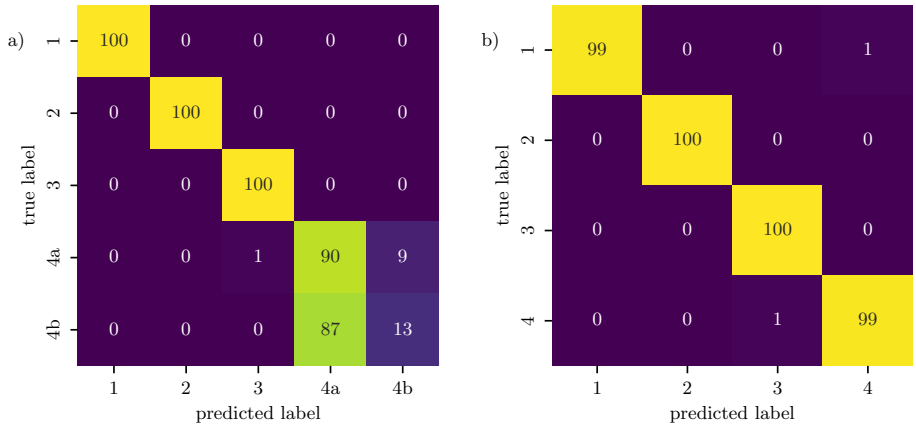

Figure 4.4: Confusion matrix for the MLP model for the classification of the classes of noise considered in this work. Each row of the matrix represents the instances of the true classes, while each column represents the instances of the predicted classes. It is apparent that classes 4*a* and 4*b*, i.e. Markovian correlated and anti-correlated, are not easily distinguishable.

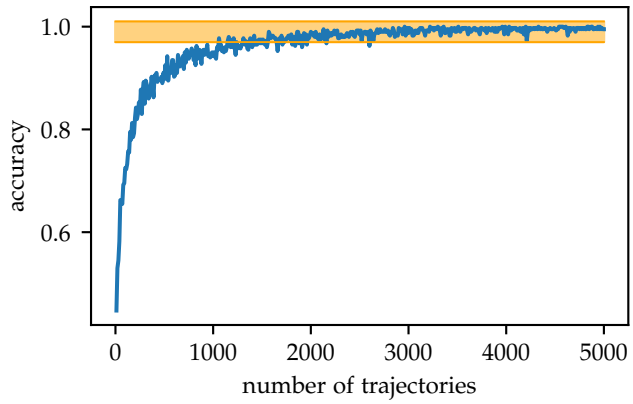

Figure 4.5: Accuracy of the classification task for four classes of noise versus the number M of projective measurements on quantum trajectories. The shaded orange region corresponds to the accuracy obtained with ideal measurements.

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 Fig. 4.5, 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 $A \approx [0.97, 1]$.

4.6 Physical implementation and interpretation of the results

4.6.1 Physical implementations of the model

Spatially correlated noise in solid-state quantum information processing is a topic of current experimental investigation in superconducting [91] and semiconducting [93] quantum devices. Charge- and spin-CTAP have been studied extensively [98, 112] and recently a similar protocol has been implemented in semiconducting quantum dots [113] while CTAP of electromagnetic excitations has been proposed as the building block of quantum operations in the ultrastrong coupling regime [114, 115] circuit-QED architectures. In superconducting devices, STIRAP has been demonstrated in Vee configuration [116–118] whereas the Lambda configuration we study in this work could be implemented either directly [119–121] or by a detuning-modulated protocol [122] which generalizes hyper-STIRAP [100] and bypasses parity selection rules at noise-protected operating points.

The model of noise studied in this work 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 [101] 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.

4.6.2 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 $\delta_{\rm p}$) for the three driving conditions, shown in Fig. 4.6

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. Fig.4.6]. The efficiency for correlated and anti-correlated noise types, is calculated as a weighed average over the line $\delta = \eta\delta_{\rm p}$, see Eq. (4.10). The corresponding efficiency depends on the ratio between $\Omega_{\rm p}^{\rm max}$ and $\Omega_{\rm s}^{\rm max}$, this dependence being different for $\eta > 0$ or $\eta < 0$. Fig. 4.6 reports as an example the lines $\delta = \eta\delta_{\rm p}$ for $\eta \gtrless 0$: it is clear that for $\eta > 0$ the efficiency of CTAP with $\Omega_{\rm p}^{\rm max} < \Omega_{\rm s}^{\rm max}$ is higher than the one with $\Omega_{\rm p}^{\rm max} > \Omega_{\rm s}^{\rm max}$, and vice versa for $\eta < 0$. For a more extensive analysis of this behavior we refer to Ref. [123]. For uncorrelated noise, instead, the average is over the whole $\delta - \delta_{\rm p}$ plane – cf. Eq. (4.9) – 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 Fig. 4.6. Instead, the noise assumes all values along the line $\delta = \eta\delta_{\rm p}$ during each repetition of the protocol.

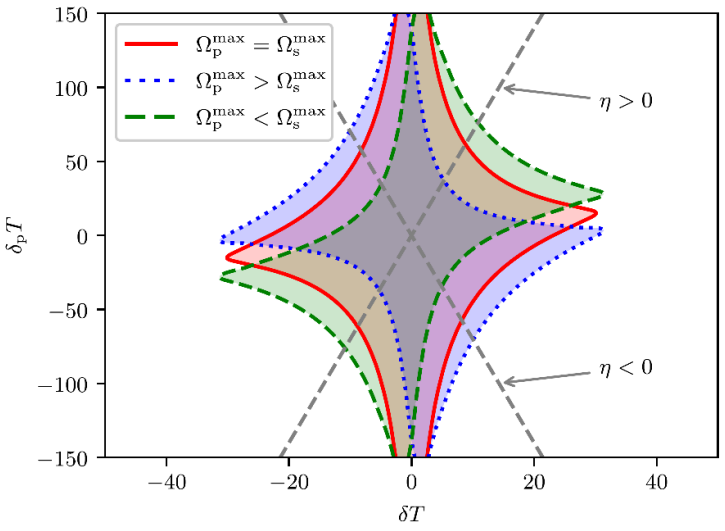

Figure 4.6: Efficiency of CTAP for the three driving conditions (i) $\Omega_p^{max} = \Omega_s^{max} = 50/T$ (red, solid line), (ii) $\Omega_p^{max} = 20\sqrt{10}/T > \Omega_s^{max} = 10\sqrt{10}/T$ (blue, dotted line), and (iii) $\Omega_p^{max} = 10\sqrt{10}/T < \Omega_s^{max} = 20\sqrt{10}/T$ (green 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 Eqs. (4.9) and (4.10). The dashed gray lines represent correlation ($\eta > 0$) or anti-correlation ($\eta < 0$) between the detunings δ and δ_p. The other parameter is $\tau = 0.7T$.

4.7 Conclusions

In this Chapter, 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 distin-

guish 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.

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 [120, 124] and extend the diagnostics to more complex probes, such as interacting multiqubit systems [92], quantum dot chains [125], strongly [126] and ultrastrongly coupled [114, 115] circuit-QED architectures. In these cases control could be implemented by unconventional CTAP/STIRAP protocols operated by detunings [55, 127]. 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 code and the data that support the findings of this study are openly available at the following URL/DOI: `https://github.com/Shreyasi31/SmallQNetNoiseML`.

Appendix

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

The Lindblad master equation Eq. (4.11) describing of the dynamics of the system subject to Markovian noise is derived as follows. We start from the total Hamiltonian of the system $H = H_{\text{sys}} + H_{\text{noise}}$, where H_{noise} is given by Eq. (4.2). Using the correlation $x_2(t) = \eta x_1(t)$, we write

$$H_{\text{noise}} = \tilde{x}_1(t)\left(|1\rangle\langle 1| + \eta|2\rangle\langle 2|\right) = \tilde{x}_1(t)O, \tag{4.17}$$

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".

$$\widetilde{\rho}(t) = e^{\frac{\mathrm{i}}{\hbar}H_{\text{sys}}t}\rho(t)e^{-\frac{\mathrm{i}}{\hbar}H_{\text{sys}}t}, \tag{4.18}$$

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

$$\dot{\widetilde{\rho}}(t) = -\frac{\mathrm{i}}{\hbar}[\tilde{H}_{\text{noise}}(t), \widetilde{\rho}(t)], \tag{4.19}$$

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

$$\widetilde{\rho}(t) = \widetilde{\rho}(0) - \frac{\mathrm{i}}{\hbar}\int_0^t [\tilde{H}_{\text{noise}}(t'), \widetilde{\rho}(t')]dt'. \tag{4.20}$$

Substituting $\tilde{\rho}(t)$ in Eq. (4.19) we obtain

$$\begin{aligned}\dot{\widetilde{\rho}}(t) = &-\frac{\mathrm{i}}{\hbar}e^{\frac{\mathrm{i}}{\hbar}H_{\text{sys}}t}[\widetilde{x}_1(t)O\rho(0) - \rho(0)O\widetilde{x}_1(t)]e^{-\frac{\mathrm{i}}{\hbar}H_{\text{sys}}t} \\ &-\frac{1}{\hbar^2}\int_0^t [\tilde{H}_{\text{noise}}(t), [H_{\text{noise}}(t'), \widetilde{\rho}(t')]]dt'.\end{aligned} \tag{4.21}$$

Averaging over the stochastic variable at time t, we obtain

$$\langle\dot{\widetilde{\rho}}(t)\rangle = -\frac{1}{\hbar^2}\left\langle\int_0^t [\widetilde{x}(t)\widetilde{O}(t), [\widetilde{x}(t')\widetilde{O}(t'), \widetilde{\rho}(t')]]dt'\right\rangle, \tag{4.22}$$

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

$$\langle\dot{\widetilde{\rho}}(t)\rangle = -\frac{\gamma}{2}\left(\widetilde{O}^2(t)\tilde{\rho}(t) + \widetilde{\rho}(t)\widetilde{O}^2(t) - 2\widetilde{O}(t)\widetilde{\rho}(t)\widetilde{O}(t)\right). \tag{4.23}$$

As expected, this equation can be recast in the Lindblad form 2.61. In the Schrödinger picture we obtain

$$\langle\dot{\rho}(t)\rangle = -\frac{\mathrm{i}}{\hbar}[H_{\text{sys}}, \rho(t)] - \frac{\gamma}{2}\left(O^2\rho(t) + \rho(t)O^2 - 2O\rho(t)O\right). \tag{4.24}$$

4.B Adiabatic passage in a three-node network

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

$$H_0 = \delta_{\text{p}}|1\rangle\langle 1| + \delta|2\rangle\langle 2|, \tag{4.25a}$$

$$H_{\text{c}}(t) = \frac{\Omega_{\text{p}}(t)}{2}\left(|0\rangle\langle 1| + |1\rangle\langle 0|\right) + \frac{\Omega_{\text{s}}(t)}{2}\left(|1\rangle\langle 2| + |2\rangle\langle 1|\right), \tag{4.25b}$$

and where $\delta_p = \epsilon_1 - \epsilon_0$ and $\delta = \epsilon_2 - \epsilon_0$ define the detunings. CTAP [97] 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 [100, 101] 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 $\Omega_p(t)$ and $\Omega_s(t)$. The condition $\delta \approx 0$ is important for the successful population transfer [101]. For $\delta = 0$, the instantaneous eigenstates of $H_{sys}(t)$ are

$$|\phi_D(t)\rangle = \cos\theta(t)|0\rangle - \sin\theta(t)|2\rangle, \tag{4.26a}$$
$$|\phi_-(t)\rangle = \sin\theta(t)\cos\phi(t)|0\rangle - \sin\phi(t)|1\rangle + \cos\theta(t)\cos\phi(t)|2\rangle, \tag{4.26b}$$
$$|\phi_+(t)\rangle = \sin\theta(t)\sin\phi(t)|0\rangle - \cos\phi(t)|1\rangle + \cos\theta(t)\sin\phi(t)|2\rangle, \tag{4.26c}$$

with eigenvalues

$$\lambda_D(t) = 0, \tag{4.27a}$$
$$\lambda_-(t) = -\frac{\hbar}{2}\sqrt{\Omega_p(t)^2 + \Omega_s(t)^2}\tan\phi(t), \tag{4.27b}$$
$$\lambda_+(t) = \frac{\hbar}{2}\sqrt{\Omega_p(t)^2 + \Omega_s(t)^2}\cot\phi(t), \tag{4.27c}$$

where

$$\theta(t) = \tan^{-1}\left(\frac{\Omega_p(t)}{\Omega_s(t)}\right), \tag{4.28}$$
$$\phi(t) = \tan^{-1}\left(\frac{\sqrt{\Omega_p(t)^2 + \Omega_s(t)^2}}{\delta_p + \sqrt{\delta_p^2 + \Omega_p(t)^2 + \Omega_s(t)^2}}\right). \tag{4.29}$$

The eigenstate $|\phi_D(t)\rangle$ 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 Ω_s is applied before Ω_p:

$$\lim_{t\to t_i}\frac{\Omega_p(t)}{\Omega_s(t)} = \lim_{t\to t_f}\frac{\Omega_s(t)}{\Omega_p(t)} = 0, \tag{4.30}$$

then the dark state $\phi_D(t)$ at the initial time t_i will coincide with the initial state $|0\rangle$ and at the final time t_f will coincide with the target state $|2\rangle$. If the evolution is adiabatic and the system is initially prepared in $|0\rangle = |\phi_D(t_i)\rangle$, by the adiabatic theorem [102, 103], the system evolves following the dark state $|\phi_D(t)\rangle$ 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 [97, 100]

$$\hbar|\langle\phi_\pm|\dot{\phi}_D\rangle| \ll |\lambda_0 - \lambda_\pm|. \tag{4.31}$$

Using the expressions for the dressed basis, Eqs. (4.26) and Eqs. (4.27), equation (4.31) is recast in terms of the pulses [56, 100, 101]

$$|\dot{\theta}(t)| \ll \frac{1}{2}\left|\delta_p \pm \sqrt{\delta_p^2 + \Omega_p(t)^2 + \Omega_s(t)^2}\right|, \tag{4.32}$$

yielding the so-called "local adiabaticity condition" which has to hold for all times $t \in [t_i, t_f]$. By time averaging Eq. (4.32) over τ and assuming $\delta_p \ll \Omega_p^{max}, \Omega_s^{max}$ one can obtain the weaker "global adiabaticity condition" which is often reported as [100]

$$\Omega_{p/s}^{max}\tau \geq 10, \tag{4.33}$$

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

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

$$\Omega_p(t) = \Omega_p^{max} e^{-(\frac{t-\tau}{T})^2}, \tag{4.34a}$$

$$\Omega_s(t) = \Omega_s^{max} e^{-(\frac{t+\tau}{T})^2}, \tag{4.34b}$$

and we let the system evolve in the time interval $t \in [-5T, 5T]$ with $\tau = 0.7T$, see Fig. 4.1 for an example of pulses.

Chapter 5

Detection of noise correlations in two qubit system by Machine Learning

5.1 Introduction

In the preceding chapter, we demonstrated that machine learning (ML), combined with the sensitivity of coherent control protocols like CTAP, enables effective classification of spatially correlated non-Markovian noise classes and between non-Markovian and Markovian noise classes in a three-level quantum network. Here, we extend this approach to the more fundamental and practically relevant case of a two-qubit system, which constitutes the building block of quantum processors.

While protection of *single* qubits beyond the error-correction threshold has been demonstrated [128, 129] substantial work has still to be done for multi-qubit systems. In particular, besides single-qubit dephasing and relaxation channels, spatially correlated noise may emerge from substrates, shared control lines or collective electromagnetic modes [123]. Correlations affect quantum-error-correction thresholds since otherwise safe assumptions (e.g., independent error models) become invalid.

Machine learning (ML) has recently established itself as a versatile tool offering innovative diagnostic and control methods for quantum de-

vices [49, 53–61]. ML-based techniques have been proposed for the classification of time-correlated noise in quantum networks [130] and correlation of energy fluctuations in a three-level system [131].

In this chapter, we propose a methodology to classify noise correlations in a two-qubit system by the fundamental building block of any programmable quantum processor.

We consider a system where two qubits are ultrastrongly coupled and affected by local noise. We design a control protocol based on STIRAP/CTAP that is sensitive to noise correlations and Markovianity and involves only global, symmetric drives. This is a well suited to most superconducting-qubit, trapped-ion and quantum-dot architectures.

We propose a minimal data-acquisition scheme consisting of three STIRAP efficiency measurements with different peak-amplitude ratios. Using numerical simulations, we show that the resulting three-component feature vector is sufficient to train a multi-layer perceptron that reaches 90% overall accuracy and almost perfect discrimination between Markovian and non-Markovian noise. Crucially, also the Markovian noise correlations are distinguished by this method.

This chapter is organised as follows. Sec. 5.2 introduces the two-qubit Hamiltonian and the stochastic noise models. Sec. 5.3 details the control protocol and the performance metrics used for noise characterization. The ML framework is described in Sec. 5.4, and simulation results are presented and discussed in Sec. 5.5. Finally, we summarize our findings and outline future perspectives in Sec. 5.6.

5.2 System and Model

Our analysis focuses on two ultra-strongly coupled qubits with energy ε_1 and ε_2 and coupling strength $g \sim \epsilon_{1/2}$, see Fig. 5.2.1. The Hamiltonian of the system is ($\hbar = 1$)

$$H_{\text{sys}} = -\frac{\varepsilon_1}{2}\sigma_1^z - \frac{\varepsilon_2}{2}\sigma_2^z + \frac{g}{2}\left(\sigma_1^-\sigma_2^+ + \sigma_1^+\sigma_2^-\right), \tag{5.1}$$

where σ_i^z is the Pauli z operator, and σ_i^- and σ_i^+ are the lowering and raising operator, respectively, for qubit i. If the qubits have the same

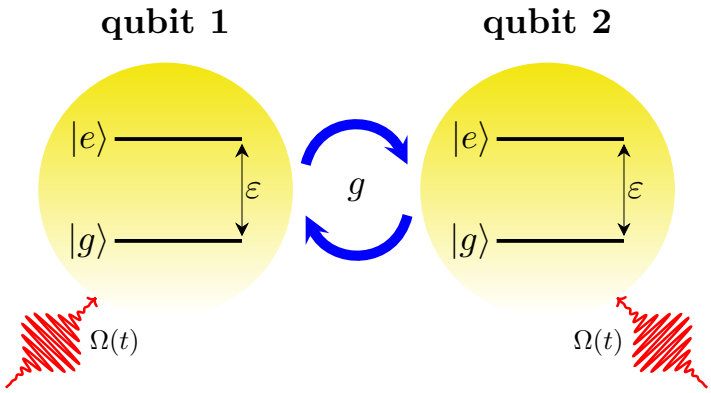

Figure 5.2.1: Schematic representation of the system under consideration: two qubits, each with energy ϵ, coupled with strength g. Both qubits are simultaneously driven by an external field characterized by a common time-dependent Rabi frequency $\Omega(t)$.

energy, i.e., $\varepsilon_1 = \varepsilon_2 = \varepsilon$, the eigenstates of Hamiltonian (5.1) are

$$\begin{aligned} |0\rangle &\equiv |gg\rangle, & |1\rangle &\equiv |ee\rangle, \\ |2\rangle &\equiv \frac{1}{\sqrt{2}}\left(|ge\rangle + |eg\rangle\right), & |3\rangle &\equiv \frac{1}{\sqrt{2}}\left(|ge\rangle - |eg\rangle\right), \end{aligned} \tag{5.2}$$

with energies $\mathcal{E} = \left\{-\varepsilon, \varepsilon, \frac{g}{2}, -\frac{g}{2}\right\}$, respectively.

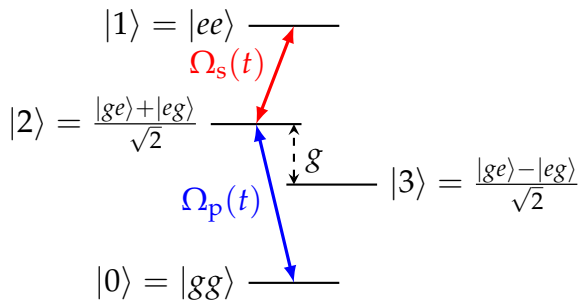

Figure 5.2.2: Energy levels of the two-qubit system, eigenenergies of Hamiltonian eq. (5.1).

5.2.1 Noise

We consider local noise, specifically noise affecting the energy levels acting individually on each qubit, which may rise, for instance, from charge noise [132]. We model it as stochastic processes whose effect is to modify the energy splittings of the qubits, as expressed by

$$H_{\text{noise}}(t) = -\frac{\delta_1(t)}{2}\sigma_1^z - \frac{\delta_2(t)}{2}\sigma_2^z. \tag{5.3}$$

Here $\delta_i(t)$, $i = 1, 2$, are random variables [29] with probability density $p_i(\delta_i, t)$.

The Hamiltonian of the system including the noise in the eigenbasis of H_{sys} reads

$$\begin{aligned}\tilde{H}_{\text{sys}}(t) &= H_{\text{sys}} + H_{\text{noise}}(t) \\ &= \left(\varepsilon + \frac{\delta_1(t) + \delta_2(t)}{2}\right)|1\rangle\langle 1| + \frac{g}{2}\left(|2\rangle\langle 2| - |3\rangle\langle 3|\right) \\ &\quad - \left(\varepsilon + \frac{\delta_1(t) + \delta_2(t)}{2}\right)|0\rangle\langle 0| + \left(\frac{\delta_1(t) - \delta_2(t)}{2}|3\rangle\langle 2| + \text{h.c.}\right).\end{aligned} \tag{5.4}$$

The effect of the noise is to shift the energy of the states $|0\rangle$ and $|1\rangle$ and to couple the states $|2\rangle$ and $|3\rangle$.

We consider three Non-Markovian and two Markovian noise classes:

- **Non-Markovian noise:** In this context we address specifically the case of *quasistatic noise*, where the noise mechanism has a long correlation time. Each of the random variables $\delta_1(t)$ and $\delta_2(t)$ thus, assume a value which is constant throughout a single time evolution of the system; however, they vary across different evolutions. The values of $\delta_i(t)$ are picked from a Gaussian distribution with zero mean and standard deviation $\sigma = 10^{-2}\epsilon$, and describes the cumulative effects of independent microscopic sources. We identify three distinct classes:

  (1) Correlated: $\delta_2(t) = \eta\delta_1(t)$ with $\eta > 0$;

  (2) Anti-correlated: $\delta_2(t) = \eta\delta_1(t)$ with $\eta < 0$;

  (3) Uncorrelated: $\delta_2(t)$ and $\delta_1(t)$, are independent of each other.

- **Markovian noise:** We consider zero-mean $\langle \delta_i(t) \rangle = 0$, delta-correlated $\langle \delta_i(t)\delta_i(t') \rangle = \gamma\delta(t-t')$ stochastic processes, making the dynamics of the two-qubits not dependent on its history. Here we consider the classes

  (4) Correlated: $\delta_2(t) = \eta\delta_1(t)$, with $\eta > 0$;

  (5) Anti-correlated: $\delta_2(t) = \eta\delta_1(t)$, with $\eta < 0$.

5.3 Population transfer in four-level systems

Our main goal is to classify the different classes of noises with ML, using the efficiency of population transfer, from $|0\rangle$ to $|1\rangle$, as a signature of the noise itself. The population transfer is achieved with a STIRAP-like protocol. In the next subsection we will show how to apply it in the two ultrastrongly coupled qubits system.

Stimulated Raman Adiabatic Passage [100, 101] (STIRAP) is a protocol that enables robust and efficient population transfer in a three-level system, by adiabatically following its dark state. In order to keep the notation consistent with the four-level system under consideration in this chapter, see eqs. (5.2) and Fig. 5.2.2. Let's consider a three-level ladder system composed by the states $\{|0\rangle, |2\rangle, |1\rangle\}$ and population transfer from state $|0\rangle$ to state $|1\rangle$. Two time-dependent classical fields drive the system: the *pump pulse* $\Omega_{\rm p}(t)$ drives the transition $|0\rangle - |2\rangle$, while the *Stokes pulse* drives the transition $|2\rangle - |1\rangle$. The Hamiltonian in an appropriate rotating frame reads [56]

$$H_{\rm S}(t) = \Delta|1\rangle\langle 1| + \Delta_p|2\rangle\langle 2| + \frac{1}{2}\left(\Omega_{\rm p}(t)|0\rangle\langle 2| + \Omega_{\rm s}(t)|2\rangle\langle 1| + \text{h.c.}\right), \tag{5.5}$$

where $\Delta_p = \varepsilon + \frac{g}{2} - \omega_{\rm p}$ and $\Delta = 2\varepsilon - \omega_{\rm p} - \omega_{\rm s}$ are the single and two-photon detunings, respectively, and $\omega_{\rm p}$ and $\omega_{\rm s}$ are the frequencies of pump and Stokes fields, respectively. For more details on the population transfer via STIRAP/CTAP protocol, see Subsec. 4.3.1 and Appendix 4.B.

It is well established in the literature [101] that STIRAP is robust against variations in the pulse shapes. Among the many possible choices [104],

in this work, we consider only Gaussian pulses, given by

$$\Omega_{\mathrm{p}}(t) = \Omega_{\mathrm{p}}^{\max} e^{-\left(\frac{t-\tau}{T}\right)^2}, \quad \Omega_{\mathrm{s}}(t) = \Omega_{\mathrm{s}}^{\max} e^{-\left(\frac{t+\tau}{T}\right)^2}, \tag{5.6}$$

and evolve the system over the time interval $[-5T, 5T]$, with $\tau = 0.7T$. Adiabaticity is ensured by the global condition [100]

$$\Omega_{\mathrm{p/s}}^{\max} \tau \geq 10. \tag{5.7}$$

The efficiency of the protocol is defined as the population of the target state $|1\rangle$ at the end of the evolution. The formal definition is given directly in Sec. 5.3.3 for the system of interest.

5.3.1 STIRAP in two ultrastrongly coupled qubits

The STIRAP protocol can be implemented in our system (Fig. 5.2.2) to transfer population from the ground state $|0\rangle = |gg\rangle$ to the double excited state $|1\rangle = |ee\rangle$ by symmetrically driving the two qubits. The control Hamiltonian is given by

$$H_{\mathrm{c}}(t) = \frac{\Omega(t)}{\sqrt{2}} \left(\sigma_1^x + \sigma_2^x\right), \tag{5.8}$$

where the $\sqrt{2}$ factor is introduced for convenience. It is important to note that, since the qubits are resonant, i.e., $\varepsilon_1 = \varepsilon_2 = \varepsilon$, and are symmetrically driven, the state $|3\rangle$ remains uncoupled from the others in the absence of noise. Consequently, the system effectively reduces to a three-level configuration with no direct transition between $|0\rangle$ and $|1\rangle$, forming a ladder-type structure. The full Hamiltonian (in absence of noise) is given by

$$\begin{aligned} H_{\mathrm{sys}} + H_{\mathrm{c}}(t) = 2\varepsilon|1\rangle\langle 1| + \left(\varepsilon + \frac{g}{2}\right)|2\rangle\langle 2| + \\ \left(\varepsilon - \frac{g}{2}\right)|3\rangle\langle 3| + \Omega(t)\left(|0\rangle\langle 2| + |2\rangle\langle 1| + \mathrm{h.c.}\right). \end{aligned} \tag{5.9}$$

In order to implement STIRAP in this system, we set $\Omega(t)$ as a two-tone field where one tone is resonant with the $|0\rangle - |2\rangle$ transition and one with the $|2\rangle - |1\rangle$ transition

$$\Omega(t) = \Omega_{\mathrm{s}}(t)\cos\omega_{\mathrm{s}} t + \Omega_{\mathrm{p}}(t)\cos\omega_{\mathrm{p}} t, \tag{5.10}$$

where $\Omega_{p/s}(t)$ are slowly varying pulse envelopes, $\omega_p = \varepsilon + \frac{g}{2}$ and $\omega_s = \varepsilon - \frac{g}{2}$.

In a convenient rotating frame, applying the Rotating Wave Approximation (RWA) and considering a full resonance scenario, the Hamiltonian simplifies in

$$H_{RWA}(t) = \frac{\Omega_p(t)}{2}|0\rangle\langle 2| + \frac{\Omega_s(t)}{2}|1\rangle\langle 2| + \text{h.c.}. \tag{5.11}$$

Oscillating terms with frequency equal to g were also neglected, due to the RWA. This is possible only within the ultrastrong coupling regime ($g \sim \varepsilon$). More details about the derivation and the approximations can be found in Appendix 5.A.

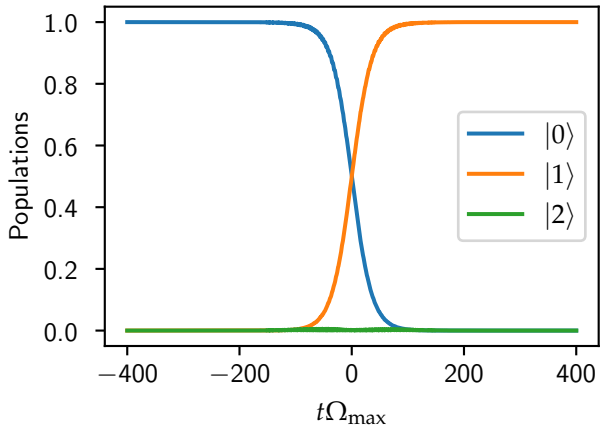

Figure 5.3.1: Population dynamics of the 2-qubits system governed by the full Hamiltonian eq. (5.9). The pulses are given by eq. (5.6) and the parameters are $\Omega_{p/s}^{max} = \Omega_{max} = 0.05\varepsilon$, $T\varepsilon = 2000$, $\tau = 0.7T$, $g = 0.5\varepsilon$. With these parameters the RWA holds, see Appendix 5.A, so the evolution closely resemble the STIRAP evolution described by eq. (5.11).

Eq. (5.11) is the Hamiltonian of a three-level ladder system with two time dependent drivings $\Omega_p(t)$ and $\Omega_s(t)$. If they are counterintuitively ordered and fulfill the adiabatic conditions, see eqs.(5.6) and (5.7), then STIRAP can effectively be applied to transfer population from state $|0\rangle$ to

$|1\rangle$. Fig. 5.3.1 reports the evolution of the 2-qubits system with the Hamiltonian given in eq. (5.9): It is apparent that the population is transfered from state $|0\rangle$ to the target state $|1\rangle$ while populating only negligibly state $|2\rangle$, as expected by STIRAP.

5.3.2 Effect of noise

In the previous section we have seen that under the condition of symmetric driving eq. (5.8) and qubit resonance ($\varepsilon_1 = \varepsilon_2$), state $|3\rangle$ is decoupled from the dynamics of the other states. This allows us to choose the Rabi frequency $\Omega(t)$ as in eq. (5.10) such that STIRAP is performed on the remaining 3 states. The presence of noise as described in Sec. 5.2 disrupts this configuration introducing a coupling between the states $|2\rangle$ and $|3\rangle$, and shifting the levels $|0\rangle$ and $|1\rangle$.

The full Hamiltonian describing the dynamics of the system with noise is the sum of the qubits Hamiltonian H_{sys}, the noise Hamiltonian $H_{\text{noise}}(t)$ and the control Hamiltonian $H_{\text{c}}(t)$

$$\begin{aligned} H_{\text{tot}} = \frac{g}{2}\left(|2\rangle\langle 2| - |3\rangle\langle 3|\right) + \left(\varepsilon + \frac{\delta_1(t) + \delta_2(t)}{2}\right)\left(|1\rangle\langle 1| - |0\rangle\langle 0|\right) \\ + \frac{\delta_1(t) - \delta_2(t)}{2}\left(|3\rangle\langle 2| + |2\rangle\langle 3|\right) + \Omega(t)\left(|0\rangle\langle 2| + |2\rangle\langle 1| + \text{h.c.}\right), \end{aligned} \tag{5.12}$$

where the noise-induced coupling between states $|2\rangle$ and $|3\rangle$, along with the energy shifts of $|0\rangle$ and $|1\rangle$, is clearly evident.

Notably, even with small noise $\delta_1, \delta_2 \ll \varepsilon$, the state $|3\rangle$ can still become populated during the evolution (see Fig. 5.3.2). This occurs because the eigensystem of the Hamiltonian is altered, eliminating the dark state that was present under ideal conditions.

5.3.3 Efficiency

The efficiency of the transfer protocol is defined as the population on the target state $|1\rangle$, that is

$$\xi = \lim_{N\to\infty} \frac{1}{N} \sum_{r=1}^{N} \xi^{(r)} \tag{5.13}$$

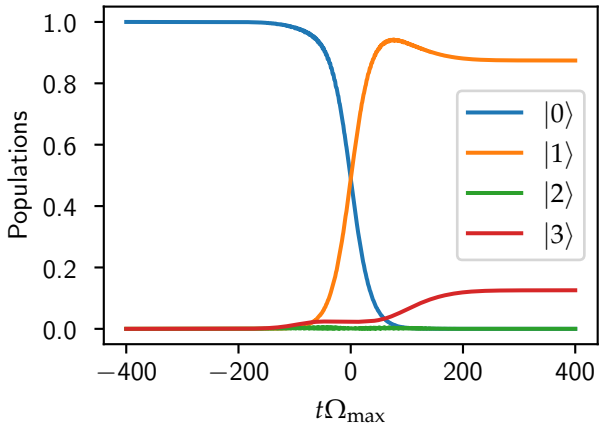

Figure 5.3.2: Population dynamics of the system, with $\delta_1 = 0.002\varepsilon$, $\delta_2 = -0.002\varepsilon$, $\omega_p = \varepsilon + g/2$, $\omega_s = \varepsilon - g/2$. The other parameters are the same as Fig. 5.3.1

where $\xi^{(r)} = \langle 1|\rho_{\mathrm{f}}^{(r)}|1\rangle$ and $\rho_{\mathrm{f}}^{(r)}$ if the density matrix of the system at the final time t_{f} for the r-th noise realization.

Since the quasistatic noise remains constant along any individual trajectory, the index r can be uniquely associated with the two real values $\delta_1^{(r)}$ and $\delta_2^{(r)}$, which are the specific realizations of the random processes δ_1 and δ_2, respectively. In other words, we can write $\xi^{(r)} = \xi(\delta_1^{(r)}, \delta_2^{(r)})$. Consequently, the efficiency defined in Eq. (5.13) is obtained by integrating over the joint probability distribution $p(\delta_1, \delta_2)$

$$\xi = \int \mathrm{d}\delta_1 \, d\delta_2 \, \xi(\delta_1, \delta_2) \, p(\delta_1, \delta_2). \tag{5.14}$$

For correlated and anti-correlated noise quasistatic noise, where $\delta_2^{(r)} = \eta\delta_1^{(r)}$, the above equation can be simplified in

$$\xi = \int d\delta_1 \, \xi(\delta_1, \eta\delta_1) \, p(\delta_1). \tag{5.15}$$

For the case of Markovian noise, the efficiency is obtained by solving a master equation in the Lindblad form (for more details see Appendix 5.B)

$$\langle \dot{\rho}(t) \rangle = -i \, [H(t), \rho(t)] - \frac{\gamma}{2} \left(O^2 \rho(t) - 2O\rho(t)O + \rho(t)O^2(t) \right), \tag{5.16}$$

where O is defined as

$$O = \left(\frac{1+\eta}{\sqrt{|\eta|}}\right)(|1\rangle\langle 1| - |0\rangle\langle 0|) + \left(\frac{1-\eta}{\sqrt{|\eta|}}\right)(|2\rangle\langle 3| - |3\rangle\langle 2|)\,. \tag{5.17}$$

Fig. 5.3.3 reports the stability plot of the efficiency of the protocol as a function of constant energy shifts, i.e., with δ_1 and δ_2 constant during the evolution.

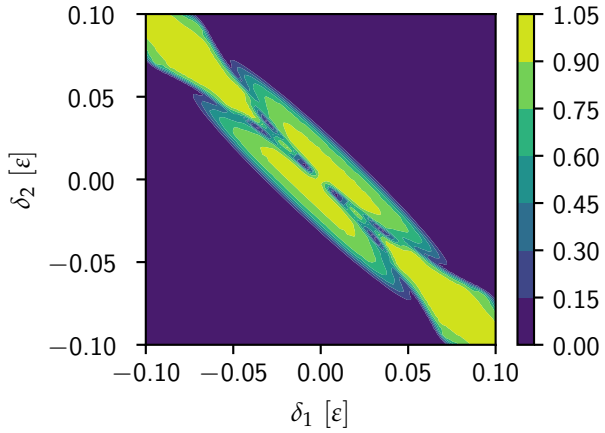

Figure 5.3.3: Stability plot for the efficiency of the STIRAP-like protocol versus the noises δ_1 and δ_2. The parameters are the same of Fig. 5.3.2.

We emphasize that our goal is not to achieve efficient population transfer but rather to leverage the protocol's sensitivity to extract information about the noise affecting the system, which reduces its efficiency.

5.4 Machine Learning model and Data generation

5.4.1 ML model

We employ a neural network (NN), namely a Multi Layer Perceptron (MLP), to perform supervised learning and classify the type of noise affecting the two-qubit system. A NN is a parameterized function capable

of approximating a broad class of target functions, the specific architecture we use is reported in Table 5.4.1. Further details about NNs are provided in Sec. 3.2.

Layer	**# Neurons**	**Activation Function**
Input	3	
Hidden 1	128	LeakyReLU
Hidden 2	128	LeakyReLU
Output	5	Softmax

Table 5.4.1: Layout of the neural network for classifying the types of noise.

Supervised Learning is performed by fitting a function approximator (in this case the NN) to a labeled dataset (see Sec. 3.1 for more details about supervised learning). This process is called *training* of the neural network [49, 99]. The dataset contains pairs $(\boldsymbol{x}_i, \hat{\boldsymbol{y}}_i)_{i=1,2,\dots,N}$, where each input $\boldsymbol{x}_i$ has an associated output $\hat{\boldsymbol{y}}_i$, also called label.

In this case the label $\hat{\boldsymbol{y}}_i$ for each sample i is an integer $a = 0, 1, 2, 3, 4$ representing one of the noise classes described in Sec. 5.2.1. The output $\boldsymbol{y}_i$ of the NN is a 5 dimensional Real vector whose components $y_{i,j}$ represent the probabilities that the noise affecting the system of sample i is of class j. The use of the *Softmax* activation in the output layer (see Tab. 5.4.1) ensures that the model's outputs form a normalized probability distribution. We minimize the *sparse categorical crossentropy* cost function

$$C = -\frac{1}{N}\sum_{i=1}^{N} \log\left(y_{i,\hat{y}_i}\right), \tag{5.18}$$

which is well-suited for multi-class classification tasks. This choice encourages the model to assign high probability to the correct label. The training process is monitored through the classification accuracy, defined as

$$A = \frac{1}{N}\sum_{i=1}^{N} \delta\big(\arg\max_j y_{ij},\ \arg\max_j \hat{y}_{ij}\big), \tag{5.19}$$

where δ is the Kronecker delta, and $\arg\max_j$ returns the index j of the largest component of the corresponding vector.

5.4.2 Data generation

Following the proposal in Ref. [131], we use as input the STIRAP efficiencies measured under three different driving conditions:

(i) $\Omega_{\mathrm{p}}^{\max} = \Omega_{\mathrm{s}}^{\max}$,

(ii) $\Omega_{\mathrm{p}}^{\max} = 2\Omega_{\mathrm{s}}^{\max}$,

(iii) $\Omega_{\mathrm{p}}^{\max} = \Omega_{\mathrm{s}}^{\max}/2$,

with the constraint $(\Omega_{\mathrm{p}}^{\max})^2 + (\Omega_{\mathrm{s}}^{\max})^2$ constant. This approach is advantageous as it provides high sensitivity to noise while requiring low experimental effort and is straightforward to implement in most physical systems. In fact the measurement has to be performed only at the end of (and not during) the evolution. This eliminates the necessity for time series data and may drastically reduce the duration of the experiment.

To generate the dataset, we numerically simulate the dynamics of the two-qubits system and evaluate the efficiency as reported in Sec. 5.3.3. Each input data point is then a three dimensional real vector
$\boldsymbol{x} = (\xi_{\Omega_{\mathrm{p}}=\Omega_{\mathrm{s}}}, \xi_{\Omega_{\mathrm{p}}>\Omega_{\mathrm{s}}}, \xi_{\Omega_{\mathrm{p}}<\Omega_{\mathrm{s}}})$.

For correlated (anti-correlated) non-Markovian noise we generate each data sample by randomly selecting the correlation parameter η within the range $[0.1, 5]$ ($[-5, -0.1]$). For each choice of η, we numerically evaluate the efficiency, eq. (5.15), under the three pulse conditions outlined earlier, yielding a single sample.

For uncorrelated non-Markovian noise instead, as the two random variables δ_1 and δ_2 are independent, we vary the standard deviation of the Gaussian distributions $p_1(\delta_1)$ and $p_2(\delta_2)$ respectively within the range $\sigma_1, \sigma_2 \in [\frac{\sigma}{5}, 5\sigma]$, with $\sigma = 10^{-2}\epsilon$ before calculating every data sample. We then numerically calculate the efficiencies under the three pulse condition using eq. (5.14).

In the context of Markovian, correlated (anti-correlated) noise, the efficiencies are calculated numerically solving the Lindblad master equation,

eq. (5.16). In this particular instance, not only does η fluctuate within the specified ranges like correlated (anti-correlated) non-Markovian scenarios from one sample to another, but also we randomly select the decay parameter $\gamma \in [10^{-4}, 10^{-3}]\epsilon$. For each noise class, we generate 500 samples. The total dataset is split into a $3:1:1$ ratio for training, validation, and test set respectively.

5.5 Results

We train the NN model described in Tab. 5.4.1. The output of the neural network for a given input vector $\boldsymbol{x}_i$, denoted as $\boldsymbol{y}_i$, is compared against the corresponding true label $\hat{\boldsymbol{y}}_i$. The training progress of the model is illustrated in Fig. 5.5.1, where panel (a) shows the classification accuracy A, eq. (5.19). The accuracy is observed to increase as the number of

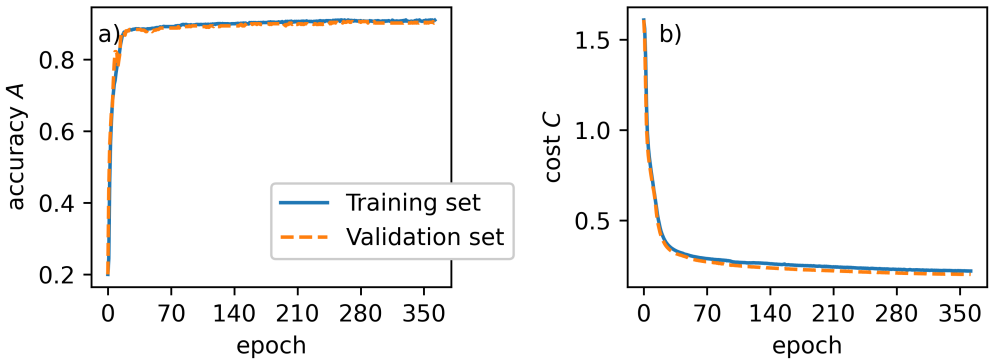

Figure 5.5.1: (a) Accuracy, and (b) value of the cost function, eq. (5.18), for the training (solid blue line) and validation (solid orange line) sets versus the number of epochs of training. The accuracy on the test set is $A \approx 90\%$.

training epochs progresses. Panel (b) shows that the cost function C, eq. (5.18), decreases accordingly. After 56 training epochs, the model achieves a classification accuracy of approximately 90%. However, this accuracy may vary slightly (around 1%) due to the random initialization

of model parameters, as well as the random shuffling and splitting of the dataset.

The classification performance of the model is further analyzed using the confusion matrix in Fig. 5.5.2, which compares the true noise classes with the predicted ones. In the matrix, each row corresponds to a true noise class, while each column represents the predicted class. The five diagonal elements indicate the percentage of correct predictions for each noise class. The model achieves an accuracy of 99.4% in distinguishing be-

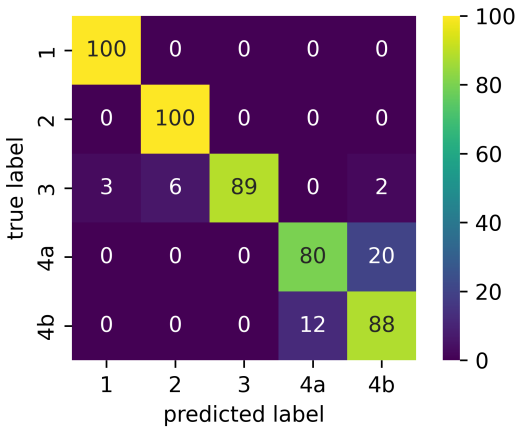

Figure 5.5.2: . Confusion matrix of the MLP model for classifying the noise types described in Sec. 5.2.1. Each row represents the true noise class, while each column corresponds to the predicted class. It is visible that all the 5 classes of noise can be classified with an accuracy up to $\approx 90\%$.

tween non-Markovian and Markovian noise. Within the non-Markovian noise class, it correctly classifies correlations with an accuracy of 95%, whereas within the Markovian class, the classification accuracy is 82.5%. This is in contrast from the three-level system case analyzed in Ref. [131], where the model successfully distinguished non-Markovian noise classes and also differentiated non-Markovian from Markovian noises, however it was unable to discriminate between the two distinct Markovian noise types (correlated vs anti-correlated). This limitation is overcome in the two-qubits system explored here, albeit at the cost of a moderate reduc-

tion in overall classification accuracy.

The enhanced classification capability observed in this four-level scenario can be attributed to the noise-induced mixing effect of the intermediate states $|2\rangle$ and $|3\rangle$ (see Sec. 5.3.2 and eq. (5.12)). Specifically, the noise not only shifts the states $|0\rangle$ and $|1\rangle$, but also mixes the states $|2\rangle$ and $|3\rangle$ causing them to no longer be eigenstates of the system Hamiltonian eq. (5.1). As a result, the external pulses simultaneously drive both states. The mixing depends to the relative sign of the noises, leading to different responses to correlated and anti-correlated noise types.

The ability to classify the non-Markovian noises can be further clarified by examining the plot of the efficiency versus the constant energy shifts δ_1 and δ_2 for the three driving conditions, as reported in Fig. 5.5.3. The effect of non-Markovian quasistatic noise is to shift the energies of

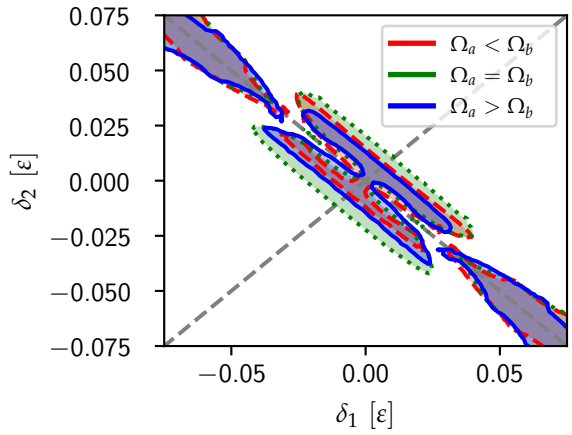

Figure 5.5.3: Contour plot for the efficiency of the STIRAP-like protocol, with different ratios between the pulses. The parameters are the same of Fig. 5.3.3. Each line refers to a threshold of 0.85 for the efficiency of the population transfer.

the qubits at each individual repetition of the protocol. Consequently, for (anti-)correlated quasistatic non-Markovian noise, the overall efficiency is obtained as the weighted integral over the line $\delta_2 = \eta\delta_1$, see eq. (5.15). Since the efficiency profile varies with the driving condition, the resulting efficiency also depends on the driving condition, and this dependence is

different for $\eta > 0$ or $\eta < 0$. Although in the three-level system scenario this dependence is understood, in the present case the effect of the noise correlation's sign is not immediately evident; nonetheless, ML techniques provides a robust solution.

The situation becomes even more complex for uncorrelated non-Markovian quasistatic noise, where the weighted integral is over the whole $\delta_1 - \delta_2$ plane (see eq. (5.14)), or for the Markovian noise where the efficiency for each trajectory cannot be represented as a single point in $\delta_1 - \delta_2$ plane, since the noise assumes all possible values during each repetition of the protocol.

5.6 Conclusions

In this work, we have proposed and numerically validated a machine-learning protocol to categorize different classical noise types affecting two ultrastrongly coupled qubits, based on their spatial correlation and temporal character (Markovianity). We defined and successfully classified five distinct noise classes: three non-Markovian quasistatic types (correlated, anti-correlated, and uncorrelated) and two Markovian types (correlated and anti-correlated). Our approach leverages the sensitivity of a STIRAP-like population transfer protocol, where variations in the population transfer efficiency serve as distinctive fingerprints of the underlying noise.

By measuring only the final population transfer efficiencies under three distinct driving conditions, we achieved over 90% classification accuracy among the five considered noise classes, as well as nearly perfect discrimination between Markovian and non-Markovian noise. Remarkably, this approach requires only three inputs to a shallow neural network, eliminating the need for time-series data or real-time monitoring. Furthermore, the protocol we propose consists of symmetric driving of both qubits, thus having a minimal experimental overhead.

Compared to our earlier study on three-level quantum networks [131], where ML effectively detected non-Markovian noise correlations and distinguished between Markovian and non-Markovian classes, in the present two-qubit setup we successfully classified different correlation signs even

in the Markovian noise class. This enhanced sensitivity arises from the richer structure of the four-level system, where noise-induced mixing between the two intermediate states $|ge\rangle \pm |eg\rangle$ imprints a distinct signature on the transfer efficiency that depends on the sign of the noise correlations.

Future work include identifying alternative input features in order to broaden the scope of the approach to include the classification of additional noise types and the application of unsupervised learning. Incorporating realistic imperfections such as pulse distortions and measurement noise will help assess the robustness of the method in practical settings. Additionally, exploring alternative coherent-control schemes and ML architectures may further enhance classification sensitivity and broaden applicability to diverse quantum platforms. We believe that the synergy between control protocols and ML approaches will play a key role in the next generation of quantum-hardware diagnostics.

Appendix

5.A Rotating frame, validity of RWA and efficiency of STIRAP with detunings

The unitary transformation used to obtain eq. (5.11) is given by the operator

$$U_x(t) = e^{-i\left(\omega_s+\omega_p\right)t}|1\rangle\langle 1| + e^{-i\omega_p t}|2\rangle\langle 2|. \tag{5.20}$$

Here, for the sake of completeness, we will allow for small detunings in the drivings, whose frequencies are defined as

$$\omega_p = \varepsilon + \frac{g}{2} - \delta_p, \tag{5.21a}$$

$$\omega_s = \varepsilon - \frac{g}{2} - \delta_s. \tag{5.21b}$$

The transformed Hamiltonian in the rotating frame is derived using the expression

$$\tilde{H}(t) = U_x(t)^\dagger H(t) U_x(t) - iU_x(t)^\dagger \partial_t U_x(t), \tag{5.22}$$

which, when applied to the Hamiltonian in eq. (5.9), yields

$$\begin{aligned}\tilde{H}_1(t) = &\left(\delta_p + \delta_s\right)|1\rangle\langle 1| + \left(\delta_p\right)|2\rangle\langle 2| + \left(\varepsilon + \frac{g}{2}\right)|3\rangle\langle 3| + \\ &\left[\frac{\Omega_p}{2}\left(e^{-2i\omega_p t} + 1\right) + \frac{\Omega_s}{2}\left(e^{i\left(\omega_s-\omega_p\right)t} + e^{-i\left(\omega_s+\omega_p\right)t}\right)\right]|0\rangle\langle 2| + \\ &\left[\frac{\Omega_s}{2}\left(e^{-2i\omega_s t} + 1\right) + \frac{\Omega_p}{2}\left(e^{-i\left(\omega_s-\omega_p\right)t} + e^{-i\left(\omega_s+\omega_p\right)t}\right)\right]|2\rangle\langle 1| + h.c.\end{aligned} \tag{5.23}$$

Since $2\omega_s, 2\omega_p, \omega_s + \omega_p \gg \omega_s, \omega_p$ we can neglect the respective oscillating terms[133], obtaining

$$\tilde{H}_1(t) = \left(\delta_p\right)|2\rangle\langle 2| + \left(\delta_p + \delta_s\right)|1\rangle\langle 1| + \left(\varepsilon + \frac{g}{2}\right)|3\rangle\langle 3| + \left[\frac{\Omega_p}{2} + \frac{\Omega_s}{2}e^{i\left(\delta_s - \delta_p - g\right)t}\right]|0\rangle\langle 2| + \left[\frac{\Omega_s}{2} + \frac{\Omega_p}{2}e^{-i\left(\delta_s - \delta_p - g\right)t}\right]|2\rangle\langle 1| + h.c., \tag{5.24}$$

Within the full resonance, the remaining terms oscillate at a frequency equal to g. This means that in the ultrastrong coupling regime ($g \sim \varepsilon$), even the last oscillating term can be neglected under the RWA. The Hamiltonian simplifies in

$$H_{\mathrm{RWA}} = \left\{\frac{\Omega_p}{2}|0\rangle\langle 2| + \frac{\Omega_s}{2}|1\rangle\langle 2| + h.c.\right\}, \tag{5.25}$$

as shown in the main text.

We want to briefly survey the validity of the RWA, comparing the efficiency and adiabaticity of the STIRAP protocol, with Gaussian pulses, for the full system, without any RWA, for different values of pulses amplitude. From eq.(5.7) the adiabaticity condition has to be valid. Unfortunately the value of $\Omega_{\mathrm{p/s}}^{\max}$ is also constrained from above, because the RWA that led to eq. (5.25) requires $\Omega_{\mathrm{p/s}}^{\max} \ll \omega_s, \omega_p$. We can see this behaviour in the Fig. 5.A.1, where the inefficiency of the protocol, defined as

$$\mathcal{I} = 1 - \langle 1|\rho_f|1\rangle, \tag{5.26}$$

whith ρ_f the final state, is plotted against the value $\Omega_{\mathrm{p/s}}^{\max}\tau$.

We also want to account the fact that, during the dynamic, $\langle 2|\rho(t)|2\rangle \neq 0$. As an ulterior measure of the adiabaticity of the protocol, we define

$$A_N = \frac{1}{t_f - t_i}\int_{t_i}^{t_f}\langle 2|\rho(t)|2\rangle dt \tag{5.27}$$

In Fig. 5.A.1 is shown how A_N is affected by $\Omega_{\mathrm{p/s}}^{\max}\tau$. This trade-off led to set $\Omega_{\mathrm{p/s}}^{\max} = 0.05\varepsilon$ for all the future simulations. Within the considerations presented above, the STIRAP protocol, is clearly very effective. In order

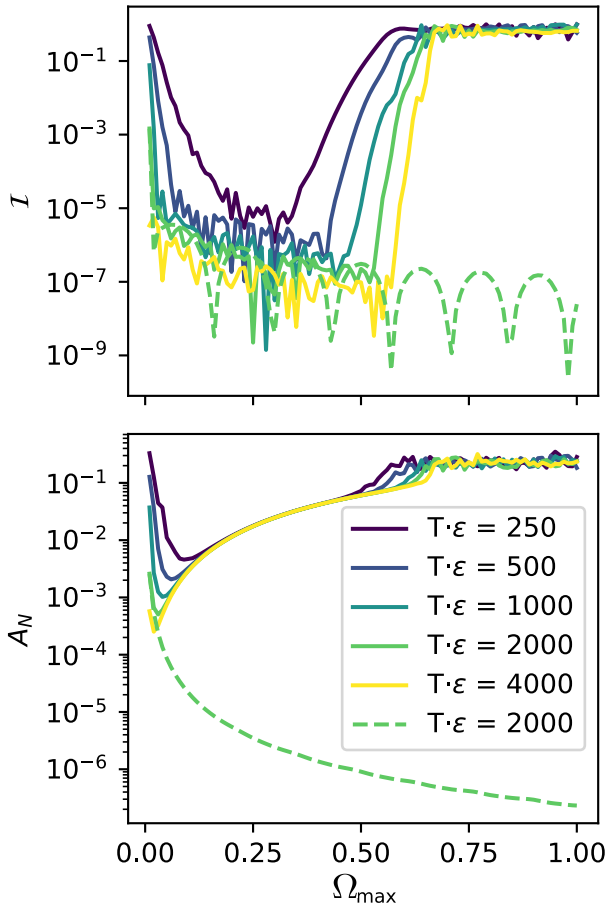

Figure 5.A.1: (Upper image) Inefficiency versus $\Omega_{\max}$ in units of ϵ for different values of T. (Lower image) A_N versus $\Omega_{\max}$ with different values of T. The other parameters are $T/\varepsilon = 2000, \tau = 0.7T$, $g = 0.5\varepsilon, \delta_a = -g/2, \delta_b = g/2$. The dashed line depicts the behavior of the system with the full RWA eq. (5.11). We can clearly see that by increasing the adiabaticity, the STIRAP protocol becomes more efficient.

to assess its validity outside the full resonance scenario, we resort to numerical simulations. Nonetheless, in Fig. 5.A.3 we show a full stability plot of the final population of the state $|1\rangle$. The leftmost cluster is linked to the full resonance scenario, as in eq. (5.25). The rightmost feature is

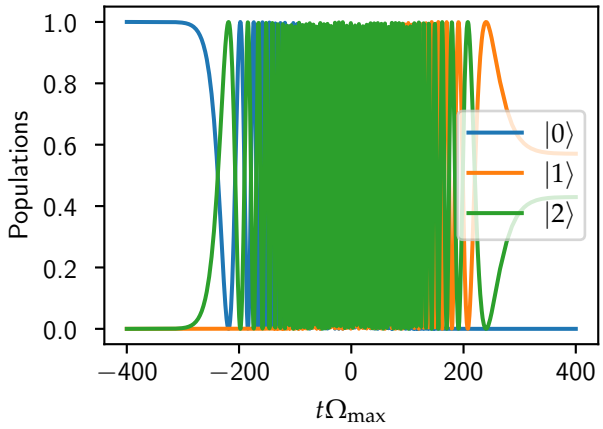

Figure 5.A.2: Population dynamics of the effective 3-level system described by eq. (5.9). The dynamic was simulated with the following parameters: $\Omega_{p,s} = (0.05\varepsilon)e^{-\left(\frac{t\mp\tau}{T}\right)^2}$, $T/\varepsilon = 2000, \tau = 0.7T$, $g = 0.5\varepsilon, \delta_p = g, \delta_s = -g$.

linked again to a full resonance scenario, but with an "intuitive" sequence of the pulses. Despite having a high efficiency population transfer, the intermediate state will be macroscopically populated during the protocol (see Fig. 5.A.4).

5.B Derivation of Lindblad master equation

In this section we outline the derivation of the Master Equation for Markovian (anti)correlated noise. The system Hamiltonian is defined by the sum of eqs. (5.1) and (5.8), that is $H_s(t) = H_{\text{sys}} + H_c(t)$ While the effect of the noise is described by eq. (5.3), with $\delta_1 = \frac{1}{\sqrt{\eta}}\chi$ and $\delta_2 = \sqrt{\eta}\chi$, obtaining:

$$H_n(t) = \frac{\chi(t)}{2}O, \tag{5.28}$$

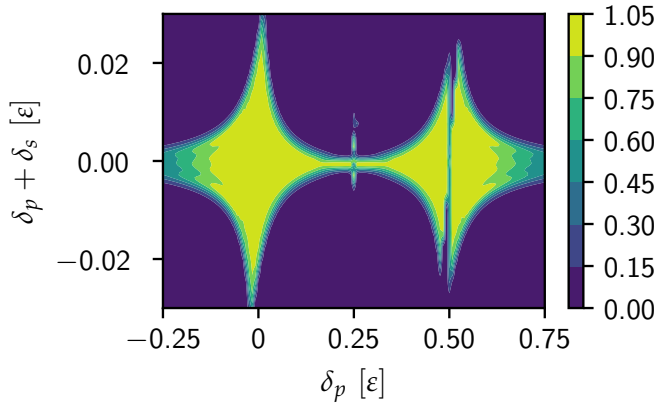

Figure 5.A.3: Stability plot for the efficiency of the STIRAP-like protocol, the parameters are the same of Fig. 5.3.1, δ_p and δ_s are in units of ε. The left high-efficiency region is where standard STIRAP occurs.

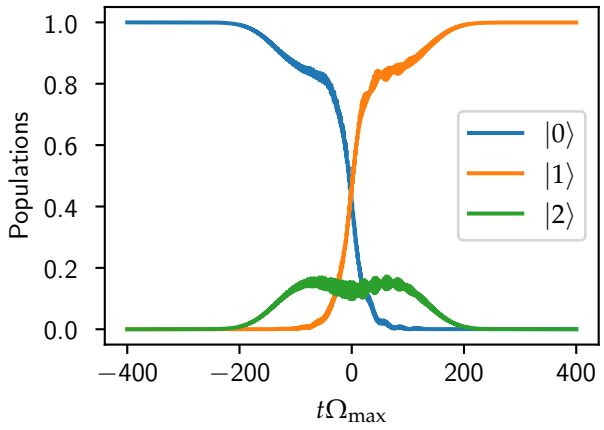

Figure 5.A.4: Population dynamics of the effective 3-level system described by eq. (5.9). The dynamic was simulated with the following parameters: $\Omega_{p,s} = (0.05\varepsilon)e^{-\left(\frac{t\mp\tau}{T}\right)^2}$, $T/\varepsilon = 2000, \tau = 0.7T$, $g = 0.5\varepsilon, \delta_a = \frac{g}{2} + \Omega_{max}, \delta_b = -\frac{g}{2} - \Omega_{max}$.

with

$$O = \left(\frac{1+\eta}{\sqrt{|\eta|}}\right)(|1\rangle\langle 1| - |0\rangle\langle 0|) + \left(\frac{1-\eta}{\sqrt{|\eta|}}\right)(|2\rangle\langle 3| - |3\rangle\langle 2|)\,. \tag{5.29}$$

Being zero-mean Markovian noise, we clearly have $\langle\chi(t)\rangle = 0$, $\langle\chi(t)\chi(t')\rangle = \gamma\delta(t-t')$. We derive the master eqation in the *instantaneous interaction picture* defined by:

$$\tilde{\rho}(t) = e^{iH_s t}\rho(t)e^{-iH_s t}, \tag{5.30}$$

from which we derive

$$\dot{\tilde{\rho}}(t) \approx -i\left[\tilde{H}_n(t), \tilde{\rho}(t)\right]. \tag{5.31}$$

Where we have neglected the time dependency of H_s, due to its adiabaticity. By formally integrating eq. (5.31), substituting it into the commutator and taking the stochastic average at time t, we get:

$$\langle\dot{\tilde{\rho}}(t)\rangle = -\left\langle \int_0^t \left[\chi(t)\tilde{O}(t), \left[\chi(t')\tilde{O}(t'), \tilde{\rho}(t')\right]\right] dt' \right\rangle. \tag{5.32}$$

We can make the substitution $\langle\chi(t)\chi(t')\tilde{\rho}(t')\rangle = \langle\chi(t)\chi(t')\rangle\,\tilde{\rho}(t')$, since $t' < t$, and $\tilde{\rho}(t')$ cannot depend on t, except for $t' = t$, obtaining

$$\langle\dot{\tilde{\rho}}(t)\rangle = -\frac{\gamma}{2}\left(\tilde{O}^2\tilde{\rho}(t) - 2\tilde{O}\tilde{\rho}(t)\tilde{O} + \tilde{\rho}(t)\tilde{O}^2(t)\right). \tag{5.33}$$

The above equation can be recast into the Schroedinger's picture giving

$$\langle\dot{\rho}(t)\rangle = -i\left[H_s(t), \rho(t)\right] - \frac{\gamma}{2}\left(O^2\rho(t) - 2O\rho(t)O + \rho(t)O^2(t)\right) \tag{5.34}$$

Chapter 6

Characterization of noise Gaussianity in two level-quantum systems

6.1 Introduction

As discussed in the former chapters , environmental noise remains a major challenge for real-world quantum platforms such as superconducting qubits [39, 134], trapped ions [135, 136], and semiconductor quantum dots [137, 138]. In particular, the noise encountered in these systems is often non-Markovian and non-Gaussian, meaning that it exhibits memory effects and statistical properties that deviate from simple Gaussian behavior, thus it cannot fully be characterized with second-order statistics and higher-order correlations have to be taken into account [34, 139, 140]. These features significantly complicate the theoretical modeling and control of open quantum systems, as standard approximations and master equation techniques often fail in such regimes [141, 142].

Accurate and efficient noise characterization is therefore essential to enhance the fidelity of quantum gates, designing robust quantum control protocols and error correction schemes. A precise understanding of noise allows one to tailor mitigation strategies to the specific characteristics of the environment, potentially extending coherence times and improving

system performance. In this chapter, we focus on the notion of Gaussianity, a fundamental property os stochastic processes. We analyze noise processes that range from fully non-Gaussian (e.g., Random Telegraph Noise (RTN)) to Gaussian (e.g., Ornstein-Uhlenbeck (OU) noise), as well as hybrid combinations of both. This investigation is not only of theoretical interest but also of practical importance, as many realistic noise environments exhibit a mixture of Gaussian and non-Gaussian traits. Traditional approaches to noise characterization include noise spectroscopy[143–145], repeated measurements on the quantum probe[146], and methods based on quantum estimation theory [28, 34]. While powerful, these methods often require intricate experimental setups, are sensitive to measurement errors, or rely on assumptions that may not hold in realistic non-Gaussian regimes.

Here, we adopt a different approach: we recast the problem in a data-driven framework using supervised machine learning (ML). By treating the quantum system itself as a probe and collecting time-series data of its dynamical observables we train neural networks to infer key features of the noise. This method leverages the predictive power of neural networks to learn complex, nonlinear mappings between observable dynamics and underlying noise parameters. Our ML-based framework is adaptable to different noise models and enables simultaneous multi-parameter estimation, allowing for comprehensive characterization of hybrid noise environments.

In sec. 6.2, we define our physical model, describing a single qubit interacting with a Gaussian and non-Gaussian environment, and introduce the parameters governing the noise. In sec. 6.3, we describe the protocol and detail the strategy for generating time-series data suitable for supervised learning. Finally in sec. 6.4 we present and analyze the performance of the model, and conclude in 6.5 with a discussion of improvements and future directions.

6.2 Physical model

We consider a qubit interacting with an environment composed by N bistable fluctuators. Each fluctuator is modeled as a classical RTN pro-

cess. RTN is an archetypical example of a non-Gaussian process [147–149].

The Hamiltonian of the total system is $H(t) = H_{\text{sys}} + H_{\text{noise}}(t)$, where

$$H_{\text{sys}} = -\frac{\epsilon}{2}\sigma_z - \frac{\Delta}{2}\sigma_x, \tag{6.1}$$

$$H_{\text{noise}}(t) = -\sum_{n=1}^{N} \frac{v_n}{2} x^{(n)}(t)\sigma_z. \tag{6.2}$$

Here σ_x and σ_z are the Pauli matrices, ϵ is the qubit energy splitting and Δ is the coupling between the two levels. The noise is described by the random processes $x^{(n)}$, n referring to the n-th instance of RTN with coupling strength v_n. The random variable $x^{(n)}(t)$ toggles between two values -1 and $+1$, with rates $\Gamma_n^{\pm}$ (see 2.5.1).

For simplicity, we consider only pure dephasing, i.e. $\Delta = 0$, resulting in fluctuations of the qubit's energy levels. This causes the coherence to decay while the populations remain unchanged. In the interaction picture with respect to the qubit Hamiltonian H_{sys}, the temporal evolution of the qubit coherence influenced by the N Random Telegraph Noises is given by [147, 148]

$$\rho_{01}(t)_{\text{RTN}} = \rho_{01}(0) \prod_{n=1}^{N} \left[A_n e^{-\Gamma_n(1-\alpha_n)t/2} + (1 - A_n)e^{-\Gamma_n(1+\alpha_n)t/2}\right], \tag{6.3}$$

where

$$\alpha_n = \sqrt{1 - g_n^2 - 2\mathrm{i}g_n\bar{\delta p}_n}, \tag{6.4a}$$

$$A_n = (1 + \alpha_n - \mathrm{i}\delta p^0 g_n)/(2\alpha_n), \tag{6.4b}$$

$$g_n = 2v_n/\Gamma_n, \tag{6.4c}$$

$$\Gamma_n = \Gamma_n^+ + \Gamma_n^-, \tag{6.4d}$$

$$\bar{\delta p}_n = (\Gamma_n^- - \Gamma_n^+)/\Gamma_n, \tag{6.4e}$$

$$\delta p^0 = \rho_{00}(0) - \rho_{11}(0). \tag{6.4f}$$

For a single RTN with $\Gamma^+ = \Gamma_- = \Gamma/2$ and $\delta p^0 = 0$, the two-time correlation function is

$$\langle x(t_1)x(t_2)\rangle = e^{-\Gamma|t_1 - t_2|}. \tag{6.5}$$

In the limit of many ($N \to \infty$) identical ($g_n = g$, $\Gamma_n = \Gamma$) illustrates weakly coupled ($g \to 0$) RTNs with $Ng^2 = C^2$ constant and in the stationary state ($\delta p^0 = \bar{\delta p}_n = 0$), the collective noise effect becomes equivalent to that of a single OU process (see 2.4.1), which is an archetypical Gaussian process. In this case the coherence of the qubit decays as

$$\rho_{01}(t)_{\mathrm{RTN}} = \rho_{01}(0) e^{-\frac{C\Gamma}{4}(t - \frac{1}{\Gamma} + \frac{1}{\Gamma} e^{-\Gamma t})}. \tag{6.6}$$

In fact, eq. (6.6) has the same form of the evolution of the coherence of a qubit subject to a single OU noise with zero mean $\langle x(t) \rangle = 0$ and two-time correlation function

$$\langle x(t_1) x(t_2) \rangle = e^{-k|t_1 - t_2|}, \tag{6.7}$$

where $k = 1/\tau_c$ is the inverse of the noise correlation time. The decay of the coherence of the qubit in this case is (see Appendix 6.A)

$$\rho_{01}(t)_{OU} = \rho_{01}(0) e^{-v^2 \tau_c (t - \tau_c + \tau_c e^{-\frac{t}{\tau_c}})} = \rho_{01}(0) e^{-\frac{v^2}{k}(t - \frac{1}{k} + \frac{1}{k} e^{-kt})}, \tag{6.8}$$

where v is the strength of the OU noise.

Our analysis explores the estimation of the noise properties in two different scenarios:

1. **Finite number of RTNs:** The qubit interacts with an environment composed of N independent, identical RTN sources, each characterized by a switching rate Γ and a coupling strength v. We aim to estimate the number N, which serves as an indicator of the degree of Gaussianity of the noise: as $N \to \infty$, the noise tends toward a Gaussian process. In addition, our objective is also to infer the common switching rate Γ, which governs the temporal correlations of the RTNs.

2. **Hybrid noise (Gaussian + non-Gaussian components):** The environment is composed of many ($N \to \infty$) weakly coupled RTNs and a single strongly coupled RTN. In the limit of many weakly coupled identical RTNs (with $g \to 0, Ng^2 = C^2$ constant), their collective effect becomes equivalent to an Ornstein-Uhlenbeck (OU) process—representing a Gaussian noise contribution while the strongly

coupled RTN introduces a non-Gaussian contribution. Our goal is to characterize both the Gaussian and non-Gaussian components of the total noise. Specifically, we aim to estimate the strengths of the OU (v_{OU}) and RTN (v_{RTN}), which indicate their respective weights, as well as estimate the RTN switching rate Γ and the OU inverse correlation time $k = \frac{1}{\tau_c}$.

6.3 Parameter estimation with quantum probe

We employ the qubit as a quantum probe to infer properties of its surrounding environment using supervised machine learning. In particular, we use a Multi Layer Perceptron (MLP) (see Sec. 3.2 for more details about MLP) trained on time-series data of the qubit's dynamics to estimate environmental noise parameters.

The qubit is initially prepared in an eigenstate of the Pauli σ_x operator

$$|\psi(0)\rangle = |+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \tag{6.9}$$

and allowed to evolve under the influence of the noisy environment. To generate the training data required for supervised learning, we consider the expectation value $\langle\sigma_x(t_i)\rangle$ at each time step t_i. These time-series trajectories form the input to the neural network, while the corresponding noise parameters serve as target outputs for training. The two scenarios introduced in Sec. 6.2 are analyzed in detail below.

6.3.1 Quantifying Gaussianity with the number N of RTNs

In this subsection, we formulate the problem introduced in Sec. 6.2, focusing on the first scenario. We consider a single qubit subject to noise arising from N independent, identical RTN sources. In this framework, the parameter N serves as a quantifier of non-Gaussianity: When $N = 1$, the noise is maximally non-Gaussian. As N increases, while keeping $C = \sqrt{N}g$ constant, the cumulative effect of the fluctuators gradually converges to that of a OU process which is Gaussian. This transition reflects the central limit theorem for classical noise sources, where many weakly coupled RTNs give rise to an effective Gaussian environment.

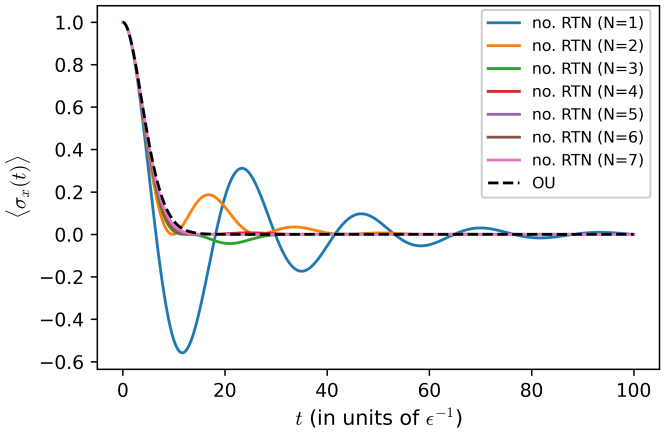

Figure 6.3.1: Time evolution of the expectation value $\langle \sigma_x(t) \rangle$ for a qubit interacting with different numbers of RTN sources. The time axis spans ($t \in [0, 100]\epsilon^{-1}$), and the total noise strength is kept constant at ($C^2 = Ng^2 = 30$). Colored lines indicate dynamics generated by RTN noise with increasing N, while the black dotted line corresponds to dynamics under OU noise with parameters $\tau_c = 1/\Gamma = 1$ and $v_{OU} = \frac{\sqrt{C}\Gamma}{2}$. For $N = 7$, the RTN-induced dynamics closely approximate the one induced by the Gaussian OU process.

To analyze the system dynamics under this type of noise, we use the analytical expression given in eq. (6.3). Fig. 6.3.1 illustrates examples of time evolutions of the qubit under the influence of N identical RTN sources, with N ranging from 1 to 7 and fixed $C^2 = 30$, along with the asymptotic dynamics reached in the limit $N \to \infty$. As already mentioned and proved in Appendix 6.B, the latter is equivalent to the dynamics of a qubit affected by a single OU process with parameters $k = \Gamma$ and noise strength $v_{OU} = \frac{\sqrt{C}\Gamma}{2}$. It is evident that with increasing N the trajectory of $\langle \sigma_x(t) \rangle$ converges to the one when the qubit is affected only by the OU process. Already for $N = 7$ that trajectory nearly overlaps with the asymptotic one, clearly demonstrating convergence to the Gaussian limit.

To characterize the noise using the qubit as a probe in a manner suitable for supervised learning, we simulate the the qubit evolution for varying numbers of RTNs, $N \in [1, 8]$, and for different noise switching rates,

$\Gamma \in [0.1, 2]\epsilon$. The time evolution is discretized into 1000 time steps over the interval $t \in [0, T]$ where $T = 100/\epsilon$. At each time step t_i, we extract the expectation value of the Pauli σ_x operator, $\langle \sigma_x(t_i) \rangle$, which serves as the input time-series data for the supervised learning model.

Each simulated qubit trajectory is labeled by the corresponding noise parameters (N, Γ) used in its generation. After training on this labeled dataset, the machine learning model can infer the values of N and Γ from previously unseen qubit dynamics affected by RTN noise, thereby enabling simultaneous estimation of the Gaussianity (via N) and temporal characteristics (via Γ) of the environment.

6.3.2 Quantifying Noise Parameters in a Hybrid Environment

In this section, we analyze the second scenario outlined in Sec. 6.2, where the environment consists of both Gaussian and non-Gaussian noise components. This can happen, for example, when one RTN is strongly coupled to the qubit accounting for non-Gaussian noise effect, while many other RTN sources are weakly coupled. In the limit $N \to \infty$ (with the constraint $Ng^2 = C^2$ constant), the collective influence of the weakly coupled RTNs converges to that of a Gaussian OU process, as described in Sec. 6.2. This configuration effectively realizes a hybrid noise environment where the non-Gaussian RTN and Gaussian OU noise coexist. Following eq. (6.3) we have

$$\begin{aligned}
\rho_{01}(t) &= \lim_{N\to\infty} \rho_{01}(0) \prod_{n=1}^{N} \left[A_n e^{-\Gamma_n(1-\alpha_n)t/2} + (1-A_n) e^{-\Gamma_n(1+\alpha_n)t/2} \right] = \\
&= \rho_{01}(0) \left(A_1 e^{-\Gamma_1(1-\alpha_1)t/2} + (1-A_1) e^{-\Gamma_1(1+\alpha_1)t/2} \right) \times \\
&\quad \times \lim_{N\to\infty} \prod_{n=2}^{N} \left[A_n e^{-\Gamma_n(1-\alpha_n)t/2} + (1-A_n) e^{-\Gamma_n(1+\alpha_n)t/2} \right] = \\
&= \frac{\rho_{01}(t)_{\mathrm{RTN}} \cdot \rho_{01}(t)_{\mathrm{OU}}}{\rho_{01}(0)},
\end{aligned} \tag{6.10}$$

where $\rho_{01}(t)_{\mathrm{RTN}}$ is the contribution from the strongly coupled RTN source, while $\rho_{01}(t)_{\mathrm{OU}}$ corresponds to the collective effect of the weakly coupled

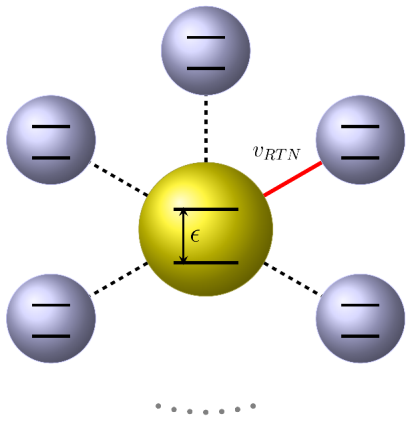

Figure 6.3.2: Schematic representation of the qubit (yellow) hybrid environment. One RTN source is strongly coupled to the qubit (shown with red line), while the remaining $N-1$ RTN sources are weakly coupled (shown with dashed line). In the limit $N \to \infty$ and $Ng^2 =$ constant, the weakly coupled RTNs collectively reproduce the effect of an Ornstein-Uhlenbeck (OU) noise process.

RTN sources.

We apply supervised learning to estimate the characteristic parameters of both noise components: the coupling strength v_{RTN} and total switching rate Γ of the RTN, as well as the strength v_{OU} and inverse correlation time $k = 1/\tau_c$ of the OU noise. Notice that the ratio between the strengths v_{RTN} and v_{OU} represents the relative strength of the non-Gaussianity of the noise.

To generate the training data, we simulate the evolution of the qubit by randomly sampling the noise parameters from the following interval:

- RTN strength: $v_{RTN} \in [0.1, 2]\,\epsilon$,
- RTN switching rate: $\Gamma \in [0.1, 2]\,\epsilon$,
- OU strength: $v_{OU} \in [0.1, 2]\,\epsilon$,
- OU inverse correlation time: $k = 1/\tau_c \in [0.1, 2]\,\epsilon$.

The parameter ranges for the RTN are chosen such that the coupling strength g spans the weak ($g < 1$), intermediate ($g = 1$), and strong

($g > 1$) coupling regimes. The same range is adopted for the OU noise to enable a consistent and meaningful comparison between the two noise types.

As in the previous scenario, the qubit evolution is computed over a time interval $t \in [0, T]$, $T = 100/\epsilon$, discretized into 1000 time steps. At each time step t_i, the expectation value $\langle \sigma_x(t_i) \rangle$ is recorded, forming the time-series input to the neural network. The four corresponding noise parameters are used as target outputs. Once trained, the neural network should be capable of estimating all four parameters from previously unseen qubit trajectories, thereby enabling a comprehensive characterization of the hybrid environment.

6.4 ML results

6.4.1 Scenario 1

We implemented the regression tasks using TensorFlow. For Scenario 1, a dataset comprising 4000 evolutions of the qubit was generated. For each integer value of $N \in [1, 8]$, we randomly sampled the switching rate $\Gamma \in [0.1, 2]\,\epsilon$ a total of 500 times. The complete dataset was divided into training, validation, and test sets using a $3 : 1 : 1$ ratio.

Layer	# Neurons	Activation Function
Input	1000	—
Hidden 1	512	ReLU
Hidden 2	256	ReLU
Hidden 3	256	ReLU
Output	2	Linear

Table 6.4.1: Architecture of the fully connected neural network used to estimate two noise parameters.

The network is a Multilayer Perceptron (MPL) comprising an input

layer, three hidden layers, and an output layer, as summarized in Table 6.4.1. A schematic illustration of the model is shown in Fig. 6.4.1.

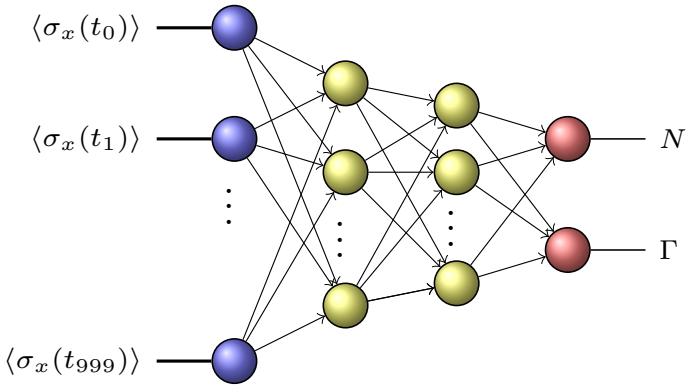

Figure 6.4.1: Schematic representation of the neural network used in Scenario 1. The input layer receives the time series of $\langle\sigma_x(t)\rangle$ and outputs estimates for the noise parameters N and Γ.

The network is trained by minimizing the 'mean squared error' cost function:

$$C(\{\boldsymbol{y}_j, \hat{\boldsymbol{y}}_j\}_j) = \frac{1}{M}\sum_j \left(\boldsymbol{y}_j - \hat{\boldsymbol{y}}_j\right)^2. \tag{6.11}$$

where, $\boldsymbol{y}_j$, $\hat{\boldsymbol{y}}_j$ are the NN predicted and the target output vectors for $j-th$ sample respectively and M is the total number of samples. The performance of the trained model on the test set is illustrated in Fig. 6.4.2. Panels (a) and (b) show the predicted versus target values of N and Γ, respectively. Black dots represent the target values, while orange dots correspond to the predictions made by the neural network. Panels (c) and (d) of Fig. 6.4.2 show the prediction errors for N and Γ, defined as the relative differences between the true and predicted values:

$$\Delta N = \frac{N^{(\text{tar})} - N^{(\text{pred})}}{N^{(\text{tar})}}, \quad \Delta\Gamma = \frac{\Gamma^{(\text{tar})} - \Gamma^{(\text{pred})}}{\Gamma^{(\text{tar})}}. \tag{6.12}$$

As previously discussed in Fig. 6.3.1, when $N = 7$, the coherence dynamics of the qubit under RTN closely resembles that generated by Gaussian OU noise. Hence, further increases in N do not substantially affect

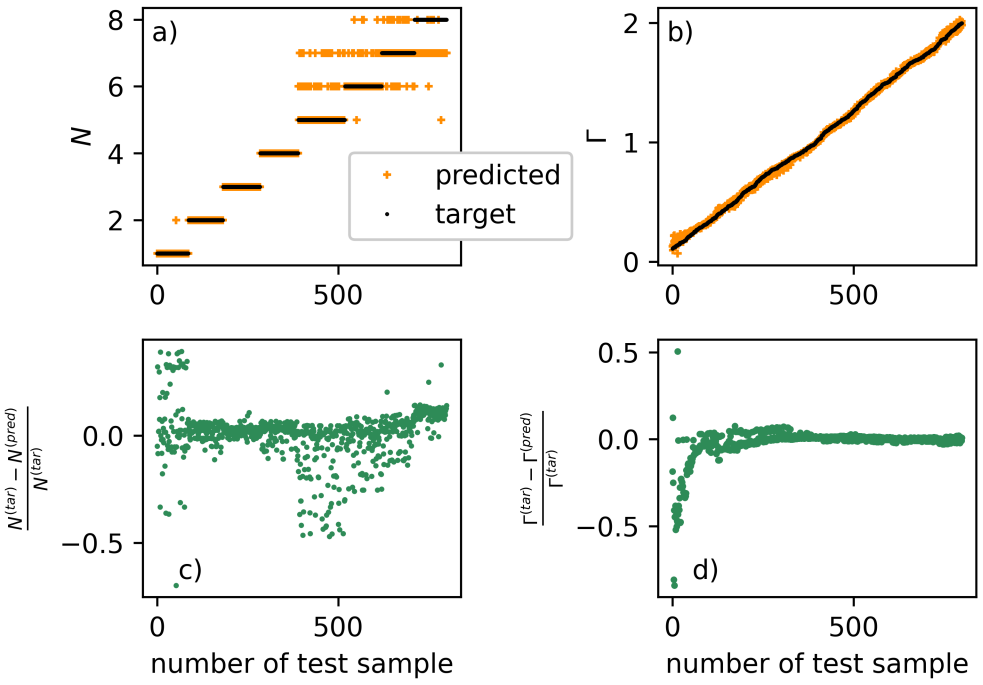

Figure 6.4.2: (a) True and predicted values of N for test samples. (b) True and predicted values of Γ. (c) Prediction error $\Delta N = \frac{N^{(\mathrm{tar})} - N^{(\mathrm{pred})}}{N^{(\mathrm{tar})}}$. (d) Prediction error $\Delta\Gamma = \frac{\Gamma^{(\mathrm{tar})} - \Gamma^{(\mathrm{pred})}}{\Gamma^{(\mathrm{tar})}}$.

the system dynamics. As a result, the NN faces difficulty distinguishing between values of $N > 7$, leading to larger prediction errors in this region. Nevertheless, the network maintains robust predictive accuracy for Γ throughout the parameter range.

6.4.2 Scenario 2

In this scenario, our objective is to estimate four noise parameters: the strength (v_{RTN}) and switching rate (Γ) of the RTN, and the strength (v_{OU}) and inverse correlation time ($k = 1/\tau_c$) of the OU noise. To this end, we generate 10000 qubit coherence evolutions under a hybrid noise envi-

ronment, with the four parameters sampled randomly within the ranges specified in subsec. 6.3.2. Also in this case, we employ a Multilayer Perceptron, its structure is detailed in Table 6.4.2. A schematic representation

Layer	# Neurons	Activation Function
Input	1000	—
Hidden 1	512	LeakyReLU
Hidden 2	512	LeakyReLU
Hidden 3	256	LeakyReLU
Hidden 4	124	ReLU
Output	4	

Table 6.4.2: Architecture of the fully connected neural network used to estimate four noise parameters.

of the neural network used in this scenario is shown in Fig. 6.4.3, and the cost function used is again 'mean square error', eq. (6.11). The perfor-

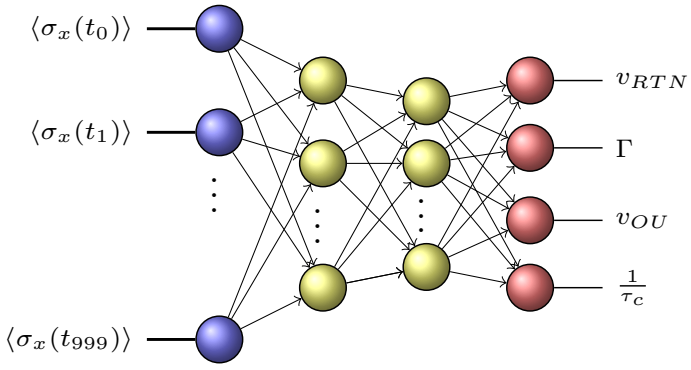

Figure 6.4.3: Schematic representation of the neural network used in Scenario 2 to estimate both RTN and OU noise parameters.

mance of the trained network on the test set is illustrated in Fig. 6.4.4. Panels (a), (b), (e), and (f) display the true (black) and predicted (orange)

values for v_{RTN}, Γ, v_{OU}, and $k = 1/\tau_c$, respectively, plotted against the test sample index. Panels (c), (d), (g), and (h) show the corresponding prediction errors for each parameter, defined as the relative difference between the true and predicted values:

$$\Delta v_{\mathrm{RTN}} = \frac{v_{\mathrm{RTN}}^{(\mathrm{tar})} - v_{\mathrm{RTN}}^{(\mathrm{pred})}}{v_{\mathrm{RTN}}^{(\mathrm{tar})}}, \quad \Delta\Gamma = \frac{\Gamma^{(\mathrm{tar})} - \Gamma^{(\mathrm{pred})}}{\Gamma^{(\mathrm{tar})}}, \tag{6.13}$$

$$\Delta v_{\mathrm{OU}} = \frac{v_{\mathrm{OU}}^{(\mathrm{tar})} - v_{\mathrm{OU}}^{(\mathrm{pred})}}{v_{\mathrm{OU}}^{(\mathrm{tar})}}, \quad \Delta k = \frac{k^{(\mathrm{tar})} - k^{(\mathrm{pred})}}{k^{(\mathrm{tar})}}. \tag{6.14}$$

The network demonstrates high predictive accuracy for all parameters except the switching rate Γ, which exhibits greater variance between predicted and target values. In order to minimize the number of measurement needed to estimate the noise parameters, we examined the ML efficiency while reducing the number of time points used for training. Specifically, we reduced the number of input time points by a factor $h = 2, \ldots, 100$, still keeping the points equidistant, and analyzed the impact on estimation accuracy. This is equivalent to increase the time interval between each measurement of $\langle \sigma_x \rangle$ by the factor h. This approach helps determine the minimum number of time points required to achieve an estimation accuracy comparable to that obtained with the full dataset. Fig. 6.4.5 illustrates the error in the estimation of each parameter defined as the component of the cost function, eq. (6.11), relative to each parameters, as a function of increasing time intervals. It is apparent that with increasing time interval between measurements, the coupling strengths v_{RTN} (v_{OU}) and γ are still predictable, while the coherence time of the OU becomes less predictable. The switching rate of the RTN is always not predictable.

6.5 Conclusion

In this chapter, we have developed a data-driven framework for characterizing Gaussian and non-Gaussian noise in quantum environments using a single qubit as a probe. By leveraging supervised machine learning, we demonstrated the ability to infer key noise parameters from the

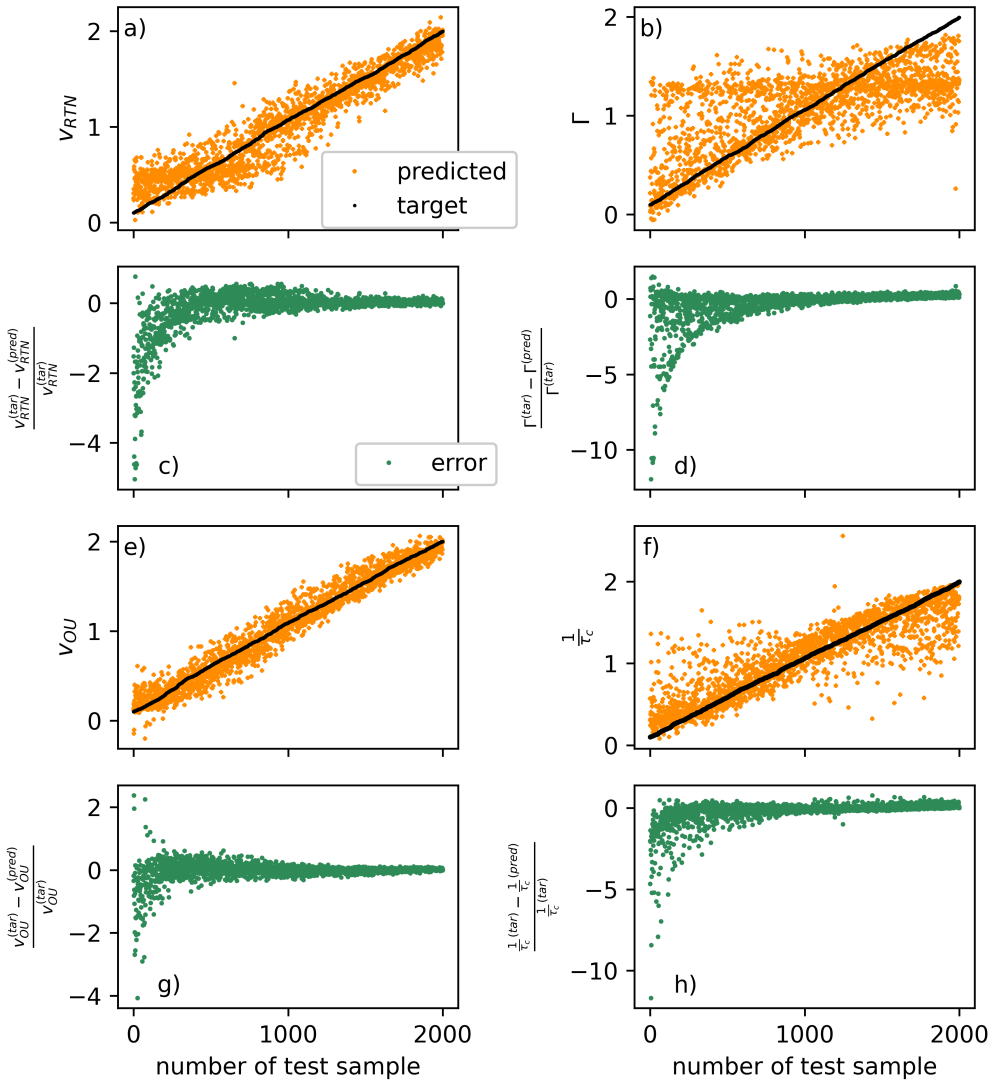

Figure 6.4.4: (a) True vs predicted values of v_{RTN}. (b) True vs predicted values of Γ. (c) Prediction error for v_{RTN}. (d) Prediction error for Γ. (e) True vs predicted values of v_{OU}. (f) True vs predicted values of $k = 1/\tau_c$. (g) Prediction error for v_{OU}. (h) Prediction error for k.

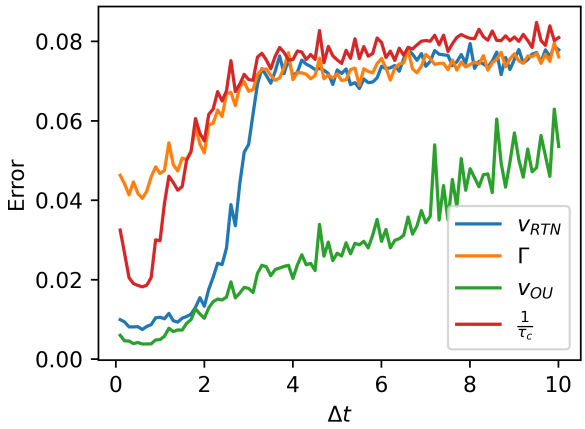

Figure 6.4.5: Error of the four parameters with respect to increasing time interval between measurements.

qubit's time-evolved observables, successfully addressing both fully non-Gaussian and hybrid noise scenarios. The results highlight the potential of machine learning as a powerful tool for noise spectroscopy.

However, this work remains a first step in a broader research direction. Several improvements are currently being developed, including optimizing the number and placement of time points for machine learning input, integrating physics-informed data reduction strategies, and exploring alternative neural network architectures.

In addition, our objective is to extend this framework through a comparative study that includes more complex and realistic noise models such as $1/f$ noise, which is highly relevant for solid-state qubits. We also plan to move beyond the pure dephasing regime by introducing transverse coupling ($\Delta \neq 0$), which allows us to investigate richer dynamics. Another interesting approach is to introduce pulses that drive the qubit either longitudinally or transversely, and analyze which pulses can enhance noise sensitivity and improve parameter estimation. These directions will help us to understand how various types of noise interact with driven dynamics and how different driving protocols influence the identifiability of noise characteristics.

Ultimately, these future developments aim to make our approach more

robust, general, and applicable to real quantum platforms. By combining insights from quantum control, open quantum systems, and machine learning, we hope to advance toward efficient, adaptive, and experimentally viable classical and quantum noise characterization tools.

Appendix

6.A Qubit dynamics under Ornstein-Uhlenbeck noise

We will explore the impact of Ornstein-Uhlenbeck noise on the qubit's evolution. The comprehensive Hamiltonian, in line with equations 6.1and 6.2, is given by

$$H(t) = H_{\text{sys}} + H_{\text{noise}}(t), \tag{6.15}$$

with the noise Hamiltonian now articulated via a single stochastic process x, namely, $H_{noise}(t) = -\frac{v}{2}x(t)\sigma_z$. Here, $x(t)$ represents a stochastic process mimicking OU noise. As we focus on the pure dephasing model, implying $\left[H_{\text{sys}}, H_{\text{noise}}\right] = 0$, and OU is a Gaussian noise, the solution for the coherent evolution of the system is (see eq. (2.71))

$$\rho_{ij}(t) = \rho_{ij}(0)e^{-\mathrm{i}\epsilon t}e^{-\frac{v^2}{2}\left\langle \phi(t)^2 \right\rangle}, \tag{6.16}$$

where $\phi(t) = \int_0^t x(s)ds$ is the noise phase. We consider an OU noise process with a correlation function given by $\langle x(t_1)x(t_2)\rangle = e^{-k|t_1-t_2|}$, where $k = \frac{1}{\tau_c}$ inverse of correlation time. Therefore, the quantity $\left\langle \phi(t)^2 \right\rangle$ can be

calculated by

$$\begin{aligned}\langle \phi(t)^2 \rangle &= \int_0^t \int_0^t \langle x(t_1)x(t_2)\rangle dt_1 dt_2, \\ &= \int_0^t \int_0^t e^{-\frac{|t_1-t_2|}{\tau_c}} dt_1 dt_2, \\ &= \frac{2}{k}(t - \frac{1}{k} + \frac{1}{k}e^{-kt}).\end{aligned} \tag{6.17}$$

Referring back to equation 6.16, we have

$$\langle \rho_{ij}(t)\rangle = \rho_{ij}(0)e^{-i\epsilon t}e^{-v^2\tau_c\left(t-\tau_c+\tau_c e^{-\frac{t}{\tau_c}}\right)}. \tag{6.18}$$

Moving to interaction picture with respect to the qubit Hamiltonian we obtain

$$\langle \rho^I_{ij}(t)\rangle = \rho_{ij}(0)e^{-v^2\tau_c\left(t-\tau_c+\tau_c e^{-\frac{t}{\tau_c}}\right)}. \tag{6.19}$$

6.B Equivalence Between Many Weakly Coupled RTNs and Ornstein-Uhlenbeck Noise

We show here that in the limit of infinitely many weakly coupled random telegraph noise (RTN) sources, the cumulative effect on the qubit becomes equivalent to that of an Ornstein-Uhlenbeck (OU) noise process.

Consider N identical RTNs, each with a coupling strength v, under the constraint $Ng^2 = C^2$, with C constant and $g = 2v/\Gamma$. As $N \to \infty$, it follows that $g \to 0$. We use the approximations: $\alpha = \sqrt{1-g^2} \approx 1 - \frac{g^2}{2}$, $A = \frac{1+\alpha}{2\alpha} \approx 1 - \frac{g^2}{4}$, $1 - A = \frac{\alpha-1}{2\alpha} \approx \frac{g^2}{4}$, where all approximations are valid up to second order in g.

Starting from the coherence evolution expression for N identical RTNs,

as given in eq. (6.3), we write:

$$\begin{aligned}\rho_{01}(t)_{\text{RTN}} &= \lim_{g\to 0}\rho_{01}(0)\left[Ae^{-\frac{\Gamma}{2}(1-\alpha)t}+(1-A)e^{-\frac{\Gamma}{2}(1+\alpha)t}\right]^N \\ &= \lim_{g\to 0}\rho_{01}(0)\left[\left(1-\frac{g^2}{4}\right)e^{-\Gamma\frac{g^2}{4}t}+\frac{g^2}{4}e^{-\Gamma\left(1-\frac{g^2}{4}\right)t}\right]^{C^2/g^2} \\ &= \lim_{g\to 0}\rho_{01}(0)\,e^{-\Gamma C^2 t/4}\left[1+\frac{g^2}{4}\left(1-e^{-\Gamma t\left(1-\frac{g^2}{2}\right)}\right)\right]^{C^2/g^2}.\end{aligned} \tag{6.20}$$

Now consider the limit:

$$\lim_{g\to 0}\left(1+ag^2\right)^{C^2/g^2}=e^{aC^2}.$$

Applying this identity with $a=\frac{1}{4}\left(1-e^{-\Gamma t\left(1-\frac{g^2}{2}\right)}\right)$, we obtain:

$$\rho_{01}(t)_{\text{RTN}}=\rho_{01}(0)\,e^{-\Gamma C^2 t/4}\cdot e^{\frac{C^2}{4}\left(1-e^{-\Gamma t}\right)}. \tag{6.21}$$

Combining the exponents, we finally get:

$$\rho_{01}(t)_{\text{RTN}}=\rho_{01}(0)\,e^{-\frac{\Gamma C^2}{4}\left(t-\frac{1}{\Gamma}+\frac{1}{\Gamma}e^{-\Gamma t}\right)}, \tag{6.22}$$

which is identical in form to the coherence decay under Ornstein-Uhlenbeck noise, as derived in eq. (6.8).

Chapter 7

Conclusions

The advancement of Quantum Technologies relies heavily on our ability to precisely control quantum systems and to mitigate the effects of environmental noise, which remains one of the central challenges in building reliable quantum devices. In this regard, machine learning (ML) has recently emerged as a powerful tool across various areas of quantum technology, including quantum control, state tomography, and error correction. Its ability to learn directly from raw measurement data has opened new avenues for analyzing complex quantum environments. Notably, ML models can extract meaningful features and infer properties of environmental noise such as correlations, memory effects, or statistical nature without requiring explicit knowledge of the system-environment interaction or a detailed microscopic model. This data driven approach not only simplifies traditionally intractable tasks but, in many cases, surpasses the limitations of conventional analytical techniques.

In this PhD thesis, the central focus is to harness ML techniques to infer and characterize noise in quantum systems.

While noise mitigation in single-qubit systems is relatively well understood, extending these techniques to multi-qubit and multi-level platforms remains a significant challenge. On one hand, noise correlations between qubits can severely degrade system performance by reducing coherence time and gate fidelities. On the other hand, as the system size increases, the analytical treatment of all possible noise correlations becomes increasingly intractable. Moreover, modern quantum error cor-

rection techniques require prior knowledge of the underlying noise characteristics, especially correlated noise. Therefore, extracting and understanding noise correlations is of paramount importance for the development of scalable and reliable quantum technologies.

In this context, we studied the dynamics of open quantum systems and the impact of environmental decoherence, specifically focusing on classically described noise models. Our approach leverages the quantum system itself as a probe to extract information about the underlying noise characteristics, along with ML to analyze the system's response to various effect of the environmental noise.

In the first part of the thesis, we employed a hybrid strategy that combines advanced quantum control protocols with machine learning tools. In particular, we investigated the behavior of a three-level quantum network driven via the Coherent Tunneling by Adiabatic Passage (CTAP) or Stimulated Raman Adiabatic Passage (STIRAP) protocol while interacting with classical environmental noise. We considered five different spatially correlated noise models: three quasistatic non-Markovian types (correlated, anti-correlated, and uncorrelated) and two Markovian types (correlated and anti-correlated). By measuring the population transfer efficiency of the CTAP/STIRAP protocol under various driving conditions, and using these data as inputs to a simple neural network, we demonstrated that it is possible to distinguish between the non-Markovian noise classes and between Markovian and non-Markovian noise, with a classification accuracy of up to 99%. The overall classification accuracy including the two Markovian classes is about 80%. Notably, distinguishing between different Markovian noise correlations was not achievable in this setup. We also showed that the protocol exhibits robustness against measurement errors, emphasising its practical viability.

Building upon these results, we extended the methodology to a more complex system composed of two ultra-strongly coupled qubits, each subjected to local, spatially correlated noise as defined previously. Remarkably, we showed that in specific parameter regimes, population transfer can also be achieved in this effective four-level system using a STIRAP-like control scheme. In this case, the machine learning model demonstrated an improved capacity, not only distinguishing among the non-Markovian noise classes but also successfully differentiating between the

Markovian noise types. This is due to the non-symmetric effect of the noise correlations on the eigenbasis of the system. Overall, the classification accuracy reached approximately 90%, highlighting the method's scalability and effectiveness in more realistic multi-qubit scenarios.

In the final part of this thesis, we addressed another important characteristic of classical environmental noise - its Gaussianity. While the Gaussian approximation is widespread in modeling environmental noise, due to their analytical tractability and well-understood statistical properties, realistic quantum environments frequently exhibit non-Gaussian features that cannot be captured by second order statistics alone. This presents a significant challenge for the design of effective control protocols and error correction schemes. To explore this aspect, we considered two representative noise models: Ornstein Uhlenbeck (OU) noise as a Gaussian process and Random Telegraph Noise (RTN) as a prototypical non-Gaussian process. We investigated two scenarios: (i) where a qubit is coupled to a finite number of RTN sources, and (ii) where a hybrid noise environment includes both a large ensemble of weakly coupled RTNs (mimicking Gaussian behavior) and a single strongly coupled RTN contributing a non-Gaussian component. To characterize these environments, we again adopted a data-driven strategy using supervised machine learning. By treating the qubit as a quantum probe and using time-series data of its coherence, we trained neural networks to infer multiple key parameters: in the first scenario, the number of RTN sources (a proxy for Gaussianity) and their switching rate; in the second, the strengths and correlation times of both the Gaussian and non-Gaussian components. This approach is simple and flexible, it does not rely on specific assumptions about the noise distribution, enables multi-parameter estimation, and can be adapted to experimental constraints, such as limited time resolution. We demonstrated that ML can characterise the noise parameters promisingly in both the above mentioned scenarios. Even in the hybrid setting, the model effectively quantifies the relative weight of Gaussian and non-Gaussian noise, offering a direct pathway to assessing the structure of environmental fluctuations.

There are several promising directions that can be explored to extend the scope of this thesis. In particular, the analysis and characterization of $1/f$ noise, which is ubiquitous in many solid-state quantum systems,

is an important and promising route. Understanding and characterizing such noise within a machine learning framework will further strengthen the applicability of data-driven methods in realistic quantum environments. Additionally, we aim to explore unsupervised learning techniques for noise classification and characterization, which could enable the automated discovery of underlying noise structures without the need for labeled training data, paving the way toward scalable, self-learning quantum diagnostics. From a control standpoint, we also intend to investigate alternative quantum driving protocols beyond STIRAP/CTAP, in order to enhance the robustness and adaptability of quantum systems under diverse noise conditions. We also look forward to the experimental implementation of our protocols, particularly in platforms where controllable noise environments can be engineered and our methods can be tested and improved. These future directions are expected to contribute to a more comprehensive, model-agnostic understanding of environmental noise and support the development of resilient, noise-aware quantum technologies.

Overall, this work marks a step toward more effective noise analysis in quantum systems, demonstrating how machine learning can be used to develop adaptable tools for both theoretical studies and experimental applications.

Bibliografia

[1] Michael A Nielsen e Isaac L Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.

[2] Dagmar Bruss e Gerd Leuchs. *Quantum information: from foundations to quantum technology applications*. John Wiley & Sons, 2019.

[3] Shalini Subramani e Santhosh Kumar Svn. "Review of security methods based on classical cryptography and quantum cryptography". In: *Cybernetics and Systems* 56.3 (2025), pp. 302–320.

[4] Artur Ekert e Richard Jozsa. "Quantum computation and Shor's factoring algorithm". In: *Reviews of Modern Physics* 68.3 (1996), p. 733.

[5] TA Kennedy et al. "Single-Qubit Operations with the Nitrogen-Vacancy Center in Diamond". In: *physica status solidi (b)* 233.3 (2002), pp. 416–426.

[6] WG Van der Wiel et al. "Semiconductor quantum dots for electron spin qubits". In: *New Journal of Physics* 8.2 (2006), p. 28.

[7] Göran Wendin. "Quantum information processing with superconducting circuits: a review". In: *Reports on Progress in Physics* 80.10 (2017), p. 106001.

[8] Frank Arute et al. "Quantum supremacy using a programmable superconducting processor". In: *Nature* 574.7779 (2019), pp. 505–510.

[9] *IBM Quantum System Two: the era of quantum utility is here | IBM Quantum Computing Blog — ibm.com*. https://www.ibm.com/quantum/blog/quantum-roadmap-2033. [Accessed 21-05-2025].

[10] Alexandre M Souza, Gonzalo A Álvarez e Dieter Suter. "Robust dynamical decoupling". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 370.1976 (2012), pp. 4748–4769.

[11] MJ Biercuk, AC Doherty e H Uys. "Dynamical decoupling sequence construction as a filter-design problem". In: *Journal of Physics B: Atomic, Molecular and Optical Physics* 44.15 (2011), p. 154002.

[12] Soheil Khadirsharbiyani et al. "Minimizing Coherence Errors via Dynamic Decoupling". In: *Proceedings of the 38th ACM International Conference on Supercomputing*. 2024, pp. 164–175.

[13] Gregory AL White et al. "Non-Markovian quantum process tomography". In: *PRX Quantum* 3.2 (2022), p. 020344.

[14] Yaakov S Weinstein et al. "Quantum process tomography of the quantum Fourier transform". In: *The Journal of chemical physics* 121.13 (2004), pp. 6117–6133.

[15] Boris I Bantysh, A Yu Chernyavskiy e Yu I Bogdanov. "Quantum tomography benchmarking". In: *Quantum Information Processing* 20.10 (2021), p. 339.

[16] Pedro Figueroa-Romero et al. "Randomized benchmarking for non-Markovian noise". In: *Prx quantum* 2.4 (2021), p. 040351.

[17] Joel J Wallman. "Randomized benchmarking with gate-dependent noise". In: *Quantum* 2 (2018), p. 47.

[18] Adarsh Chandrashekar, Soumya Das e Goutam Paul. "Characterization of Noise using variants of Unitarity Randomized Benchmarking". In: *arXiv preprint arXiv:2410.20528* (2024).

[19] Youngkyu Sung et al. "Multi-level quantum noise spectroscopy". en. In: *Nat Commun* 12.1 (2021-02). Publisher: Nature Publishing Group, p. 967.

[20] Santiago Hernández-Gómez et al. “Noise spectroscopy of a quantum-classical environment with a diamond qubit”. In: *Physical Review B* 98.21 (2018), p. 214307.

[21] David F. Wise, John J.L. Morton e Siddharth Dhomkar. “Using Deep Learning to Understand and Mitigate the Qubit Noise Environment”. In: *PRX Quantum* 2.1 (2021-01), p. 010316.

[22] Akram Youssry, Gerardo A. Paz-Silva e Christopher Ferrie. “Characterization and Control of Open Quantum Systems beyond Quantum Noise Spectroscopy”. In: *npj Quantum Inf.* 6.1 (2020-12), p. 95.

[23] E. Paladino et al. “1/F Noise: Implications for Solid-State Quantum Information”. In: *Rev. Mod. Phys.* 86.2 (2014-04), pp. 361–418.

[24] Esteban Calzetta, Albert Roura e Enric Verdaguer. “Stochastic description for open quantum systems”. In: *Physica A: Statistical Mechanics and its Applications* 319 (2003), pp. 188–212.

[25] G. Benenti et al. *Principles of Quantum Computation and Information: A Comprehensive Textbook*. World Scientific, 2019.

[26] Matteo AC Rossi, Claudia Benedetti e Matteo GA Paris. “Engineering decoherence for two-qubit systems interacting with a classical environment”. In: *International Journal of Quantum Information* 12.07n08 (2014), p. 1560003.

[27] Antti P. Vepsäläinen et al. “Impact of ionizing radiation on superconducting qubit coherence”. en. In: *Nature* 584.7822 (2020-08). Publisher: Nature Publishing Group, pp. 551–556.

[28] Claudia Benedetti et al. “Quantum probes for the spectral properties of a classical environment”. In: *Physical Review A* 89.3 (2014), p. 032114.

[29] Leonard Mandel e Emil Wolf. *Optical Coherence and Quantum Optics*. Cambridge ; New York: Cambridge University Press, 1995.

[30] Athanasios Papoulis e S Unnikrishna Pillai. *Probability, Random Variables and Stochastic Processes*. 2002.

[31] Crispin W Gardiner. “Handbook of stochastic methods for physics, chemistry and the natural sciences”. In: *Springer series in synergetics* (1985).

[32] Nicolaas Godfried Van Kampen. *Stochastic processes in physics and chemistry*. Vol. 1. Elsevier, 1992.

[33] Patricia Antunes, Sandra Ferreira e Dario Ferreira. "Higher Order Cumulants". In: *Proceedings of the 2023 6th International Conference on Mathematics and Statistics*. 2023, pp. 7–11.

[34] Claudia Benedetti e Matteo GA Paris. "Characterization of classical Gaussian processes using quantum probes". In: *Physics Letters A* 378.34 (2014), pp. 2495–2500.

[35] Claudia Benedetti. "Decoherence, non-Markovianity and quantum estimation in qubit systems subject to classical noise". In: (2015).

[36] Leigh M Norris, Gerardo A Paz-Silva e Lorenza Viola. "Qubit noise spectroscopy for non-Gaussian dephasing environments". In: *Physical review letters* 116.15 (2016), p. 150503.

[37] Heinz-Peter Breuer e Francesco Petruccione. *The theory of open quantum systems*. OUP Oxford, 2002.

[38] Decoherence Schlosshauer. "The quantum-to-classical transition". In: *The Frontiers Collection (Springer-Verlag, 2007)* (2007).

[39] J Bergli, Yu M Galperin e BL Altshuler. "Decoherence in qubits due to low-frequency noise". In: *New Journal of Physics* 11.2 (2009), p. 025002.

[40] J Bergli, YM Galperin e BL Altshuler. "Decoherence of a qubit by non-Gaussian noise at an arbitrary working point". In: *Physical Review B—Condensed Matter and Materials Physics* 74.2 (2006), p. 024509.

[41] E Paladino et al. "Decoherence and 1/f noise in Josephson qubits". In: *Physical review letters* 88.22 (2002), p. 228304.

[42] Aurélien Géron. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. Third edition. Beijing Boston Farnham Sebastopol Tokyo: O'Reilly, 2023.

[43] Stephen Marsland. *Machine Learning: An Algorithmic Perspective, Second Edition*. 2nd. Chapman & Hall/CRC, 2014.

[44] Sebastian Raschka e Vahid Mirjalili. *Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2*. Packt publishing ltd, 2019.

[45] A. Burkov. *The Hundred-Page Machine Learning Book*. Vol. 1. Andriy Burkov Quebec City, QC, Canada, 2019.

[46] Akash Goel, Amit Kumar Goel e Adesh Kumar. "The role of artificial neural network and machine learning in utilizing spatial information". In: *Spatial Information Research* 31.3 (2023), pp. 275–285.

[47] Oludare Isaac Abiodun et al. "State-of-the-art in artificial neural network applications: A survey". In: *Heliyon* 4.11 (2018).

[48] Harshit Poddar. "From neurons to networks: Unravelling the secrets of artificial neural networks and perceptrons". In: *Deep Learning in Engineering, Energy and Finance*. CRC Press, 2024, pp. 25–79.

[49] Florian Marquardt. "Machine Learning and Quantum Devices". In: *SciPost Physics Lecture Notes* (2021-05), p. 029.

[50] Christiane P. Koch et al. "Quantum Optimal Control in Quantum Technologies. Strategic Report on Current Status, Visions and Goals for Research in Europe". In: *EPJ Quantum Technol.* 9.1 (2022-12), p. 19.

[51] Antonio Acín et al. "The Quantum Technologies Roadmap: A European Community View". In: *New J. Phys.* 20.8 (2018-08), p. 080201.

[52] Wojciech Hubert Zurek. "Decoherence, Einselection, and the Quantum Origins of the Classical". In: *Rev. Mod. Phys.* 75.3 (2003-05), pp. 715–775.

[53] Mario Krenn et al. "Artificial Intelligence and Machine Learning for Quantum Technologies". In: *Phys. Rev. A* 107.1 (2023-01), p. 010101.

[54] Valentin Gebhart et al. "Learning Quantum Systems". In: *Nat Rev Phys* (2023-02), pp. 1–16.

[55] J. Brown et al. "Reinforcement Learning-Enhanced Protocols for Coherent Population-Transfer in Three-Level Quantum Systems". In: *New J. Phys.* 23.9 (2021-09), p. 093035.

[56] Luigi Giannelli et al. “A Tutorial on Optimal Control and Reinforcement Learning Methods for Quantum Technologies”. In: *Physics Letters A* 434 (2022-05), p. 128054.

[57] Murphy Yuezhen Niu et al. “Universal Quantum Control through Deep Reinforcement Learning”. In: *npj Quantum Inf* 5.1 (2019-04), pp. 1–8.

[58] Pierpaolo Sgroi, G. Massimo Palma e Mauro Paternostro. “Reinforcement Learning Approach to Nonequilibrium Quantum Thermodynamics”. In: *Physical Review Letters* 126.2 (2021), p. 020601.

[59] L. Banchi et al. “Modelling Non-Markovian Quantum Processes with Recurrent Neural Networks”. In: *New J. Phys.* 20.12 (2018-12), p. 123030.

[60] Giacomo Torlai et al. “Neural-Network Quantum State Tomography”. In: *Nature Phys.* 14.5 (2018-05), pp. 447–450.

[61] Adriano Macarone Palmieri et al. “Experimental Neural Network Enhanced Quantum Tomography”. In: *npj Quantum Inf* 6.1 (2020-02), pp. 1–5.

[62] Raphaël Couturier et al. “Characterization of a Driven Two-Level Quantum System by Supervised Learning”. In: *Entropy* 25.3 (2023-03), p. 446.

[63] Élie Genois et al. “Quantum-Tailored Machine-Learning Characterization of a Superconducting Qubit”. In: *PRX Quantum* 2.4 (2021-12), p. 040355.

[64] S. Bandyopadhyay et al. “Applications of Neural Networks to the Simulation of Dynamics of Open Quantum Systems”. In: *Chemical Physics*. Ultrafast Photoinduced Processes in Polyatomic Molecules:Electronic Structure, Dynamics and Spectroscopy (Dedicated to Wolfgang Domcke on the Occasion of His 70th Birthday) 515 (2018-11), pp. 272–278.

[65] Di Luo et al. “Autoregressive Neural Network for Simulating Open Quantum Systems via a Probabilistic Formulation”. In: *Phys. Rev. Lett.* 128.9 (2022-02), p. 090501.

[66] I. A. Luchnikov et al. “Machine Learning Non-Markovian Quantum Dynamics”. In: *Phys. Rev. Lett.* 124.14 (2020-04), p. 140502.

[67] Felipe F. Fanchini et al. “Estimating the Degree of Non-Markovianity Using Machine Learning”. In: *Phys. Rev. A* 103.2 (2021-02), p. 022425.

[68] Miha Papič e Inés de Vega. “Neural-Network-Based Qubit-Environment Characterization”. In: *Phys. Rev. A* 105.2 (2022-02), p. 022605.

[69] J. Barr et al. “Spectral Density Classification for Environment Spectroscopy”. In: *Mach. Learn.: Sci. Technol.* 5.1 (2024-03), p. 015043.

[70] Stefano M. et al. “Learning the Noise Fingerprint of Quantum Devices”. In: *arXiv:2109.11405 [quant-ph]* (2021-09).

[71] Stefano Martina et al. “Deep Learning Enhanced Noise Spectroscopy of a Spin Qubit Environment”. In: *Mach. Learn.: Sci. Technol.* 4.2 (2023-05), 02LT01.

[72] D. Vion et al. “Manipulating the Quantum State of an Electrical Circuit”. In: *Science* 296.5569 (2002-05), pp. 886–889.

[73] J. Roffe. “Quantum error correction: an introductory guide”. In: *Contemporary Physics* 60 (2019), p. 226.

[74] E. Campbell. “A series of fast-paced advances in Quantum Error Correction”. In: *Nat. Rev. Phys.* 6 (2024), p. 160.

[75] Lorenza Viola, Emanuel Knill e Seth Lloyd. “Dynamical Decoupling of Open Quantum Systems”. In: *Phys. Rev. Lett.* 82.12 (1999-03), pp. 2417–2421.

[76] G. Falci et al. “Dynamical suppression of telegraph and $1/f$ noise due to quantum bistable fluctuators”. In: *Phys. Rev. A* 70.4 (2004-10). Publisher: American Physical Society, p. 040101.

[77] Dieter Suter e Gonzalo A. Álvarez. “Colloquium: Protecting quantum information against environmental noise”. In: *Rev. Mod. Phys.* 88.4 (2016-10). Publisher: American Physical Society, p. 041001.

[78] Antonio D’Arrigo et al. *Open-Loop Quantum Control of Small-Size Networks for High-Order Cumulants and Cross-Correlations Sensing*. 2024-01.

[79] Giuseppe Falci, Pertti J. Hakonen e Elisabetta Paladino. “1/f noise in quantum nanoscience”. In: *Encyclopedia of Condensed Matter Physics (Second Edition)*. A cura di Tapash Chakraborty. Oxford: Academic Press, 2024-01, pp. 1003–1017.

[80] Alexander A. Balandin, Elisabetta Paladino e Pertti J. Hakonen. “Electronic noise—From advanced materials to quantum technologies”. In: *Applied Physics Letters* 124.5 (2024-01), p. 050401.

[81] Francesco M. D. Pellegrino, Giuseppe Falci e Elisabetta Paladino. “1/f critical current noise in short ballistic graphene Josephson junctions”. en. In: *Commun Phys* 3.1 (2020-01). Publisher: Nature Publishing Group, pp. 1–8.

[82] Michael J. Biercuk et al. “Optimized dynamical decoupling in a model quantum memory”. In: *Nature* 458 (2009), p. 996.

[83] S. Damodarakurup et al. “Experimental Inhibition of Decoherence on Flying Qubits via “Bang-Bang” Control”. In: *Phys. Rev. Lett.* 103 (4 2009-07), p. 040502.

[84] Adeline Orieux et al. “Experimental on-demand recovery of entanglement by local operations within non-Markovian dynamics”. en. In: *Sci Rep* 5.1 (2015-02). Publisher: Nature Publishing Group, p. 8575.

[85] E. R. Jenista et al. “Optimized, unequal pulse spacing in multiple echo sequences improves refocusing in magnetic resonance”. In: *J. Chem. Phys.* 131 (2009), p. 204510.

[86] Boris Naydenov et al. “Dynamical decoupling of a single-electron spin at room temperature”. In: *Phys. Rev. B* 83 (8 2011-02), p. 081201.

[87] Jonas Bylander et al. “Noise spectroscopy through dynamical decoupling with a superconducting flux qubit”. en. In: *Nature Phys* 7.7 (2011-07). Publisher: Nature Publishing Group, pp. 565–570.

[88] Morten Kjaergaard et al. “Superconducting Qubits: Current State of Play”. en. In: *Annu. Rev. Condens. Matter Phys.* 11.1 (2020-03), pp. 369–395.

[89] A. D'Arrigo et al. "Effects of low-frequency noise cross-correlations in coupled superconducting qubits". en. In: *New J. Phys.* 10.11 (2008-11), p. 115006.

[90] A. B. Zorin et al. "Background Charge Noise in Metallic Single-Electron Tunneling Devices". In: *Phys. Rev. B* 53.20 (1996-05), pp. 13682–13687.

[91] Uwe von Lüpke et al. "Two-Qubit Spectroscopy of Spatiotemporally Correlated Quantum Noise in Superconducting Qubits". In: *PRX Quantum* 1.1 (2020-09). Publisher: American Physical Society, p. 010305.

[92] J. Yoneda et al. "Noise-correlation spectrum for a pair of spin qubits in silicon". en. In: *Nat. Phys.* 19.12 (2023-12). Number: 12 Publisher: Nature Publishing Group, pp. 1793–1798.

[93] Ji Zou, Stefano Bosco e Daniel Loss. *Spatially correlated classical and quantum noise in driven qubits: The good, the bad, and the ugly*. arXiv:2308.03054 [cond-mat, physics:quant-ph]. 2023-08.

[94] G. Chiriacò et al. "Diagrammatic Method for Many-Body Non-Markovian Dynamics: Memory Effects and Entanglement Transitions". In: *Phys. Rev. B* 108.7 (2023-08), p. 075151.

[95] Mikheil Tsitsishvili et al. "Measurement induced transitions in non-Markovian free fermion ladders". In: *SciPost Phys. Core* 7 (2024), p. 011.

[96] T. J. Pope et al. "Coherent Trapping in Small Quantum Networks". In: *Journal of Statistical Mechanics: Theory and Experiment* 2019.12 (2019-12), p. 124024.

[97] Andrew D. Greentree et al. "Coherent Electronic Transfer in Quantum Dot Systems Using Adiabatic Passage". In: *Phys. Rev. B* 70.23 (2004-12), p. 235317.

[98] M. J. Gullans e J. R. Petta. "Coherent Transport of Spin by Adiabatic Passage in Quantum Dot Arrays". In: *Phys. Rev. B* 102.15 (2020-10), p. 155404.

[99] Ian Goodfellow, Yoshua Bengio e Aaron Courville. *Deep Learning*. Adaptive Computation and Machine Learning. Cambridge, Massachusetts: The MIT Press, 2016.

[100] K. Bergmann, H. Theuer e B. W. Shore. "Coherent Population Transfer among Quantum States of Atoms and Molecules". In: *Rev. Mod. Phys.* 70.3 (1998-07), pp. 1003–1025.

[101] Nikolay V Vitanov et al. "Stimulated Raman Adiabatic Passage in Physics, Chemistry, and Beyond". In: *Reviews of Modern Physics* 89.1 (2017), p. 015006.

[102] M. Born e V. Fock. "Beweis Des Adiabatensatzes". In: *Zeitschrift für Physik* 51.3-4 (1928), pp. 165–180.

[103] A. Messiah. *Quantum Mechanics, Vol. 1/2*. North-Holland, Amsterdam, 1961.

[104] Luigi Giannelli e Ennio Arimondo. "Three-Level Superadiabatic Quantum Driving". In: *Phys. Rev. A* 89.3 (2014-03), p. 033419.

[105] Crispin W. Gardiner e Peter Zoller. *Quantum Noise: A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods with Applications to Quantum Optics*. 3. ed., [Nachdr.] Springer Series in Synergetics. Berlin Heidelberg: Springer, 2010.

[106] Andrew L Maas, Awni Y Hannun e Andrew Y Ng. "Rectifier Nonlinearities Improve Neural Network Acoustic Models". In: *ICML Workshop on Deep Learning for Audio, Speech, and Language Processing (WDLASL 2013)*. 2013.

[107] Scott C. Douglas e Jiutian Yu. "Why RELU Units Sometimes Die: Analysis of Single-Unit Error Backpropagation in Neural Networks". In: *2018 52nd Asilomar Conference on Signals, Systems, and Computers*. Pacific Grove, CA, USA: IEEE, 2018-10, pp. 864–868.

[108] Yanqiang Guo et al. "High-Speed Photon Correlation Monitoring of Amplified Quantum Noise by Chaos Using Deep-Learning Balanced Homodyne Detection". In: *Applied Physics Letters* 123.5 (2023-07-31), p. 051101.

[109] Ye-Xiong Zeng et al. “Application of Machine Learning for Predicting Strong Phonon Blockade”. In: *Applied Physics Letters* 118.16 (2021-04-22), p. 164003.

[110] Martín Abadi et al. *TensorFlow: Large-scale Machine Learning on Heterogeneous Systems*. 2015.

[111] Vladimir B. Braginsky, Farid Ya Khalili e Kip S. Thorne. *Quantum Measurement*. en. Google-Books-ID: 42xfQgAACAAJ. Cambridge University Press, 1992-09.

[112] R. Menchon-Enrich et al. “Spatial adiabatic passage: a review of recent progress”. en. In: *Rep. Prog. Phys.* 79.7 (2016-06). Publisher: IOP Publishing, p. 074401.

[113] Yadav P. Kandel et al. “Adiabatic quantum state transfer in a semiconductor quantum-dot spin chain”. en. In: *Nat Commun* 12.1 (2021-04). Publisher: Nature Publishing Group, p. 2156.

[114] G. Falci et al. “Ultrastrong coupling probed by Coherent Population Transfer”. en. In: *Sci Rep* 9.1 (2019-06). Publisher: Nature Publishing Group, p. 9249.

[115] L. Giannelli et al. “Detecting virtual photons in ultrastrongly coupled superconducting quantum circuits”. In: *Phys. Rev. Res.* 6.1 (2024-01). Publisher: American Physical Society, p. 013008.

[116] K. S. Kumar et al. “Stimulated Raman adiabatic passage in a three-level superconducting circuit”. en. In: *Nat Commun* 7.1 (2016-02). Publisher: Nature Publishing Group, p. 10628.

[117] H. K. Xu et al. “Coherent population transfer between uncoupled or weakly coupled states in ladder-type superconducting qutrits”. en. In: *Nat Commun* 7.1 (2016-03). Publisher: Nature Publishing Group, p. 11018.

[118] Musang Gong et al. “Two-Photon-Transition Superadiabatic Passage in a Nitrogen-Vacancy Center in Diamond”. In: *Phys. Rev. A* 109.3 (2024-03), p. 032626.

[119] Jens Siewert, Tobias Brandes e G. Falci. “Advanced control with a Cooper-pair box: Stimulated Raman adiabatic passage and Fock-state generation in a nanomechanical resonator”. In: *Phys. Rev. B* 79.2 (2009-01). Publisher: American Physical Society, p. 024504.

[120] G. Falci et al. “Design of a Lambda system for population transfer in superconducting nanocircuits”. In: *Phys. Rev. B* 87.21 (2013-06). Publisher: American Physical Society, p. 214515.

[121] N. Earnest et al. “Realization of a $\mathrm{\ensuremath{\Lambda}}$ System with Metastable States of a Capacitively Shunted Fluxonium”. In: *Phys. Rev. Lett.* 120.15 (2018-04). Publisher: American Physical Society, p. 150504.

[122] P. G. Di Stefano et al. “Coherent manipulation of noise-protected superconducting artificial atoms in the Lambda scheme”. In: *Phys. Rev. A* 93.5 (2016-05). Publisher: American Physical Society, p. 051801.

[123] G Falci et al. “Effects of Low-Frequency Noise in Driven Coherent Nanodevices”. In: *Phys. Scr.* T151 (2012-11), p. 014020.

[124] B. Spagnolo et al. “Relaxation Phenomena in classical and Quantum systems”. en. In: *Acta Phys. Pol. B* 43.5 (2012), p. 1169.

[125] A. R. Mills et al. “Shuttling a single charge across a one-dimensional array of silicon quantum dots”. en. In: *Nat Commun* 10.1 (2019-03). Publisher: Nature Publishing Group, p. 1063.

[126] Alexandre Blais, Steven M. Girvin e William D. Oliver. “Quantum information processing and quantum optics with circuit quantum electrodynamics”. en. In: *Nat. Phys.* 16.3 (2020-03). Publisher: Nature Publishing Group, pp. 247–256.

[127] P. G. Di Stefano et al. “Population transfer in a Lambda system induced by detunings”. In: *Phys. Rev. B* 91.22 (2015-06). Publisher: American Physical Society, p. 224506.

[128] Jeffrey M Gertler et al. “Protecting a bosonic qubit with autonomous quantum error correction”. In: *Nature* 590.7845 (2021), pp. 243–248.

[129] Florentin Reiter et al. “Dissipative quantum error correction and application to quantum sensing with trapped ions”. In: *Nature communications* 8.1 (2017), p. 1822.

[130] Stefano Martina, Stefano Gherardini e Filippo Caruso. "Machine learning classification of non-Markovian noise disturbing quantum dynamics". In: *Physica Scripta* 98.3 (2023), p. 035104.

[131] Shreyasi Mukherjee et al. "Noise Classification in Three-Level Quantum Networks by Machine Learning". In: *Mach. Learn.: Sci. Technol.* 5.4 (2024-11), p. 045049.

[132] J Yoneda et al. "Noise-correlation spectrum for a pair of spin qubits in silicon". In: *Nature Physics* 19.12 (2023), pp. 1793–1798.

[133] Bruce Shore. "The Theory of Coherent Atomic Excitation". In: *Fusion Technology* 19.3P1 (1991-05), pp. 576–577.

[134] Piotr Szańkowski et al. "Environmental noise spectroscopy with qubits subjected to dynamical decoupling". In: *Journal of Physics: Condensed Matter* 29.33 (2017), p. 333001.

[135] Miguel Orszag. *Quantum optics: including noise reduction, trapped ions, quantum trajectories, and decoherence*. Springer Nature, 2024.

[136] S Brouard e J Plata. "Decoherence of trapped-ion internal and vibrational modes: The effect of fluctuating magnetic fields". In: *Physical Review A—Atomic, Molecular, and Optical Physics* 70.1 (2004), p. 013413.

[137] L Jacak et al. "Unavoidable decoherence in semiconductor quantum dots". In: *Physical Review B—Condensed Matter and Materials Physics* 72.24 (2005), p. 245309.

[138] Guillaume Cassabois e Robson Ferreira. "Dephasing processes in a single semiconductor quantum dot". In: *Comptes Rendus. Physique* 9.8 (2008), pp. 830–839.

[139] Youngkyu Sung et al. "Non-Gaussian Noise Spectroscopy with a Superconducting Qubit Sensor". In: *Nature Communications* 10.1 (2019-09), p. 3715.

[140] Leigh M. Norris, Gerardo A. Paz-Silva e Lorenza Viola. "Qubit Noise Spectroscopy for Non-Gaussian Dephasing Environments". In: *Physical Review Letters* 116.15 (2016-04), p. 150503.

[141] Akram Youssry, Gerardo A Paz-Silva e Christopher Ferrie. “Characterization and control of open quantum systems beyond quantum noise spectroscopy”. In: *npj Quantum Information* 6.1 (2020), p. 95.

[142] Clemens Neuenhahn et al. “Recent progress in open quantum systems: Non-Gaussian noise and decoherence in fermionic systems”. In: *physica status solidi (b)* 246.5 (2009), pp. 1018–1023.

[143] Jonas Bylander et al. “Noise spectroscopy through dynamical decoupling with a superconducting flux qubit”. In: *Nature Physics* 7.7 (2011), pp. 565–570.

[144] Gerardo A Paz-Silva, Leigh M Norris e Lorenza Viola. “Multiqubit spectroscopy of Gaussian quantum noise”. In: *Physical review a* 95.2 (2017), p. 022121.

[145] Antonio D’Arrigo et al. “Open-loop quantum control of small-size networks for high-order cumulants and cross-correlations sensing”. In: *Scientific Reports* 14.1 (2024), p. 16681.

[146] Jingfu Zhang et al. “Effect of system level structure and spectral distribution of the environment on the decoherence rate”. In: *Physical Review A—Atomic, Molecular, and Optical Physics* 75.4 (2007), p. 042314.

[147] E. Paladino, L. Faoro e G. Falci. “Decoherence Due to Discrete Noise in Josephson Qubits”. In: *Advances in Solid State Physics*. A cura di Bernhard Kramer. Berlin, Heidelberg: Springer, 2003, pp. 747–762.

[148] E. Paladino et al. “Background Charges Induced Stochastic Fluctuations in Josephson Qubits”. In: *Journal of the Physical Society of Japan* 72.Suppl.A (2003-01), pp. 165–166.

[149] Antonio D’Arrigo et al. “Open-Loop Quantum Control of Small-Size Networks for High-Order Cumulants and Cross-Correlations Sensing”. In: *Scientific Reports* 14.1 (2024-07), p. 16681.