Machine learning non-Markovian two-level quantum noise spectroscopy

Juan M. Scarpetta [1,2,*] John H. Reina [1,2] and Morten Hjorth-Jensen [2]

[1]Departament of Physics and Centre for Bioinformatics and Photonics—CIBioFi, Universidad del Valle, Cali 760032, Colombia

[2]Department of Physics and Center for Computing in Science Education, University of Oslo, Oslo N-0316, Norway

(Received 6 June 2025; accepted 8 November 2025; published 11 December 2025)

We develop machine learning models for the automated characterization of quantum noise spectroscopy for non-Hermitian two-level systems. We use the Random Forest, Support Vector, and Feedforward Neural Network regression algorithms to perform a highly accurate regression of the two-level system-bath coupling strength. High accuracy Ohmicity classification was implemented to provide a complete characterization of the spectral density function. We define a time-averaged trace-distance metric to feed the machine learning algorithms, which, together with numerically exact populations as inputs, produce a highly accurate non-Markovian regression spanning the transition from fast to slow baths and from weak to strong coupling regimes of the interaction. The dynamics database of the non-Hermitian systems has been built up within the independent spin-boson and pure dephasing model.

DOI: 10.1103/2lzl-vpjd

I. INTRODUCTION

Quantum spectroscopy of dissipative two-level systems (TLSs) or quantum bits (qubits) is a key tool for understanding and benchmarking the conditional dynamics that take place in quantum processors [1–5], and for the design of quantum noise tolerant networks and computing platforms [6–9]. Quantum technologies, such as quantum computing, quantum sensing, and quantum information engineering, rely on the precise control of system dynamics and the exploitation of their quantum properties [1,10–18]. However, quantum systems are inevitably immersed in an environment that induces dissipative effects that have a non-negligible impact on the dynamics of the system [19–22].

To address this challenge, the study of open quantum systems [19–24] becomes fundamental for understanding their interaction with their environment, aiming to characterize dissipation phenomena by developing models that describe such complex interactions [21–26]. Several studies have sought to improve our understanding of this process, proposing efficient ways to control its effects [7–9,12,13]. Such manipulations can be achieved, for example, through quantum coherent control techniques [8,27,28] or error correction protocols [7,9,12,13], which require probing the system to gather information about its environment.

We use TLSs or qubits as the building blocks of models of open quantum systems that represent the minimal units of information with which we can study the primary effects of a large reservoir in statistical equilibrium on a TLS undergoing relaxation and decoherence. Models such as the spin-boson model [19,23,25] are commonly used for this purpose. Due to the complexity of the system-environment interactions, such models are often heuristic and oversimplified. For instance, harmonic environments assume that the bath is coupled to the target system within a predefined frequency range. Even in these cases, the analysis is rather complicated and requires sophisticated analytical and numerical methods.

In the Markovian regime, where a large separation of timescales between the system and the environment is required, a key method for describing open quantum systems is the Lindblad master equation. This simplifies the description by assuming a featureless bath in which all dissipative effects are condensed into jump operators with their respective decay rates [19,20]. More robust methods that account for memory (non-Markovian) effects in the system-environment interactions, where the system and bath timescales are comparable, include the hierarchical equations of motion (HEOM) [29,30], the quasiadiabatic path integral (QUAPI) [27,31,32], and the time-evolving matrix product operator [33]. While these methods are numerically exact, they tend to be computationally expensive.

Building on this line of research, we focus on the spin-boson model, which incorporates an interacting bosonic bath of harmonic oscillators and can therefore be characterized by the spectral density (SD) function. The SD function can be used to calculate the correlation function [34] of the environment, determining the timescales of influence, and it encodes all information about the system-bath (S-B) coupling [32,35]. This SD characterization is based on key parameters such as Ohmicity, cutoff frequency, and coupling strength regimes. However, this function is rarely available for arbitrary quantum systems and is typically inferred from experimental data or imposed by *ad hoc* assumptions.

Despite the fact that the SD function encoding the system-environment interaction is inaccessible in practice unless the system is probed [36], one still faces the challenge of *learning*

*Contact author: scarpetta.juan@correounivalle.edu.co

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

something about an unknown environment [26] in order to realistically reproduce the system's dynamics. For this reason, in this work we implement machine learning (ML) tools to predict the parameters that fully characterize the properties of a bosonic environment, using the system's time evolution as input. Extracting these parameters allows us to determine the coupling regimes, Ohmicity, and non-Markovian nature of the interaction.

Recent works [37,38] have attempted to investigate dissipative effects in TLSs using neural networks (NNs). They achieved high accuracy in classifying the Ohmicity of an environment through its SD function, by using the time evolution of the system's density matrix under a pure dephasing model. However, crucial parameters influencing the dynamics, such as coupling strength, cutoff frequency, and tunneling energy (in the case of spin-boson systems), were not considered during training and classification. In this study, we therefore further explore the parameter space of the environment. Here, we have successfully identified scenarios ranging from fast to slow baths and from weak to strong coupling regimes of interaction, enabling the identification and quantification of non-Markovian dynamics.

Previous developments in quantum technologies explored the usage of machine learning techniques to efficiently solve complex data-driven problems. The intersection of ML and quantum physics has been extensively exploited in various fields [39–42]. These methods include reinforcement learning for optimal coherent control [43], quantum thermodynamics [44], quantum circuits for modeling non-Markovian dynamics [45], and NNs for simulating the quantum dynamics of open quantum systems [46]. The work of Krenn *et al.* [42] has several interesting perspectives on machine learning and quantum technologies.

While NNs are the most widely used method due to their adaptable architecture, which can easily be customized for any regression or classification task, this work investigates alternative ensemble methods for regression, such as the Random Forest Regressor (RFR) and the Support Vector Regressor (SVR). These methods also achieve high precision when predicting parameters from complex inputs. Additionally, the metrics used to define various prediction measures are analyzed, including accuracy, mean squared error (MSE), R^2 score, recall, and $F1$ score. These are widely used in such tasks [47–49].

Several metrics have been proposed to determine the non-Markovianity of a quantum evolution process Φ_t. Each metric depends on the physical context (e.g., quantum computing, quantum optics, and few-level systems) [22] whose features can vary considerably depending on the physical scenario under consideration (see, e.g., Ref. [22]). Thus, quantifying the non-Markovianity of a quantum process requires precise metrics, such as the so-called Breuer-Laine-Piilo (BLP) [50,51] and Rivas-Huelga-Plenio (RHP) [52], among others. While these quantify the degree of non-Markovianity, they depend on the choice of initial states [53] or are difficult to compute. Furthermore, these metrics do not always agree, as they capture different aspects of non-Markovianity. They are based on different physical properties, such as information, entanglement, and divisibility, all of which allow memory in quantum systems to be quantified.

Here, we report the successful prediction of the direct, automated characterization of the degree of non-Markovianity of Φ_t using ML methods with different labeled targets, thus avoiding the need for lengthy and complex computations of these metrics.

The paper is structured as follows: Sec. II presents the theoretical framework, including the spin-boson and pure dephasing models. Section III describes the methods used, particularly the dataset generated to train and evaluate the ML algorithms. It also briefly outlines the Random Forest, Support Vector Regressor, and Feedforward Neural Networks (FFNN) algorithms, the technical details of which are provided in Appendix A. Sections IV and V present the main results and discussion. Finally, Sec. VI presents the conclusions. Some of the code that we developed has been open-sourced.

II. THEORETICAL FRAMEWORK

We analyze non-Hermitian quantum systems in which a TLS interacts with a reservoir. We consider the paradigmatic spin-boson model, which captures the interactions between the system and the reservoir [23], in order to construct ML models capable of automating non-Markovian quantum spectroscopy of TLSs.

Two main processes are identified for the interacting dissipative TLS: decoherence, where the phase of the TLS state is randomized by its entanglement with the bath without energy exchange [24], and damping or relaxation, where the TLS exchanges energy with its surrounding environment, eventually reaching thermal equilibrium [21–23]. Mathematically, a non-Hermitian system exhibiting the latter process means that the TLS operator does not commute with the total Hamiltonian [21–25]. Conversely, the former process defines a different type of non-Hermitian system, where the TLS operator commutes with the total Hamiltonian. This is the so-called pure dephasing model, in which there is a loss of coherence in the TLS state, but the total energy of the system remains constant over time.

We use the spin-boson model to analyze the physics of described non-Hermitian quantum systems. Consider a quantum system **S** with two energy levels E_0 and E_1 and a transition frequency ω_0. The system **S** interacts with a bath **B**, leading to dephasing effects and relaxation between the energy levels [21–25,27]. The bath is modeled as a large set of quantum harmonic oscillators, where the kth oscillator couples to the system at frequency ω_k through a coupling constant g_k, and can be described by the so-called *bosonic bath* Hamiltonian $\hat{H}_B = \sum_k \hbar\omega_k(\hat{a}_k^\dagger \hat{a}_k + \frac{1}{2})$, where $\hat{a}_k^\dagger$ ($\hat{a}_k$) are the creation (annihilation) operators of mode k, which satisfy the usual bosonic commutation relations [19,20,54]. Figure 1 illustrates the relevant energies and S-B couplings.

The total Hamiltonian describing the system is given by $\hat{H} = \hat{H}_S + \hat{H}_B + \hat{H}_{\rm SB}$, where $\hat{H}_S$ denotes the TLS Hamiltonian. The interaction Hamiltonian, $\hat{H}_{\rm SB}$, is described as a linear coupling between the system operator $\hat{Q}$ and the bath operator $\hat{B}$ as follows:

$$\hat{H}_{\rm SB} = \hbar\,\hat{Q} \otimes \hat{B}. \tag{1}$$

In turn, $\hat{B}$ is expressed as the sum of the creation and annihilation operators of each mode: $\hat{B} = \sum_k (g_k \hat{a}_k^\dagger + g_k^* \hat{a}_k)$, and the

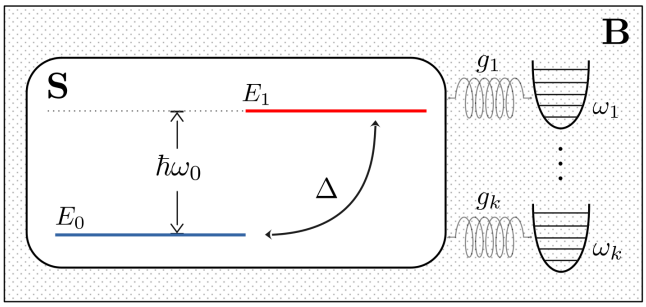


FIG. 1. Sketch of a generic TLS **S** of transition frequency ω_0 and tunneling energy Δ, coupled to a bosonic bath **B** of infinitely many oscillators. The coupling constants g_k define the S-B interaction strength of each mode ω_k.

TLS-bath coupling can be written as

$$\hat{H}_{\rm SB} = \tfrac{1}{2}\hbar\hat{\sigma}_z \otimes \sum_k (g_k \hat{a}_k^\dagger + g_k^* \hat{a}_k). \tag{2}$$

The TLS coupling with bath modes k is characterized by an SD function $J(\omega) = \sum_k |g_k|^2 \delta(\omega - \omega_k)$. Typically, k denotes a much larger number of degrees of freedom than the TLS, meaning that J becomes a continuous function of the frequency ω. Different types of spectral densities have been studied [23–25,32,55–58], each with different properties and effective frequency regimes. Several analytical and numerical techniques can be employed to calculate the dissipative time evolution of the TLS [19–24].

We consider a general SD function of the form

$$J_s(\omega) = \eta \omega_c^{1-s} \omega^s \mathrm{e}^{-\omega/\omega_c}, \tag{3}$$

with an exponential cutoff at frequency ω_c, where η is the system-bath coupling strength and s is the Ohmicity parameter, usually expressed as *Class* 0: $0 < s < 1$, sub-Ohmic; *Class* 1: $s = 1$, Ohmic; and *Class* 2: $s > 1$, super-Ohmic.

Another type of spectral density considered here is the Lorentz-Drude function

$$J_L(\omega) = \frac{2\gamma\omega_c\omega}{\omega^2 + \omega_c^2}, \tag{4}$$

where γ is the coupling strength energy between the system and the bosonic environment with bandwidth ω_c. This SD function is used during the dataset generation of the non-Markovian dynamics in the spin-boson model (cf. Sec. III A), as it provides a convenient implementation of the Matsubara exponential decomposition in the correlation function. Furthermore, by appropriately fine-tuning the parameters s and η in Eq. (3), it is possible to recover Eq. (4) [57]. Figure 2 shows a comparison between $J_s(\omega)$ and $J_L(\omega)$.

The quantum dissipative dynamics associated with such SD functions is plotted in Fig. 3. The construction of machine learning models for the characterization of TLSs' quantum noise spectroscopy using these Ohmic-type and Lorentz-Drude spectral densities will be considered below.

A. Pure dephasing model

As there is no energy exchange between the TLS and the bath, the system's populations remain unchanged. Coherence

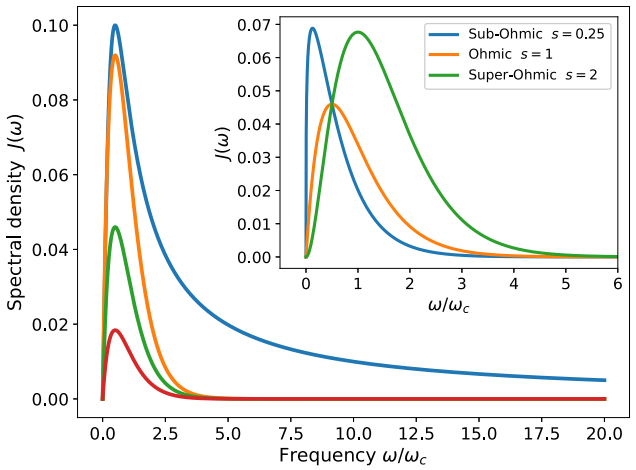


FIG. 2. Comparison of different spectral density functions. The Lorentz-Drude SD function with $\gamma = 0.1$ and $\omega_c = 0.5$ is shown in blue. The orange, green, and red curves represent Ohmic SD functions $J_1(\omega)$ with $\eta = 0.5$, $\eta = 0.25$, and $\eta = 0.1$, respectively, and cutoff $\omega_c = 0.5$. The inset shows the SD functions from Eq. (3) for fixed $\omega_c = 0.5$, $\eta = 0.25$, and different s values.

decay can therefore be calculated exactly, since the TLS operator $\hat{Q} \equiv \hat{\sigma}_z$ commutes with the total Hamiltonian $\hat{H}$. The coupling Hamiltonian, $\hat{H}_{\rm SB}$, is given by Eq. (2) and the TLS Hamiltonian [20,24]

$$\hat{H}_S = \tfrac{1}{2}\hbar\omega_0\hat{\sigma}_z. \tag{5}$$

The analytical solution for the TLS reduced density matrix in the Schrödinger picture is given by

$$\hat{\rho}_S(t) = \begin{pmatrix} \rho_{00} & \rho_{01} e^{-\Gamma(t) + i\omega_0 t} \\ \rho_{01}^* e^{-\Gamma(t) - i\omega_0 t} & 1 - \rho_{00} \end{pmatrix}, \tag{6}$$

where $\Gamma(t)$ is the characteristic decoherence function, which admits the analytical expression [20,24]

$$\Gamma(t) = 4 \int_0^\infty \mathrm{d}\omega\, J_s(\omega) \coth\left(\frac{\hbar\omega}{2k_B T}\right) \frac{1 - \cos(\omega t)}{\omega^2}. \tag{7}$$

The values of ρ_{00} correspond to the initial population of the ground state $|0\rangle$, while ρ_{01} to the initial coherence, which is determined by the initial condition $\hat{\rho}_S(t = 0)$. Since for this model the populations of the system remain constant, the relevant information lies in the coherences, which exhibit damped oscillations over time depending on the bath parameters s, η, and ω_c. Examples of damped coherence oscillations are shown in Figs. 3(a) and 3(b), where the SD function [Eq. (3)] has been used. This model is particularly useful for studying several physical systems, since decoherence usually occurs much faster than relaxation processes. For instance, it has recently been employed to identify basic chemical principles behind electronic quantum decoherence of molecules in the condensed phase [58–60].

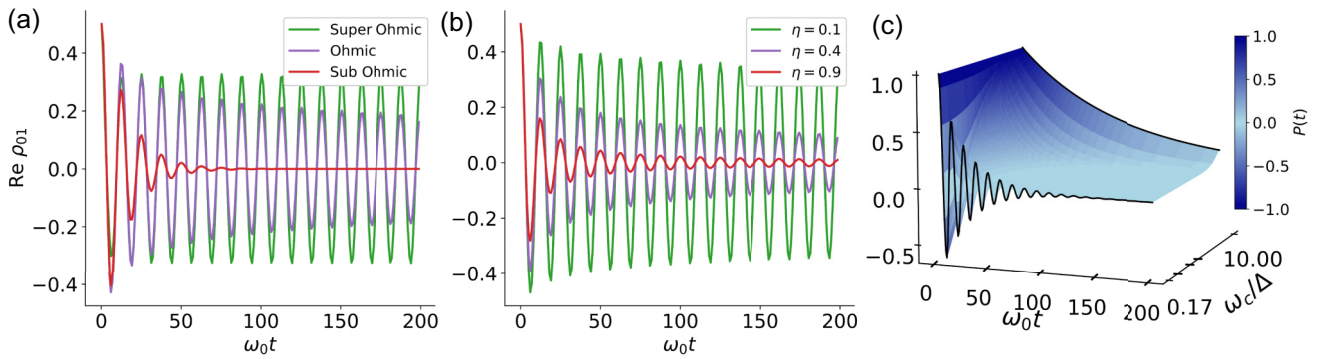


FIG. 3. Case $\Delta = 0$: dynamics of the real part of the coherence Re ρ_{01} as a function of Ohmicity s and coupling strength η. (a) Sub-Ohmic ($s = 0.1$), Ohmic ($s = 1$), and super-Ohmic ($s = 3.5$) spectral densities; $\omega_c = 0.5$, $\eta = 0.25$, and $\omega_0 = 1$. (b) Ohmic dynamics for $\eta = 0.1, 0.4, 0.9$, $\omega_c = 0.5$, and $\omega_0 = 1$. In both cases, Eq. (6) was used with initial condition $\hat{\rho}_S(0) = |+\rangle\langle+|$ at $T = 0$. (c) Case $\Delta \neq 0$: TLS population difference $P(t)$ as a function of time and the ratio of the bath bandwidth to the tunneling energy, ω_c/Δ ($\hbar = 1$). The numerically exact calculation allows the full transition from fast (Markovian) to slow (non-Markovian) baths; $\gamma = 0.25$, $\omega_0 = 1$, $k_B T = 0.5\Delta$, and TLS initial state $\hat{\rho}_S(0) = |1\rangle\langle 1|$.

B. Spin-boson model

The interacting quantum system **S** (see Fig. 1) is described more generally by the Hamiltonian [23]

$$\hat{H}_S = \tfrac{1}{2}\hbar\omega_0\hat{\sigma}_z + \Delta\hat{\sigma}_x, \tag{8}$$

where Δ denotes the tunneling energy and $\hat{\sigma}_j$ ($j = x, z$) are the Pauli matrices. The interaction Hamiltonian is given by Eq. (2); therefore $\hat{H}_S$ [Eq. (8)] does not commute with the $\hat{H}_{SB}$ operator, and a general analytical expression for the $\hat{\rho}_S$ components is no longer available. In this case, oscillations in the coherences and changes in the populations also occur, and numerically exact algorithms such as HEOM [29,30] or QUAPI [27,31,32] can be employed to solve the dynamics.

Throughout the dynamics, the bosonic bath remains in thermal equilibrium at temperature T and can be described by a Gibbs thermal state $\hat{\rho}_B = \frac{1}{Z_B}e^{-\beta\hat{H}_B}$, with associated bath partition function $Z_B = \mathrm{Tr}(e^{-\beta\hat{H}_B})$. For the initial condition at $t = 0$, we consider the separable state $\hat{\rho}(0) = \hat{\rho}_S(0) \otimes \hat{\rho}_B$. In the interaction picture, the $\hat{B}$ operator transforms as $e^{i\hat{H}_B t}\hat{B}e^{-i\hat{H}_B t}$ to obtain $\hat{B}(t) = \sum_k (g_k \hat{a}_k^\dagger e^{i\omega_k t} + g_k^* \hat{a}_k e^{-i\omega_k t})$, with bath autocorrelation function $C(t) = \mathrm{Tr}\{\hat{B}(0)\hat{B}(t)\hat{\rho}_B\}$. We obtain

$$C_\beta(t) = \int_0^\infty J(\omega)\left[\cos(\omega t)\coth\frac{\beta\hbar\omega}{2} - i\sin(\omega t)\right]d\omega, \tag{9}$$

where $\beta = 1/k_B T$ is the inverse temperature, k_B is the Boltzmann constant, and $k_B T$ is the thermal energy.

$C_\beta(t)$ enables all bath correlations relevant to the system to be determined once its spectral density function, $J(\omega)$, has been established. In the numerically exact results reported in Fig. 3(c), we set $J(\omega) \equiv J_L(\omega)$. The bandwidth ω_c sets the timescale for the bath correlation dynamics $\tau_c \equiv 1/\omega_c$. The reorganization energy $E_r = 2\gamma$, as obtained from Eq. (4), where γ is the S-B coupling strength energy. The full structure of $C_\beta(t)$ (including both the real and imaginary parts) has been incorporated into the HEOM method employed here to compute the TLS dynamics, thus enabling a complete non-Markovian description of the system dynamics.

Figure 3(c) plots the TLS population difference, $P(t) \equiv \rho_{11}(t) - \rho_{00}(t)$, which illustrates the transition from fast ($\omega_c \gg \Delta$) to slow ($\omega_c \ll \Delta$) bath, within the strong coupling regime of the TLS-bath interaction for $\gamma = 0.25$. If the bath evolves quickly compared to the TLS time evolution scale within the weak coupling limit $\gamma \ll 1$ (not shown), it quickly loses its memory, rendering the dynamics Markovian [27]. Figure 3(c) shows results covering the scaling limit ($\omega_c \gg \Delta$) and sweeping over the value of ω_c such that the TLS reaches the crossover to the adiabatic limit ($\omega_c \sim \Delta$). In the latter case (and for $\omega_c \ll \Delta$), the bath responds more slowly than the TLS timescale, and non-Markovian effects become dominant. These results demonstrate how the transition between incoherent and coherent behavior depends on the interplay between ω_c and Δ in the strong coupling limit, in agreement with the reported crossover from coherent to incoherent dynamics in damped quantum systems [27,61].

III. METHODS

A. Dataset generation

In order to identify and classify the effect of the bath on the dynamics of the system, the time evolution of the density matrix components of the system was calculated for different parameter values. We first consider the pure dephasing scenario ($\Delta = 0$), with reduced density matrix components given by Eq. (6), and set $T = 0$. The real part of the coherence ρ_{01} is then considered as the dephasing information. The Ohmicity and coupling strength regime are completely determined by the parameters s and η, respectively. In all cases, an evolution of 200 time steps was taken. Therefore, a first scenario is created where both regimes are relatively easy to distinguish, as the distribution of these parameters is well separated. This creates a first classification criterion that is easy to distinguish for machine learning methods.

For the classification of the parameter s, in the scenario with separated data, 10 000 different values were generated in the intervals $s \in (0.1, 0.25) \cup \{1\} \cup (2, 4)$, while for the scenario with nonseparated data, 10 000 values were also taken, uniformly distributed in the interval $s \in (0.1, 4)$. On the other hand, for the regression of the η parameter, a logarithmic distribution was used in the intervals $\eta \in (10^{-2}, 10^{-1}) \cup (0.7, 1)$

for the separated data. In contrast, for the continuous data, a distribution of 10 000 values was generated in the interval $\eta \in (10^{-3}, 1)$. Figures 3(a) and 3(b) show some curves for different values of s and η, calculated from Eq. (6), under the initial condition $\hat{\rho}_S(0) = |+\rangle\langle+|$, $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$.

For the case $\Delta \neq 0$, a more complex dynamics arises, in which both populations and coherences change over time. The characteristic timescale due to the bath response τ_c, the TLS energy Δ, and the strength coupling γ introduce a set of parameters that define dynamical Markovian and non-Markovian regimes [27,62]. These can lead to population difference showing damped oscillations over time or incoherent exponential decay. The ratio ω_c/Δ, together with the TLS-bath coupling strength γ, defines the transition from a fast to a slow bath and from weak to strong coupling regimes of the S-B interaction. The population difference $P(t)$ is plotted as a function of time and of the ratio ω_c/Δ in Fig. 3(c) for $\gamma = 0.25$. This shows the full crossover from a Markovian to a non-Markovian bath, generated using 100 values for such a ratio. The calculations were performed at high temperature, $k_B T = 0.5\Delta$, using the numerically exact HEOM technique. For the regression of the ω_c/Δ parameter, a distribution of 10 000 values was generated in the interval $\frac{\omega_c}{\Delta} \in [0.166, 10.0]$.

B. Non-Markovianity identification and quantification

For the purposes of this discussion, we introduce the ratio of the bosonic bath bandwidth to the tunneling energy, $\alpha \equiv \omega_c/\Delta$, a parameter that captures two fundamental properties of the system under consideration. In the asymptotic limit $\Delta \to 0$, the system can be treated within the pure dephasing model [Eq. (6)], yielding flat population inversion curves at the base of the surface plot [see Fig. 3(c)]. However, for tunneling energies in the scaling limit $\Delta \ll \omega_c$, the population difference exhibits exponential decay over time for finite α, as is clearly shown in Fig. 3(c). This monotonic decay behavior indicates an *irreversible* energy flow from the system to the environment, characteristic of a Markovian process [22,23]. This reference dynamics, denoted $\hat{\rho}_S^{\text{ref}}(t)$, is used to determine the (non-)Markovianity of other processes with different α values. To directly compare two dynamics with different labels, it is necessary to quantify the distinguishability between their states, described by their respective density matrices. The distinguishability between two states $\hat{\rho}_1(t)$ and $\hat{\rho}_2(t)$ at time t can be defined using the trace-distance metric [63]

$$D_{12}(t) = \tfrac{1}{2}\text{Tr}|\hat{\rho}_2(t) - \hat{\rho}_1(t)|, \tag{10}$$

where the operator norm $|\hat{A}| := \sqrt{\hat{A}^\dagger \hat{A}}$. This metric satisfies $0 \leqslant D \leqslant 1$ and exhibits properties relevant to the analysis of non-Markovianity, including contractivity under completely positive maps [64,65]. To measure the global similarity between two dynamics up to time T, we define a parameter σ as the time-averaged $D(t)$:

$$\sigma = \frac{1}{T}\int_0^T D_{12}(t)\, \mathrm{d}t. \tag{11}$$

By setting a reference dynamics to $\hat{\rho}_{\text{ref}} = \hat{\rho}_1$ corresponding to a well-defined Markovian regime, we reparametrize the remaining dynamics in terms of α as

$$\sigma_\alpha = \frac{1}{T}\int_0^T \frac{1}{2}\,\text{Tr}|\hat{\rho}_\alpha(t) - \hat{\rho}_{\text{ref}}(t)|\, \mathrm{d}t. \tag{12}$$

These σ_α values define labels for the regression problem, encoding the deviation from the Markovian scenario by sweeping Δ across the regimes $\Delta \ll \omega_c$ and $\Delta \sim \omega_c$ for $\gamma \ll 1$ and $\gamma \sim 1$. In our study, we perform training and prediction at fixed parameters $\gamma = 0.25$ and $\omega_c = 0.5$, while varying α in the range [0.166, 10.0]. As a Markovian limit, the reference density matrix $\hat{\rho}_{\text{ref}}$ is taken to correspond to the dynamics at $\alpha = 10$. See Appendix C for details.

TABLE I. Metrics for the s classification of the FFNN for non-separated data.

Class	Precision	Recall	$F1$ score
0	1.000 000	0.970 266	0.984 909
1	0.951 359	1.000 000	0.975 073
2	1.000 000	0.978 448	0.989 107
Average	0.983 827	0.983 000	0.983 099

C. Machine Learning models

Using all this information as input, ML models are built to classify and predict the Ohmicity, coupling strength regime, and non-Markovianity of the system dynamics. We have used the Feedforward Neural Network algorithm for both classification and regression problems. In addition, we have implemented the Random Forest Regressor and Support Vector Regressor models. The implementation of these methods for the automated characterization of quantum noise spectroscopy developed here is described in detail in Appendix A.

IV. RESULTS

For the classification of the Ohmicity in the dynamics, after the neural network training process, it was expected that the FFNN would perform well in the case of separated data, as the curves are easy to distinguish. As shown in Fig. 3(a), the sub-Ohmic dynamics show a faster and more pronounced decay, while the super-Ohmic dynamics show a slower oscillatory behavior. In the Ohmic case, an intermediate region between these two behaviors is observed. After about 200 epochs, 100% accuracy is achieved on the test data, confirming the excellent performance of the model in this pilot classification task.

In a second step, the training process was repeated for the case of continuous and uniformly distributed values of s, ensuring a smooth transition between the Ohmicity categories. After 3000 epochs, approximately 98% accuracy was achieved on the test data. The classification results are shown in Fig. 4 and Table I.

Moving onto the regression of the parameter η, the results of the RFR are shown in Fig. 5. We can see how the model predicts the values of η based on the input features. As expected, the results for the distinction of the strong and weak coupling regimes gave the better R^2 score and the lowest Mean Squared

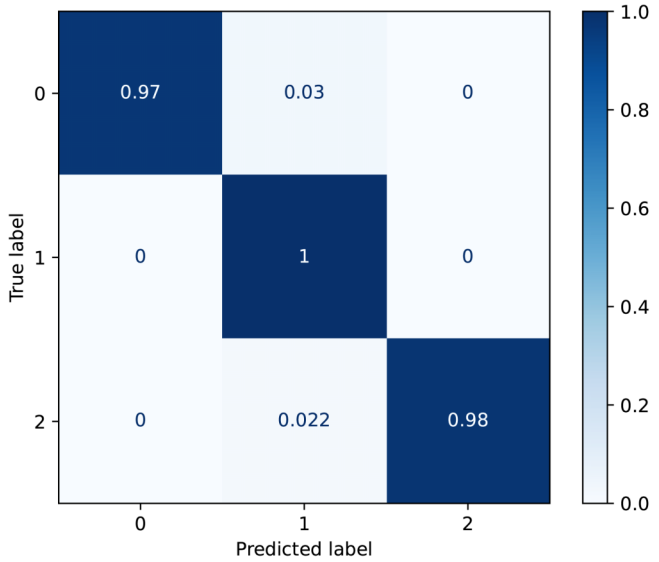


FIG. 4. Normalized confusion matrix showing the results of Ohmicity classification using an FFNN on the test set with nonseparated data.

Error of 6.227×10^{-10} given by the very well-separated data. For the regression with uniformly distributed values, we obtained an R^2 score of 0.9996 and an MSE of 6.0258×10^{-9}, which is an order of magnitude higher than the separated case. These results highlight the robustness of the model even in scenarios with overlapping parameter distributions.

In order to capture the sensitiveness of the overall dynamics with respect to the coupling strength parameter, we use different types of kernels in the SVR to map the temporal features onto different feature spaces in order to reduce the complexity of the dataset and simplify the regression of the data. Table II shows the metrics of the SVR regression for different kernels.

For the FFNN regression model, it was trained using the gradient descent with Adaptive Momentum (ADAM) method described in Eq. (A5) to obtain the model for predicting the values of η. In Fig. 6, it can be observed that after a total of 500 epochs, an effective minimization of the cost function given by Eq. (A1) is achieved.

After the training phase, the model was tested on the test set resulting in a final MSE of 3.3244×10^{-7} for separated η data and 1.2355×10^{-7} for nonseparated data.

For the non-Markovian regime of the dynamics, the dataset generated with the population differences $P(t)$ was used. These components were obtained using the HEOM method, which captures the dynamics for cases of strong coupling and bidirectional flow of information and energy from the bath **B** back to the system **S**.

TABLE II. Metrics of the η regression using SVR.

Kernel	MAE	MSE	R^2 score
Linear	0.064 2568	0.004 982 04	0.9576
RBF	0.057 9192	0.004 195 03	0.9643
Poly	0.064 6946	0.005 028 20	0.9572

TABLE III. Non-Markovianity regression metrics using different labels and the three ML methods discussed in this work (see Appendix A).

Labels	Method	MSE	R^2 score
$\alpha \equiv \frac{\omega_c}{\Delta}$	RFR	$7.592\,11 \times 10^{-6}$	0.999 9937
	SVR—linear	0.005 680 07	0.995 2057
	SVR—RBF	0.005 591 25	0.995 2807
	SVR—poly	0.006 411 22	0.994 5886
	FFNN	$2.299\,32 \times 10^{-6}$	0.999 9984
$\log \alpha$	RFR	$7.767\,82 \times 10^{-8}$	0.999 9994
	SVR—linear	0.006 577 16	0.952 0644
	SVR—RBF	0.005 779 94	0.957 8756
	SVR—poly	0.006 236 88	0.954 5443
	FFNN	$3.764\,51 \times 10^{-8}$	0.999 9997
σ_α	RFR	$1.822\,07 \times 10^{-9}$	0.999 9998
	SVR—linear	0.004 630 15	0.660 1655
	SVR—RBF	0.004 735 60	0.652 4264
	SVR—poly	0.005 172 00	0.620 3962
	FFNN	$1.859\,83 \times 10^{-9}$	0.999 9998

The population dynamics $P(t)$ change drastically depending on the ratio $\alpha := \omega_c/\Delta$ [see Fig. 3(c)]. This parameter distinguishes between coherent and incoherent dynamics, and is therefore used, together with the γ, to categorize non-Markovianity in time evolution. For small values of α, population oscillations occur over time; for larger values, however, a monotonic decay of the excited state is observed. α was logarithmically distributed within the range $[10^{-2}, 10^{1}]$, so the $\log \alpha$ values were used as regression targets to facilitate prediction. The RFR, SVR, and FFNN methods, which are discussed in Appendix A, were applied to this regression task using these input and target sets. The results using all three methods are shown in Fig. 7.

Regressions were performed using three different machine learning methods, with targets including α, $\log \alpha$, and σ_α. After the training process, the Mean Squared Error and the R^2 score were used as error metrics on the test set. These values are summarized in Table III for the different targets.

V. DISCUSSION

We aim to fully characterize the properties of the environment through the time evolution of the components of the system density matrix $\hat{\rho}_S$. Since the bath properties are set by the SD function contained in the correlation function $C_\beta(t)$, the determination of the parameters s, η, ω_c, and γ was sufficient for this task. A first dataset using $J_s(\omega)$ [Eq. (3)] with separate values of s and η was created to establish a scenario where the Ohmicity and coupling strength regimes are easily identifiable from a physical perspective, as they exhibit sub(super)damped oscillations in extreme cases. This task was successfully accomplished, achieving 100% accuracy in the classification of s on the test set and an MSE on the order of 10^{-10} for the η regression, values significantly lower than those obtained for a uniformly distributed dataset. For the latter case, as shown in Table I and the confusion matrix in Fig. 4, the model also achieves 100% accuracy in predicting the sub-Ohmic case (class 0), characterized by a stronger damping of

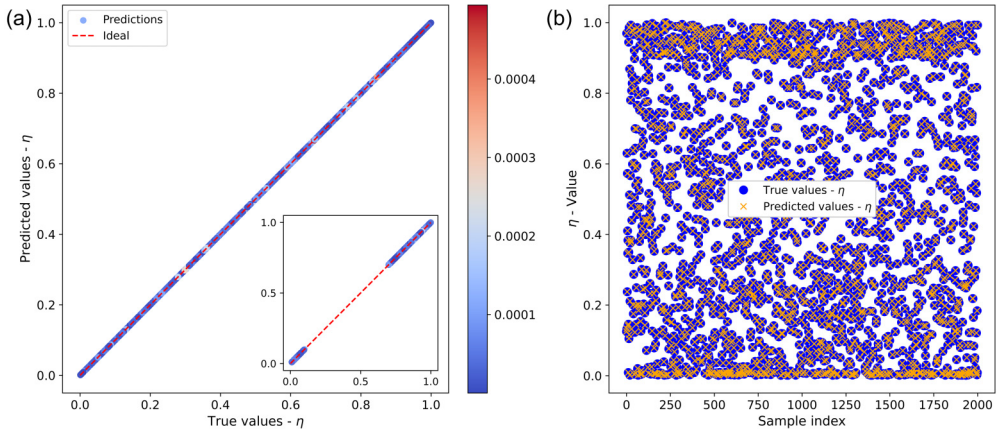


FIG. 5. (a) Results of the η regression using RFR showing the comparison between predicted and true values. The inset shows the regression for separated data and the colorbar shows the absolute error. (b) The contrast between actual values and model predictions is shown over the parameter space of the test set.

the coherence oscillations. For the remaining cases, however, the model does not fully discriminate the transition between Ohmic and super-Ohmic regimes as these cases exhibit oscillations that extend beyond the given timescale. Nevertheless, these cases showed few false positives and false negatives.

To determine the value of η, the three most common regression methods in the literature were used [49]. Each of these methods has its own advantages in terms of numerical implementation and underlying algorithmic logic. The RFR outperformed the other methods, achieving an overall error on the order of 10^{-9} and an R^2 score close to unity. In the case of the SVR, different types of kernels were used to determine which data mapping most efficiently captured the features of the quantum dynamics. Each kernel resulted in an MSE $\sim 10^{-2}$ that is significantly higher than the MSE obtained with RFR or FFNN.

Finally, a neural network was used due to its flexibility to be tuned through various parameters, such as its architecture and training methods. In this case, a simple architecture, previously used in similar applications [37], was adopted to avoid adding unnecessary complexity to the model. The accuracy of the neural network for regression or classification tasks also depends on the values of the learning rate, the batch size, and the number of epochs. To standardize the training process, a preliminary characterization was performed by sweeping over different learning rates and batch sizes for a fixed number of epochs.

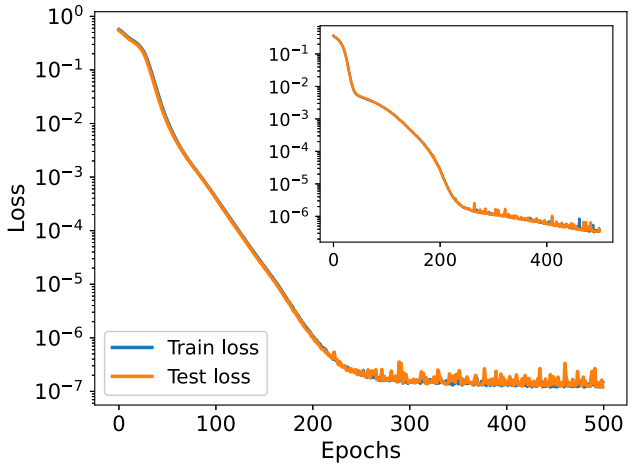


FIG. 6. Training process of the FFNN using the ADAM optimizer with a learning rate $\lambda = 10^{-5}$ and a batch size $\mathrm{Bs} = 64$ to perform η-regression tasks with nonseparated data. Inset: loss evolution for the separated data case.

The main parameters were selected based on their ability to achieve a convergent minimization of the cost function in fewer epochs. For the η regression task, errors were within the same order of magnitude as those obtained with RFR, and a lower error was observed for nonseparated data. These results demonstrate that the FFNN approach performs well in both cases, with a significantly low MSE, indicating a remarkable accuracy in predicting the values of η. Furthermore, the fact that the MSE is lower for nonseparated data suggests that the model can efficiently handle more complex and less differentiated data distributions.

With this comprehensive analysis for the case of pure dephasing, it is evident that models have been established to recover the information related to the dissipative effect of the bath, encoded in the SD function, using only the evolution of the components $\hat{\rho}_S^{(ij)}$. This has been achieved without the need to explicitly calculate the dynamics for the test values or to have prior knowledge of the bath type. This is justified by the fact that a parametrization of the SD $J_s(\omega)$, as shown in Eq. (3), allows the approximation of different scenarios, such as Lorentz-Drude spectral densities by Ohmic approaches [66] or underdamped spectral densities [57], which reproduce the dynamics previously studied [23,24,67]. This lays the groundwork for the study of more complex spectral densities obtained experimentally or baths modeled as superpositions of multiple spectral densities with multiple peaks [58].

For the spin-boson system under consideration, Fig. 3(c) shows that the values of ω_c/Δ and γ can be used to distinguish between Markovian and non-Markovian regimes, with the populations $P(t)$ decaying monotonically or oscillating over time. This suggests using these parameters to determine the non-Markovianity of the process by comparing them with known Markovian reference dynamics.

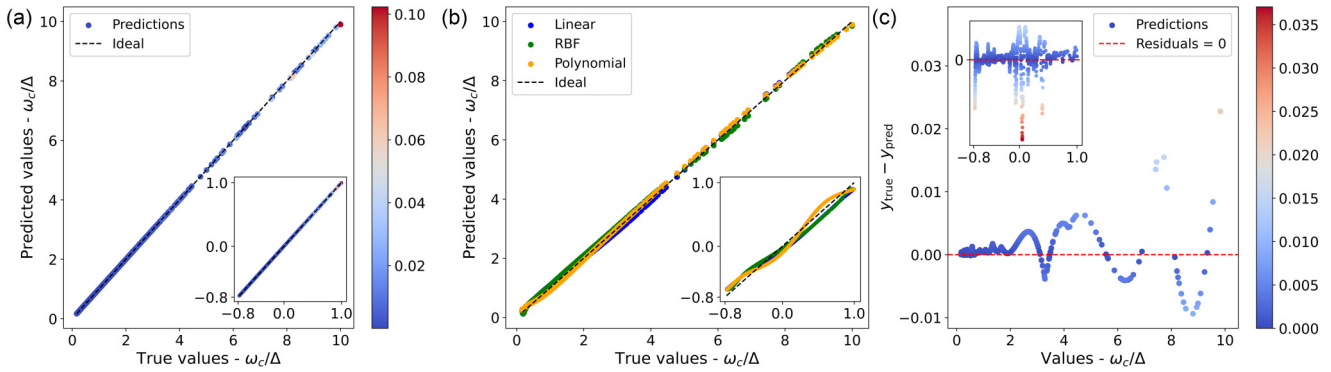


FIG. 7. Regression results for the parameter $\alpha \equiv \omega_c/\Delta$. (a) RFR, comparing the actual values with the predicted values; the colorbar represents the squared absolute error. (b) Regression results using different types of kernels. (c) The error between actual and predicted data (residuals) is shown as a function of ω_c/Δ, and the colorbar represents the squared absolute error. In all cases, the insets show the results of the respective method on test data with a logarithmic scale.

Our analysis covers the entire TLS-bath parameter space. By sweeping over the parameter Δ, we covered both the weak coupling limit, $\gamma \ll 1$, and the strong coupling limit $\gamma \sim 1$ (see Appendix C). We also covered all the regimes in terms of α, analyzing the scaling limit, $\alpha \gg 1$, and the crossover to the adiabatic limit, $\alpha \sim 1$.

In addition, we have proposed a metric that uses the trace distance between operators to identify the flow of information from the environment to the system [Eq. ((11). Although this may be somewhat related to the BLP metric [51], which is sensitive to the initial state of the system, our label avoids the need to sweep over all possible initial conditions, a task that can be computationally intensive, as is the case when computing the BLP metric for arbitrary systems. Our approach simplifies the regression task using three different labels and lays the groundwork for determining non-Markovianity in few-level systems. An extension to systems with more degrees of freedom [45] and more complex dynamics, where regimes are not as easily distinguishable as in the spin-boson model, is a perspective of this work.

VI. CONCLUSIONS

We have demonstrated that, in a generic two-level system under the influence of a bosonic environment, spectral density function classification and regression can be performed with high accuracy. Starting from the baseline case of the spin-boson model under pure dephasing dynamics (i.e., no tunneling), we were able to contrast the model predictions with exact analytical solutions. For the SD Ohmicity classification, we achieved an accuracy above 98%, which is consistent with previous studies [37], but our approach used direct input from the time evolution and a larger dataset. For the coupling strength in the regression analysis, we used the RFR, the SVR, and the FFNN. We investigated the performance of all models, obtaining excellent regression results in all three cases, although RFR and FFNN outperformed SVR. With this classification and regression of the parameters s and η, respectively, we were able to fully characterize the SD function and thus predict and identify environmental effects from the system dynamics.

A crucial outcome of this work was generalizing the dynamics analysis to non-Markovian spin-boson systems using a numerically exact method to generate a dataset. Using the numerical solutions generated for the general dissipative TLS, we successfully performed a regression analysis that quantified the degree of non-Markovianity using three different target labels. Once again, the RFR and FFNN models outperformed the SVR.

These results demonstrate the complete characterization of an unknown SD and the quantification of non-Markovian dynamics using regression and classification approaches. We expect that these methods can be combined to characterize and control the environment under realistic conditions or from direct experimental data where no prior information on possible non-Markovian effects is available. In particular, we expect these models to be useful in the study of quantum dissipation and noise spectroscopy of fermionic and spin environments.

ACKNOWLEDGMENTS

We thank the financial support of the Colombian Ministry of Science (MINCIENCIAS), Grant No. BPIN 2022000100068, and the Norwegian Partnership Programme for Global Academic Cooperation (NORPART), Grant No. NORPART-2021/10436 (QTECNOS). J.H.R. and J.M.S. thank the Department of Physics at the University of Oslo. M.H.-J. was partially supported by U.S. National Science Foundation (NSF) under Grant No. PHY-2310020.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The database and Python codes used to develop this work are available upon reasonable request to the authors.

APPENDIX A: MACHINE LEARNING METHODS

The FFNN algorithms for both classification and regression, together with the RFR and SVR models for the implemented non-Markovian quantum noise spectroscopy characterization, are described below.

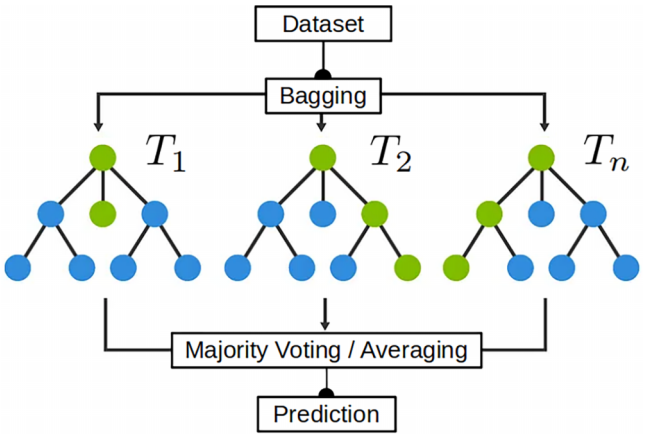


FIG. 8. Sketch of the implementation of the Ensemble Random Forest Regressor with $n = 100$ trees T and majority averaging for the regression task.

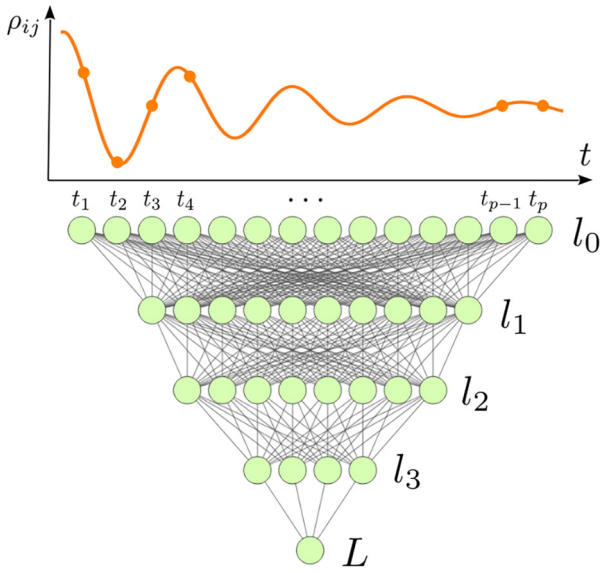


FIG. 9. Architecture of the neural network used for regression and classification. The input layer l_0 consists of $p = 200$ neurons, corresponding to the components of ρ_{ij}. The subsequent hidden layers have dimensions $[l_1, l_2, l_3] = [128, 64, 32]$, each with the ReLU$(\cdot)$ activation function. The output layer contains $L = 3$ ($L = 1$) neurons for the classification (regression) problem.

1. Random forest regressor

RFRs are part of ensemble methods that allow classification and regression tasks to be performed based on the different features present in the data. These methods combine multiple weak learners (decision trees) and make decisions based on majority voting for prediction. These models avoid using a single tree with all the features, instead splitting them into different decision trees, thus reducing strong correlations between predictors and trees [68]. This ensemble approach addresses the overfitting problem that affects decision trees and also reduces the variance in the fit.

For our dataset, the main features are the height of the oscillations and their damping, represented by the values of the density matrix $\rho_t^{(ij)}$ during all 200 time steps of the dynamics. Given that we have a dataset with a total of $n = 10\,000$ entries and a total of $p = 200$ time steps (features), the bagging in the trees consists in randomly splitting the trees into $\sqrt{p} \approx 14$ features. For the regression case, majority averaging is implemented to fit the values and make predictions on a test set. Figure 8 outlines the implementation of the Random Forest.

The architecture for the implementation of the RFR consisted of 100 *number of estimators*, a *minimum samples split* of 2 in order to prevent the data from splitting rapidly along each branch, and an undefined *max depth*, so that each decision tree has as many branches as necessary until pure leaf nodes are obtained, or until all leaves contain fewer than *min-samples-split* samples.

2. Support vector regressor

Support Vector Machine analysis is a popular machine learning tool for classification and regression tasks, developed by Vapnik *et al.* [69]. This algorithm uses support vectors to identify and construct a hyperplane that best fits a given set of data points [70]. For regression problems, SVR is used to find the optimal function $f(x)$ that separates the datasets and minimizes the distance between the targets and the predictions within a precision threshold ϵ. Unlike traditional regression methods, which focus on minimizing the mean squared error, SVR ensures that the predictions remain within a predefined error margin ϵ. This is achieved by optimizing a loss function that penalizes deviations greater than ϵ, allowing greater control over the accuracy of the model. For datasets with more complex features, SVR can work efficiently in higher dimensional spaces. This is because SVR can take as a parameter a kernel that maps the feature space to a hyperdimensional space where finding a separating hyperplane is often asier [69]. Common kernels include linear, polynomial, and radial basis function (RBF). Each of these kernels transforms the data differently, allowing SVR to adapt to a wide variety of regression problems. In this work, an SVR is implemented using the 200 features of dissipative quantum dynamics, and kernels such as linear, polynomial, and RBF are used for regression tasks. The choice of the appropriate kernel depends on the nature of the data and the specific problem to be solved.

3. Feedforward neural network

As a final method, an FFNN model was implemented, which takes the time evolution of the quantum dynamics as an input parameter and returns the regression value or classification, depending on the task (see, for example, Ref. [71], for an in depth discussion of deep learning methods).

The network architecture consists of 200 units in the input layer, followed by hidden layers with sizes [128, 64, 32, out], where out $= 1$ for regression and out $= 3$ for Ohmicity classification. The architecture is shown in Fig. 9. The network was implemented in PyTorch [72] to perform supervised learning for regression and classification tasks. The former is trained by minimizing the cost function

$$C(y, \tilde{y}) = \sum_{k}^{N/\mathrm{Bs}} \frac{1}{|M_k|} \sum_{i \in M_k} (y_i - \tilde{y}_i)^2, \tag{A1}$$

where y_i ($\tilde{y}_i$) is the target (predicted) value. The latter is achieved by minimizing the cross-entropy loss function

$$C(y,\tilde{y}) = -\sum_{k}^{N/\mathrm{Bs}} \frac{1}{|M_k|} \sum_{i\in M_k} y_i \log \tilde{y}_i, \tag{A2}$$

where M_k is the k-mini batch with size $\mathrm{Bs} = |M_k|$. The training process was performed using training and validation sets of size $\frac{3}{5}N$ and $\frac{1}{5}N$, respectively, where N is the total number of data points. Training used 3000 epochs, batch sizes of 32 and 64, and learning rates between 10^{-4} and 10^{-6} to ensure proper convergence of the parameters.

For the classification task, out = 3, since we have three classes and the output vector is passed through a *softmax* function:

$$\sigma(x) = \frac{e^x}{\sum_i e^{x_i}}, \tag{A3}$$

where the exponential function is applied elementwise to the vector x. This function transforms the output vector into a probabilistic vector, such that the ith component represents the probability of belonging to class i. On the other hand, for the regression task, out = 1 and the activation function used is the ReLU function, defined as

$$\mathrm{ReLU}(x) = \max\{x, 0\}. \tag{A4}$$

This function ensures a positive output, which is suitable for predicting the values of η and ω_c/Δ, since these parameters are always positive in the context of the problem. The network is trained to determine the optimal weights and biases $\theta = \{W_i, b_i\}$ that minimize the cost function given in Eq. (A2). For this task, the Adam optimizer [73] was used, which calculates the nth gradient of the cost function $g^{(n)} = \nabla_\theta C(\theta^{(n)})$ and the moments

$$\begin{aligned} s &\leftarrow \rho_1 s + (1-\rho_1)\, g^{(n)} \\ r &\leftarrow \rho_2 r + (1-\rho_2)\, g^{(n)} \odot g^{(n)} \\ t &\leftarrow t+1 \\ \hat{s} &= \frac{s}{1-\rho_1^t}, \quad \hat{r} = \frac{r}{1-\rho_2^t}, \end{aligned}$$

in order to follow the updating rule for the neural network parameters:

$$\theta^{(n+1)} = \theta^{(n)} - \lambda \frac{\hat{s}}{\sqrt{\hat{r}}+\epsilon}, \tag{A5}$$

where λ is the learning rate, $\rho_1 = 0.9$, $\rho_2 = 0.999$, and $\epsilon = 10^{-8}$ for stabilization. After random initialization, the training process starts by following the recursive rules in Eq. (A5) until the stopping criterion is met. This approach allows for efficient and robust optimization of the network parameters, ensuring fast and stable convergence during training.

APPENDIX B: HIERARCHICAL EQUATIONS OF MOTION FOR A DRUDE-LORENTZ BATH

The HEOM method [29,30] is a nonperturbative approach for solving and evaluating the evolution of the density matrix of a quantum dissipative system. This method treats the system-bath interaction nonperturbatively and captures non-Markovian noise correlation times without relying on assumptions such as the Born, Markovian, and rotating-wave approximations.

In the case of a bosonic environment with a Drude-Lorentz spectral density [see Eq. (4)], the bath correlation function can be expressed as an exponential series [34]. In particular, it admits a decomposition into real and imaginary parts,

$$C(t) = C^{\mathrm{R}}(t) + i\, C^{\mathrm{I}}(t) = \sum_{k=0}^{\infty} (L_{1,k} + i\, L_{2,k})\, e^{-\nu_k t}, \tag{B1}$$

with decay rates

$$\nu_0 = \gamma, \quad \nu_k = \frac{2\pi k}{\beta\hbar}\ (k \geqslant 1), \tag{B2}$$

which correspond to Matsubara frequencies. The coefficients $L_{1,k}$ and $L_{2,k}$, which are determined by the Padé expansion [74] for the Drude-Lorentz case, take the form

$$L_{1,0} = \lambda\,\gamma\, \cot\!\left(\frac{\beta\hbar\gamma}{2}\right), \quad L_{2,0} = -\lambda\,\gamma, \tag{B3}$$

$$L_{1,k} = \frac{4\lambda\,\gamma}{\beta\hbar}\, \frac{\nu_k}{\nu_k^2 - \gamma^2}, \quad L_{2,k} = 0\ (k \geqslant 1). \tag{B4}$$

In a numerical implementation, the series is truncated at a finite K.

The HEOM algorithm propagates the full open-system dynamics by introducing a hierarchy of auxiliary density operators (ADOs) $\{\rho_{\mathbf{n}}(t)\}$, which are labeled by the multi-index $\mathbf{n} = (n_0, n_1, \ldots, n_{K-1}) \in \mathbb{N}_0^K$. The physical reduced density matrix of the system corresponds to the root ADO

$$\hat{\rho}_S(t) = \hat{\rho}_{\mathbf{0}}(t) \equiv \hat{\rho}_{0,0,\ldots,0}(t), \tag{B5}$$

and all system observables are evaluated from $\hat{\rho}_S(t)$ as $\langle \hat{O}(t)\rangle = \mathrm{Tr}\{\hat{O}\,\hat{\rho}_S(t)\}$.

Recalling the notation $\hat{H}_S$ for the system Hamiltonian and $\hat{V}$ for the system coupling operator, we have $\hat{H}_{\mathrm{tot}} = \hat{H}_S + \hat{H}_B + \hat{V}\otimes\hat{B}$. In this work, we set $\hat{V} = \hat{\sigma}_z$ and define $\hat{B}(t) = \sum_k (g_k \hat{a}_k^\dagger e^{i\omega_k t} + g_k^* \hat{a}_k e^{-i\omega_k t})$ (Sec. II B). Defining the superoperators $\mathcal{L}_S\hat{X} \equiv [\hat{H}_S, \hat{X}]$, $\hat{A}^\times\hat{X} \equiv [\hat{A}, \hat{X}]$, and $\hat{A}^\circ\hat{X} \equiv \{\hat{A}, \hat{X}\}$, and the unit vectors $\mathbf{e}_k$ that increase n_k by one, the HEOM for the components $\hat{\rho}_{\mathbf{n}}(t)$ read

$$\begin{aligned} \frac{d}{dt}\hat{\rho}_{\mathbf{n}}(t) = &-\left(\frac{i}{\hbar}\,\mathcal{L}_S + \sum_{k=0}^{K-1} n_k\,\nu_k\right)\hat{\rho}_{\mathbf{n}}(t) \\ &- \frac{i}{\hbar}\sum_{k=0}^{K-1} \hat{V}^\times\,\hat{\rho}_{\mathbf{n}+\mathbf{e}_k}(t) \\ &- \frac{i}{\hbar}\sum_{k=0}^{K-1} n_k (c_k\,\hat{V}\,\hat{\rho}_{\mathbf{n}-\mathbf{e}_k}(t) \;-\; c_k^*\,\hat{\rho}_{\mathbf{n}-\mathbf{e}_k}(t)\,\hat{V}). \end{aligned}$$

Here, $c_k = L_{1,k} + iL_{2,k}$ is the complex coefficient, obtained using the decomposition in Eqs. (B3) and (B4). During the iterations, the hierarchy is truncated by limiting the total index depth $\sum_k n_k \leqslant N_{\mathrm{cut}}$ or by using a terminator at the boundary. Convergence must also be checked with respect to K and N_{cut}.

APPENDIX C: NON-MARKOVIANITY METRIC

The metric used for the regression problem, Eq. (12), is based on the so-called BLP measure [50,51], which relies on the distinguishability and indistinguishability of a pair of states represented by their density matrices at an arbitrary time t.

To categorize the Markovianity of a system using the BLP metric, one monitors all time intervals where the trace distance between the states increases, indicating an enhancement in the distinguishability of the states. Under a Markovian evolution governed by a completely positive and divisible dynamical map, the trace distance never increases, reflecting the fact that information flows irreversibly from the system to the environment. Accordingly, the BLP measure identifies non-Markovian behavior under the condition

$$S(t) \equiv \frac{d}{dt}\frac{1}{2}\mathrm{Tr}|\hat{\rho}_2(t) - \hat{\rho}_1(t)| > 0, \tag{C1}$$

which signals a backflow of information from the environment to the system and is interpreted as non-Markovianity. Therefore, the BLP measure of the non-Markovianity of a given dynamics, $\mathcal{N}$, is given by

$$\mathcal{N} = \max_{\hat{\rho}_{1,2}(0)} \int_{S>0} S(t)\, dt, \tag{C2}$$

where the maximization is taken over all possible pairs of initial states and the integration is restricted to the intervals where $S(t) > 0$, where the trace distance exhibits the largest revivals and distinguishability increases. Typical values show that when $\mathcal{N} = 0$, dynamics are purely Markovian, with no information backflow. In contrast, when $\mathcal{N} > 0$, dynamics are non-Markovian. Computing this value numerically requires solving multiple dynamical evolutions over a sweep of all possible initial conditions on the Bloch sphere, thereby characterizing the dissipative dynamical map.

In the main text, the same idea underlying this metric is employed; however, in this case, deviations of the dynamical trajectories are compared with respect to a reference dynamics taken in the Markovian limit. For example, setting $\gamma = 0.25$, $\omega_c = 0.5$, and $\alpha = 10$ yields $\mathcal{N} = 0$, confirming the system's Markovian nature (see Fig. 10). This approach establishes a criterion for identifying and quantifying the Markovianity of a system, using the trace distance to measure these deviations without sweeping over all possible initial conditions. Instead, the method uses observables as inputs for regression and classification, which is one of the main objectives and central ideas of this work.

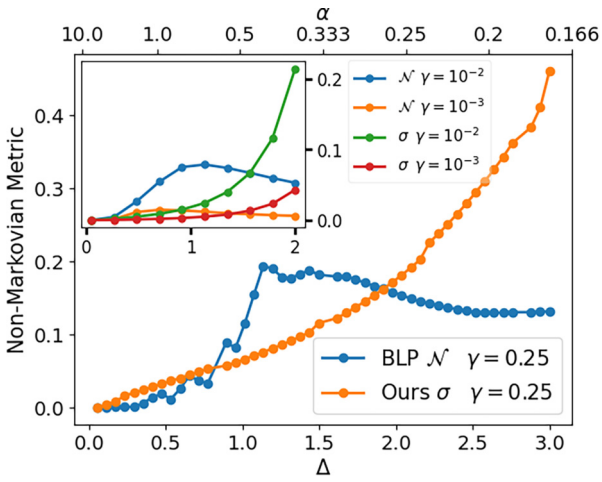


FIG. 10. Comparison between the BLP metric and the target σ_α introduced in this work for the regression and prediction of non-Markovianity. The dynamics were generated using different values of γ, with $\omega_c = 0.5$ and 30 initial conditions of pure states sampled uniformly over the Bloch sphere. The results cover both the weak coupling (inset) and the strong coupling regime (main).

Figure 10 presents a comparison between the two metrics. The analysis was performed in both the weak and strong coupling regimes. The orange curve in the main graph shows our result, which is consistent with that reported in Fig. 3(c) in the main text for the dynamics of the TLS population difference. As we transition from a fast to a slow bath by varying α from left to right on the top axis of the figure within the strong coupling regime ($\gamma = 0.25$), our metric gives higher values monotonically, capturing non-Markovianity. This contrasts with the metric $\mathcal{N}$, which, after some time, gives decreasing values as the bath becomes slower. In the weak coupling scenario (inset, $\gamma = 10^{-2}$ and 10^{-3}), our metric reflects Markovianity ($\sigma = 0$) in regions where the scaling limit is met, deviating slowly and monotonically from it as the bath transitions toward slower dynamics. These results are plotted equivalently, in both cases, on the bottom axis as a function of Δ.

[1] Y. Alexeev *et al.*, Quantum computer systems for scientific discovery, PRX Quantum **2**, 017001 (2021).

[2] S. Kotler, N. Akerman, Y. Glickman, and R. Ozeri, Nonlinear single-spin spectrum analyzer, Phys. Rev. Lett. **110**, 110503 (2013).

[3] L. M. Norris, G. A. Paz-Silva, and L. Viola, Qubit noise spectroscopy for non-Gaussian dephasing environments, Phys. Rev. Lett. **116**, 150503 (2016).

[4] V. Maloney, Y. Oda, G. Quiroz, B. D. Clader, and L. M. Norris, Qubit control noise spectroscopy with optimal suppression of dephasing, Phys. Rev. A **106**, 022425 (2022).

[5] K. W. Chan, W. Huang, C. H. Yang, J. C. C. Hwang, B. Hensen, T. Tanttu, F. E. Hudson, K. M. Itoh, A. Laucht, A. Morello, and A. S. Dzurak, Assessment of a silicon quantum dot spin qubit environment via noise spectroscopy, Phys. Rev. Appl. **10**, 044017 (2018).

[6] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, Noisy intermediate-scale quantum algorithms, Rev. Mod. Phys. **94**, 015004 (2022).

[7] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien, Quantum error mitigation, Rev. Mod. Phys. **95**, 045005 (2023).

[8] D. Suter and G. A. Álvarez, *Colloquium*: Protecting quantum information against environmental noise, Rev. Mod. Phys. **88**, 041001 (2016).

[9] Z. Wang and H. Tang, Artificial intelligence for quantum error correction: A comprehensive review, arXiv:2412.20380.

[10] J. H. Reina, Quantum information processing in nanostructures, Ph.D. thesis, University of Oxford, 2002.

[11] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[12] D. Aasen *et al.*, Roadmap to fault tolerant quantum computation using topological qubit arrays, arXiv:2502.12252.

[13] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. van den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, and A. Kandala, Evidence for the utility of quantum computing before fault tolerance, Nature (London) **618**, 500 (2023).

[14] C. L. Degen, F. Reinhard, and P. Cappellaro, Quantum sensing, Rev. Mod. Phys. **89**, 035002 (2017).

[15] W. Ding, X. Wang, and S. Chen, Fundamental sensitivity limits for non-Hermitian quantum sensors, Phys. Rev. Lett. **131**, 160801 (2023).

[16] K. Herb and C. L. Degen, Quantum speed limit in quantum sensing, Phys. Rev. Lett. **133**, 210802 (2024).

[17] J.-X. Peng, B. Zhu, W. Zhang, and K. Zhang, Enhanced quantum metrology with non-phase-covariant noise, Phys. Rev. Lett. **133**, 090801 (2024).

[18] X. Yang, X. Long, R. Liu, K. Tang, Y. Zhai, X. Nie, T. Xin, J. Li, and D. Lu, Control-enhanced non-Markovian quantum metrology, Commun. Phys. **7**, 282 (2024).

[19] U. Weiss, *Quantum Dissipative Systems*, 4th ed. (World Scientific, 2012).

[20] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, UK, 2002).

[21] F. Campaioli, J. H. Cole, and H. Hapuarachchi, Quantum master equations: Tips and tricks for quantum optics, quantum computing, and beyond, PRX Quantum **5**, 020202 (2024).

[22] I. de Vega and D. Alonso, Dynamics of non-Markovian open quantum systems, Rev. Mod. Phys. **89**, 015001 (2017).

[23] A. J. Leggett, S. Chakravarty, A. T. Dorsey, M. P. A. Fisher, A. Garg, and W. Zwerger, Dynamics of the dissipative two-state system, Rev. Mod. Phys. **59**, 1 (1987).

[24] J. H. Reina, L. Quiroga, and N. F. Johnson, Decoherence of quantum registers, Phys. Rev. A **65**, 032326 (2002).

[25] F. Wilhelm, S. Kleff, and J. von Delft, The spin-boson model with a structured environment: A comparison of approaches, Chem. Phys. **296**, 345–353 (2004).

[26] J. H. Reina, C. E. Susa, and F. F. Fanchini, Extracting information from qubit-environment correlations, Sci. Rep. **4**, 7443 (2014).

[27] J. Eckel, J. H. Reina, and M. Thorwart, Coherent control of an effective two-level system in a non-Markovian biomolecular environment, New J. Phys. **11**, 085001 (2009).

[28] D. Mitra, K. H. Leung, and T. Zelevinsky, Quantum control of molecules for fundamental physics, Phys. Rev. A **105**, 040101 (2022).

[29] Y. Tanimura, Reduced hierarchical equations of motion in real and imaginary time: Correlated initial states and thermodynamic quantities, J. Chem. Phys. **141**, 044114 (2014).

[30] A. Ishizaki and Y. Tanimura, Quantum dynamics of system strongly coupled to low-temperature colored noise bath: Reduced hierarchy equations approach, J. Phys. Soc. Jpn. **74**, 3131 (2005).

[31] N. Makri and D. E. Makarov, Tensor propagator for iterative quantum time evolution of reduced density matrices. I. Theory, J. Chem. Phys. **102**, 4600 (1995).

[32] M. Thorwart, J. Eckel, J. Reina, P. Nalbach, and S. Weiss, Enhanced quantum entanglement in the non-Markovian dynamics of biomolecular excitons, Chem. Phys. Lett. **478**, 234 (2009).

[33] A. Strathearn, P. Kirton, D. Kilda, J. Keeling, and B. W. Lovett, Efficient non-Markovian quantum dynamics using time-evolving matrix product operators, Nat. Commun. **9**, 3322 (2018).

[34] H. Takahashi, S. Rudge, C. Kaspar, M. Thoss, and R. Borrelli, High accuracy exponential decomposition of bath correlation functions for arbitrary and structured spectral densities: Emerging methodologies and new approaches, J. Chem. Phys. **160**, 204105 (2024).

[35] N. Lambert, S. Ahmed, M. Cirio, and F. Nori, Modelling the ultra-strongly coupled spin-boson model with unphysical modes, Nat. Commun. **10**, 3721 (2019).

[36] G. Garau Estarellas, G. L. Giorgi, M. C. Soriano, and R. Zambrini, Machine learning applied to quantum synchronization-assisted probing, Adv. Quantum Technol. **2**, 1800085 (2019).

[37] J. Barr, G. Zicari, A. Ferraro, and M. Paternostro, Spectral density classification for environment spectroscopy, Mach. Learn.: Sci. Technol. **5**, 015043 (2024).

[38] J. Barr, A. Ferraro, M. Paternostro, and G. Zicari, Machine learning-enhanced characterisation of structured spectral densities: Leveraging the reaction coordinate mapping, arXiv:2501.07485.

[39] C. W. Duncan, P. M. Poggi, M. Bukov, N. T. Zinner, and S. Campbell, Taming quantum systems: A tutorial for using shortcuts-to-adiabaticity, quantum optimal control, and reinforcement learning, PRX Quantum **6**, 040201 (2025).

[40] H. Ma, B. Qi, I. R. Petersen, R.-B. Wu, H. Rabitz, and D. Dong, Machine learning for estimation and control of quantum systems, Natl. Sci. Rev. **12**, nwaf269 (2025).

[41] M. K. Gupta, T. Rybotycki, and P. Gawron, On the status of current quantum machine learning software, arXiv:2503.08962.

[42] M. Krenn, J. Landgraf, T. Foesel, and F. Marquardt, Artificial intelligence and machine learning for quantum technologies, Phys. Rev. A **107**, 010101 (2023).

[43] L. Giannelli, S. Sgroi, J. Brown, G. S. Paraoanu, M. Paternostro, E. Paladino, and G. Falci, A tutorial on optimal control and reinforcement learning methods for quantum technologies, Phys. Lett. A **434**, 128054 (2022).

[44] S. Sgroi, G. M. Palma, and M. Paternostro, Reinforcement learning approach to nonequilibrium quantum thermodynamics, Phys. Rev. Lett. **126**, 020601 (2021).

[45] I. Luchnikov, S. Vintskevich, D. Grigoriev, and S. Filippov, Machine learning non-Markovian quantum dynamics, Phys. Rev. Lett. **124**, 140502 (2020).

[46] S. Bandyopadhyay, Z. Huang, K. Sun, and Y. Zhao, Applications of neural networks to the simulation of dynamics of open quantum systems, Chem. Phys. **515**, 272 (2018).

[47] O. Rainio, J. Teuho, and R. Klén, Evaluation metrics and statistical tests for machine learning, Sci. Rep. **14**, 6086 (2024).

[48] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference and Prediction* (Springer Verlag, Berlin, 2009).

[49] K. P. Murphy, *Machine Learning: A Probabilistic Perspective* (The MIT Press, Cambdridge, MA, 2012).

[50] H.-P. Breuer, E.-M. Laine, and J. Piilo, Measure for the degree of non-Markovian behavior of quantum processes in open systems, Phys. Rev. Lett. **103**, 210401 (2009).

[51] E.-M. Laine, J. Piilo, and H.-P. Breuer, Measure for the non-Markovianity of quantum processes, Phys. Rev. A **81**, 062115 (2010).

[52] A. Rivas, S. F. Huelga, and M. B. Plenio, Quantum non-Markovianity: Characterization, quantification and detection, Rep. Prog. Phys. **77**, 094001 (2014).

[53] L. Chen, Y. Yan, M. F. Gelin, and Z. Lü, Dynamics of the spin-boson model: The effect of bath initial conditions, J. Chem. Phys. **158**, 104109 (2023).

[54] M. Schlosshauer, A. P. Hines, and G. J. Milburn, Decoherence and dissipation of a quantum harmonic oscillator coupled to two-level systems, Phys. Rev. A **77**, 022111 (2008).

[55] D. Green, B. S. Humphries, A. G. Dijkstra, and G. A. Jones, Quantifying non-Markovianity in underdamped versus overdamped environments and its effect on spectral lineshape, J. Chem. Phys. **151**, 174112 (2019).

[56] J. Gilmore and R. H. McKenzie, Spin boson models for quantum decoherence of electronic excitations of biomolecules and quantum dots in a solvent, J. Phys.: Condens. Matter **17**, 1735 (2005).

[57] B. S. Humphries, D. Green, and G. A. Jones, Derivation of Lorentz-Drude undamped oscillator hierarchical equations of motion, arXiv:2412.14727.

[58] I. Gustin, C. W. Kim, D. W. McCamant, and I. Franco, Mapping electronic decoherence pathways in molecules, Proc. Natl. Acad. Sci. USA **120**, e2309987120 (2023).

[59] D. Madrid-Úsuga, C. E. Susa, and J. H. Reina, Room temperature quantum coherence vs. electron transfer in a rhodanine derivative chromophore, Phys. Chem. Chem. Phys. **21**, 12640 (2019).

[60] D. Madrid-Úsuga and J. H. Reina, Molecular structure, quantum coherence, and solvent effects on the ultrafast electron transport in BODIPY–C_{60} derivatives, J. Phys. Chem. A **125**, 2518 (2021).

[61] R. Egger, H. Grabert, and U. Weiss, Crossover from coherent to incoherent dynamics in damped quantum systems, Phys. Rev. E **55**, R3809 (1997).

[62] F. K. Wilhelm, Quantum oscillations in the spin-boson model: Reduced visibility from non-Markovian effects and initial entanglement, New J. Phys. **10**, 115011 (2008).

[63] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

[64] G. Guarnieri, A. Smirne, and B. Vacchini, Quantum regression theorem and non-Markovianity of quantum dynamics, Phys. Rev. A **90**, 022110 (2014).

[65] J. M. Sánchez Velázquez, A. Steiner, R. Freund, M. Guevara-Bertsch, C. D. Marciniak, T. Monz, and A. Bermudez, Dynamical quantum maps for single-qubit gates under universal non-Markovian noise, Phys. Rev. Res. **7**, 013008 (2025).

[66] H. Zheng, Dynamics of a two-level system coupled to Ohmic bath: A perturbation approach, Eur. Phys. J. B **38**, 559 (2004).

[67] A. Caldeira and A. Leggett, Quantum tunnelling in a dissipative system, Ann. Phys. **149**, 374 (1983).

[68] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*, 1st ed. (Chapman and Hall/CRC, Boca Raton, Florida, USA, 1984).

[69] C. Cortes and V. Vapnik, Support-vector networks, Mach. Learn. **20**, 273 (1995).

[70] V. N. Vapnik, *The Nature of Statistical Learning Theory* (Springer, New York, 2018).

[71] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (The MIT Press, Cambridge, MA, 2016).

[72] A. Paszke *et al.*, PyTorch: An imperative style, high-performance deep learning library, arXiv:1912.01703.

[73] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, Automatic differentiation in PyTorch, in *NIPS 2017 Workshop on Autodiff* (2017).

[74] J. Hu, M. Luo, F. Jiang, R.-X. Xu, and Y. Yan, Padé spectrum decompositions of quantum distribution functions and optimal hierarchical equations of motion construction for quantum open systems, J. Chem. Phys. **134**, 244106 (2011).