



Detecting Markovianity of quantum processes via recurrent neural networks

Angela Rosy Morgillo^{1,2} · Massimiliano F. Sacchi^{1,3} · Chiara Macchiavello^{1,2}

Received: 23 June 2024 / Accepted: 5 February 2025
© The Author(s) 2025

Abstract

We present a novel methodology utilizing recurrent neural networks (RNNs) to classify Markovian and non-Markovian quantum processes, leveraging time series data derived from Choi states. The model exhibits exceptional accuracy, surpassing 95%, across diverse scenarios, including dephasing and Pauli channels in an arbitrary basis, generalized amplitude damping dynamics, and even in the presence of noise. Additionally, the developed model shows efficient forecasting capabilities for the analyzed time series data. These results suggest the potential of RNNs in discerning and predicting the Markovian nature of quantum processes.

Keywords Markovianity · Recurrent neural networks · Choi state · Time series

1 Introduction

The study of open quantum dynamics plays a crucial role in understanding the behavior of quantum systems interacting with their surrounding environment (Rivas and Huelga 2012; Banerjee et al. 2018). Unlike closed quantum systems, which are isolated from external influences, open quantum systems experience continuous exchanges of energy, information, and particles with their surroundings. Such interactions lead to the phenomenon of quantum decoherence, where the system loses its coherence and exhibits classical-like behavior (Schlosshauer 2019).

One key aspect in characterizing the dynamics of open quantum systems is the concept of Markovianity. Markovian processes are stochastic processes where the future state of the system depends only on its present state and not on the sequence of events that preceded it (Rivas et al. 2014). In the theory of open quantum systems, Markovian dynamics provides a powerful framework for describing the time evolution of reduced open system states. This evolution is condensed in

a quantum master equation, typically expressed in the Lindblad form, in which a weak system-environment coupling is assumed (Lindblad 1976; Merkli 2020, 2022).

Nonetheless, in nature, the Markovian assumption often breaks down, especially in scenarios characterized by strong system-environment coupling and low temperatures, leading to the emergence of non-Markovian dynamics and information backflow (Hall et al. 2014; Shrikant and Mandayam 2023; Utagi et al. 2020).

In recent years, distinguishing Markovian and non-Markovian dynamics in open quantum systems has gained substantial interest in various domains such as quantum information theory (Bylicka et al. 2014; Cheong et al. 2022; Liu et al. 2013), quantum control (Reich et al. 2015), system-environment coupling modeling (Li et al. 2018), quantum error mitigation (Takagi et al. 2022; Hakoshima et al. 2021), and quantum thermodynamics (Mutti et al. 2022; Choquehuanca et al. 2023), in which non-Markovian dynamics has been shown to facilitate faster energy conversion of heat and work (Huang and Zhang 2022).

At the same time, machine and deep learning have recently achieved notable results, particularly in computer vision (Chai et al. 2021) and natural language processing (Khurana et al. 2023; Min et al. 2023), with transformer models revolutionizing language understanding (Deng and Lin 2022).

Indeed, machine learning has shown great potential also in quantum error mitigation (Kim et al. 2020), entanglement

✉ Angela Rosy Morgillo
angelarosy.morgillo01@universitadipavia.it

¹ Dipartimento di Fisica, Università di Pavia, Via Bassi 6, I-27100 Pavia, Italy

² INFN Sezione di Pavia, Via Bassi 6, I-27100 Pavia, Italy

³ CNR-Istituto di Fotonica e Nanotecnologie, Piazza Leonardo da Vinci 32, I-20133 Milan, Italy

quantification (Gray et al. 2018; Koutny et al. 2023) and classification (Harney et al. 2021), and in quantum thermodynamics for optimal control of finite time processes in quantum thermal machines (Khait et al. 2022), for predicting open quantum dynamics described by time-local generators (Mazza et al. 2021), for reducing entropy production in closed quantum systems out of equilibrium (Sgroi et al. 2021), for classifying the environmental parameters of single-qubit dephasing channels using a time series approach (Palmieri et al. 2021), for classifying the spectral density characterizing the dynamics of a system (Barr et al. 2023), for modeling non-Markovian effects in several regimes by using recurrent neural networks (Banchi et al. 2018), and for identifying Pareto-optimal cycles (Erdman et al. 2023).

In this study, we present a model based on recurrent neural networks (RNNs) (Salehinejad et al. 2017), tailored for analyzing time series data representing Choi states of quantum channels (Choi 1975; Jamiołkowski 1972). A significant characteristic of our methodology is the utilization of only the diagonal components within the Bell basis of the Choi state, resulting in results comparable to those achieved when considering all components (see Fig. 2 for more details). Our model is also effective in classifying quantum processes in the presence of noise, maintaining high accuracy even under noisy conditions. The novelty of our approach lies in using machine learning for the classification task, as, to the best of our knowledge, machine learning has only been applied to assess the degree of non-Markovianity (Fanchini et al. 2021). Unlike current state-of-the-art methods, which rely on analytical solutions and are therefore limited to specific classes of dynamical maps with explicitly solvable master equations, our machine learning-based approach provides greater flexibility by handling datasets for processes with dynamics that

are analytically intractable or computationally unresolvable. The schematic representation of our approach is illustrated in Fig. 1.

Indeed, prior studies have demonstrated the effectiveness of RNNs, along with other machine learning techniques such as artificial neural networks and support vector machines, in discerning various external noise sources, as illustrated in Martina et al. (2023). Furthermore, the Choi state has emerged as a valuable resource for assessing non-Markovianity in quantum processes, as exemplified in Chruściński et al. (2017) and Mallick et al. (2024).

Our primary aim is to tackle the classification problem of discerning between semi-group, Markovian, and non-Markovian dynamics. Here, a semi-group denotes a specific type of Markovian dynamics characterized by constant Lindblad operators and decay rates in the canonical master equation. On the other hand, (non-)Markovianity refers to the case of (non-)positive time-dependent decay rates. We apply this approach to various single-qubit channels, including dephasing, Pauli, and generalized amplitude damping.

Our model achieves accuracies exceeding 95% and demonstrates strong forecasting capabilities, providing robust and reliable predictions.

The manuscript is organized as follows. In Sect. 2, we introduce the dynamics of open quantum systems, defining the concept of (non-)Markovianity. Sect. 3 introduces the quantum models employed in constructing the temporal sequences essential to our study. Moving to Sect. 4, we show the methodologies applied, specifically introducing RNNs and time series and describing the construction of the datasets. Sect. 5 presents the results achieved from process classification and forecasting, showing the utilized network architecture. Finally, in Sect. 6, we derive our conclusions and outline promising paths for future research.

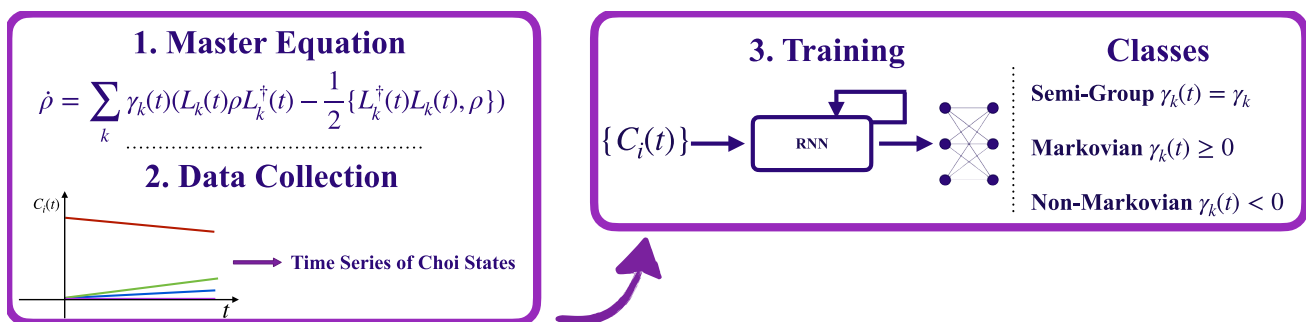


Fig. 1 Our model scheme. We start with the formulation of a master equation to depict the system evolution. We then derive the diagonal components $C_i(t)$ of the Choi states in the Bell basis at different time steps, elucidating the time evolution process outlined in the master equation.

The resulting dataset, including time series inputs and class labels (semi-group, Markovian, and non-Markovian), is then fed into a recurrent neural network, that outputs the probability distribution over the three classes

2 Theory

Exploring the dynamics of open quantum systems is a key focus in modern quantum physics, holding implications for quantum information processing and technology. In this section, we investigate the fundamental aspects of open quantum dynamics, starting from the definition of the master equation, followed by the characterization of dynamical maps and their alternative representation using Choi states. Subsequently, we define the concept of quantum (non-)Markovianity.

2.1 Dynamics of open quantum systems

The master equation Quantum systems are inherently susceptible to decoherence and dissipation when interacting with their surroundings. We use master equations to understand and model these effects systematically. The dynamics of a system coupled to an environment is described by the time non-local Nakajima-Zwanzig master equation (Rivas and Huelga 2012)

$$\dot{\rho}(t) = -\frac{i}{\hbar}[H_S, \rho(t)] + \int_{t_0}^t \mathcal{K}_{t,s}[\rho(s)]ds, \tag{1}$$

where H_S is the system Hamiltonian and the memory kernel $\mathcal{K}_{t,s}$ is a linear map representing the effects of the environment on the system. Utilizing the Born-Markov and secular approximations allows us to approximate the memory kernel, eliminating rapidly fluctuating terms and yielding a memoryless time-local Lindblad master equation:

$$\dot{\rho} = -\frac{i}{\hbar}[H_S, \rho] + \sum_k \gamma_k \left(A_k \rho A_k^\dagger - \frac{1}{2} \{ A_k^\dagger A_k, \rho \} \right), \tag{2}$$

where $\{A_k\}$ represent the Lindblad operators, and the decay rates γ_k are positive coefficients (Hall et al. 2014).

We can generalize the master equation in the canonical form with time-dependent decay rates, where the rotating approximation holds while the Born-Markov approximation does not (Shrikant and Mandayam 2023):

$$\begin{aligned} \dot{\rho} = \mathcal{L}(t)[\rho] = & -\frac{i}{\hbar}[H_S(t), \rho] \\ & + \sum_k \gamma_k(t) \left(A_k(t) \rho A_k^\dagger(t) - \frac{1}{2} \{ A_k^\dagger(t) A_k(t), \rho \} \right). \end{aligned} \tag{3}$$

The canonical form of the master equation requires that the operators $A_k(t)$ constitute a set of orthonormal traceless operators (Hall et al. 2014). This form proves particularly valuable for characterizing non-Markovianity, as discussed in Sect. 2.2, defining a *unique* set of decoherence rates.

Dynamical maps and Choi states The solution of Eq. 3 provides a set of dynamical maps $\Lambda(t, t_0)$, transforming density operators ρ into density operators $\Lambda(t, t_0)[\rho]$. In order to be physically admissible, the maps $\Lambda(t, t_0)$ have to be completely positive (CP) for all times t (Breuer and Petruccione 2007). These maps can be represented as

$$\Lambda(t, t_0)[\rho] = \sum_k E_k(t) \rho E_k^\dagger(t), \tag{4}$$

where $\{E_k(t)\}$ denotes the set of Kraus operators, satisfying $\sum_k E_k^\dagger(t) E_k(t) = I$ (Nielsen and Chuang 2010). From linearity, $\Lambda(t, t_0)$ itself satisfies $\dot{\Lambda}(t, t_0) = \mathcal{L}(t)[\Lambda(t, t_0)]$, and its solution takes the form:

$$\Lambda(t, t_0) = \mathcal{T} e^{\int_{t_0}^t \mathcal{L}(s) ds}, \tag{5}$$

with \mathcal{T} representing the time-ordering operator.

For each dynamical map, a one-to-one correspondence can be established with a bipartite state through the Choi-Jamiołkowski isomorphism (Adlam 2020). Given a completely positive dynamical operator $\Lambda(t, t_0)$, the associated Choi-Jamiołkowski state $\mathcal{C}_\Lambda(t, t_0)$ is derived by applying $\Lambda(t, t_0)$ to a party of a maximally entangled bipartite state $|\phi^+\rangle\langle\phi^+|$ (Mallick et al. 2024):

$$\mathcal{C}_\Lambda(t, t_0) = (\Lambda(t, t_0) \otimes I) |\phi^+\rangle\langle\phi^+|. \tag{6}$$

For single-qubit states, we can select the Bell state $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ and represent the Choi state on the Bell basis $\{|\phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), |\psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)\}$.

2.2 Markovian vs non-Markovian

A time-local master equation describes a Markovian process if and only if the canonical decay rates reported in Eq. 3 are positive at all times t . Conversely, if any of the coefficients are negative, we refer to the dynamics as non-Markovian.

In our study, we examine three types of processes, namely *semi-group*, *Markovian*, and *non-Markovian*, distinguished by the behavior of the decay rates $\gamma_k(t)$ in Eq. 3.

Markovianity is established when $\gamma_k(t) \geq 0 \forall k$ for all times t , suggesting a memoryless evolution. Here, Markovianity is equivalent to the property of CP-divisibility of the dynamical map (Rivas et al. 2014; Hall et al. 2014). However, the resulting set of dynamical maps $\Lambda(t)$ generally does not constitute a semi-group, i.e., they do not satisfy the semi-group property $\Lambda(t + s) = \Lambda(t) \circ \Lambda(s)$ for all $t, s \geq 0$. Specifically, we refer to a *semi-group* type for dynamical maps when all coefficients $\gamma_k(t)$ are constant and positive,

and the operators $H_S(t)$ and $A_k(t)$ are time-independent, representing a particular instance of Markovianity (Lidar et al. 2001). Finally, if one of the coefficients $\gamma_k(t)$ in Eq. 3 is negative for certain times t , non-Markovianity arises.

3 Quantum processes

In this section, we provide an extensive overview of the various quantum processes under investigation. These encompass the dephasing and the Pauli channels in arbitrary basis and the generalized amplitude damping channel. Our exploration includes a detailed examination of each channel, showing in particular the relation between the decay rates of the canonical master equation and the corresponding parameters characterizing the quantum channels.

First, we exclusively consider pure dissipative evolutions, thus disregarding the Hamiltonian $H_S(t)$ in Eq. 3. While this is not strictly necessary for the methodology, focusing on pure dissipative processes simplifies the implementation and is particularly relevant to our context. Including the coherent part of the process (i.e., the Hamiltonian) would still allow the method to remain valid. Second, we restrict our analysis to channels where all Lindblad operators $A_k(t)$ are time-independent, facilitating the analytical derivation of the Choi states by simplifying the resolution of the master equations. If the Lindblad operators were time-dependent, the semi-group class would no longer be applicable, as it requires constant decoherence coefficients and time-independent Lindblad operators. However, the methodology could be extended to consider time-dependent operators by focusing on the classification of Markovian versus non-Markovian processes.

In the upcoming paragraphs, we will utilize the notation $\mathcal{L}_A[\cdot] := A \cdot A^\dagger - \frac{1}{2}\{A^\dagger A, \cdot\}$.

3.1 Dephasing channel in a rotated basis

We begin by formulating the master equation that governs the dynamics of a dephasing channel:

$$\dot{\rho}(t) = \frac{\gamma(t)}{2} \mathcal{L}_{\sigma_z}[\rho(t)], \tag{7}$$

where $\gamma(t)$ denotes the decay rate and $\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|$ is the Pauli matrix associated with the z -axis.

The solution to this differential equation is

$$\rho(t) = \frac{1 + e^{-\Gamma(t)}}{2} \rho(0) + \frac{1 - e^{-\Gamma(t)}}{2} \sigma_z \rho(0) \sigma_z \tag{8}$$

with $\Gamma(t) = \int_0^t \gamma(s) ds$.

The Kraus operators characterizing the dephasing channel are then given by the following:

$$K_1(t) = \sqrt{\frac{1 + e^{-\Gamma(t)}}{2}} I, \quad K_2(t) = \sqrt{\frac{1 - e^{-\Gamma(t)}}{2}} \sigma_z, \tag{9}$$

where $I = |0\rangle\langle 0| + |1\rangle\langle 1|$. The Choi state of the dephasing (\mathcal{D}) channel is obtained by the evolution of the maximally entangled state $|\phi^+\rangle$ and can be represented in terms of the Bell basis as follows:

$$\mathcal{C}_{\mathcal{D}}(t) = \frac{1 + e^{-\Gamma(t)}}{2} |\phi^+\rangle\langle\phi^+| + \frac{1 - e^{-\Gamma(t)}}{2} |\phi^-\rangle\langle\phi^-|. \tag{10}$$

We can extend this analysis to a dephasing channel with respect to an arbitrary basis, by considering a master equation of the form:

$$\dot{\rho}(t) = \frac{\gamma(t)}{2} \mathcal{L}_{\sigma \cdot \mathbf{n}}[\rho(t)], \tag{11}$$

with $\sigma = (\sigma_x, \sigma_y, \sigma_z)$ being the vector of the three Pauli matrices and $\mathbf{n} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$ being a vector on the unit sphere, with $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi)$.

This scenario is more intricate compared to the standard dephasing channel. While the dynamics of the dephasing process can be effectively analyzed and classified using a feedforward neural network, addressing dephasing in a rotated basis calls for a more advanced architecture, such as a recurrent neural network.

The Choi state for the dephasing channel in a rotated basis (\mathcal{D}_{RB}) is given by the following:

$$\begin{aligned} \mathcal{C}_{\text{DRB}}(t) = & \frac{1 + e^{-\Gamma(t)}}{2} |\phi^+\rangle\langle\phi^+| + \frac{1 - e^{-\Gamma(t)}}{2} \{ \cos^2 \theta |\phi^-\rangle\langle\phi^-| + \sin^2 \theta \cos^2 \varphi |\psi^+\rangle\langle\psi^+| + \sin^2 \theta \sin^2 \varphi |\psi^-\rangle\langle\psi^-| \\ & + \sin \theta \cos \theta \cos \varphi (|\phi^-\rangle\langle\psi^+| + |\psi^+\rangle\langle\phi^-|) + i \sin \theta \cos \theta \sin \varphi (|\phi^-\rangle\langle\psi^-| - |\psi^-\rangle\langle\phi^-|) \\ & + i \sin^2 \theta \sin \varphi \cos \varphi (|\psi^+\rangle\langle\psi^-| - |\psi^-\rangle\langle\psi^+|) \}. \end{aligned} \tag{12}$$

3.2 Pauli channel

The dynamics of a Pauli channel is described by the master equation:

$$\dot{\rho}(t) = \frac{1}{2} \sum_{k=1}^3 \gamma_k(t) \mathcal{L}_{\sigma_k}[\rho(t)], \tag{13}$$

where $\sigma_k \in \{\sigma_1, \sigma_2, \sigma_3\}$ represent the Pauli matrices, and $\gamma_k(t)$ are time-dependent coefficients. The generators $\mathcal{L}_{\sigma_k}(\cdot) = \sigma_k \cdot \sigma_k - \cdot$ satisfy the commutation relation $[\mathcal{L}_{\sigma_i}, \mathcal{L}_{\sigma_j}] = 0 \forall i, j = 1, 2, 3$ (Chruściński and Siudzińska 2016). This property allows for the solution to be derived by factorizing the solutions for the dephasing channel (8) for all Pauli matrices.

The Choi state for the Pauli channel (P) is given by the following:

$$\begin{aligned} C_P(t) = & \frac{1}{4} \{ (1 + e^{-\Gamma_1(t)-\Gamma_2(t)} + e^{-\Gamma_1(t)-\Gamma_3(t)} + e^{-\Gamma_2(t)-\Gamma_3(t)}) |\phi^+\rangle\langle\phi^+| \\ & + (1 + e^{-\Gamma_1(t)-\Gamma_2(t)} - e^{-\Gamma_1(t)-\Gamma_3(t)} - e^{-\Gamma_2(t)-\Gamma_3(t)}) |\phi^-\rangle\langle\phi^-| \\ & + (1 - e^{-\Gamma_1(t)-\Gamma_2(t)} - e^{-\Gamma_1(t)-\Gamma_3(t)} + e^{-\Gamma_2(t)-\Gamma_3(t)}) |\psi^+\rangle\langle\psi^+| \\ & + (1 - e^{-\Gamma_1(t)-\Gamma_2(t)} + e^{-\Gamma_1(t)-\Gamma_3(t)} - e^{-\Gamma_2(t)-\Gamma_3(t)}) |\psi^-\rangle\langle\psi^-| \}, \end{aligned} \tag{14}$$

where $\Gamma_k(t) = \int_0^t \gamma_k(s) ds$.

Rotated basis transformation The Pauli channel can be generalized by considering a rotated basis, akin to the treatment applied to the dephasing channel. In this scenario, the modification of the Choi state presented in Eq. 14 is achieved by accounting for the rotation of the reference system. This is done by considering the total rotation $R(\alpha, \beta, \gamma) = R_z(\gamma)R_y(\beta)R_x(\alpha)$, where α, β , and γ are the Euler angles, and $R_i(\psi)$ represents a rotation around the i -axis by an angle ψ , with $i \in \{x, y, z\}$.

3.3 Generalized amplitude damping channel

The generalized amplitude damping evolution is characterized by the following master equation:

$$\dot{\rho}(t) = (\gamma_1(t)\mathcal{L}_{\sigma_-} + \gamma_2(t)\mathcal{L}_{\sigma_+})[\rho(t)]. \tag{15}$$

Here, $\sigma_+ = |1\rangle\langle 0|$ and $\sigma_- = |0\rangle\langle 1|$ denote the raising and lowering operators.

The corresponding channel can be written as $\rho(t) = \sum_{i=0}^3 E_i \rho(0) E_i^\dagger$, where

$$\begin{aligned} E_0 &= \sqrt{p} \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\lambda} \end{pmatrix} & E_2 &= \sqrt{1-p} \begin{pmatrix} \sqrt{1-\lambda} & 0 \\ 0 & 1 \end{pmatrix} \\ E_1 &= \sqrt{p} \begin{pmatrix} 0 & \sqrt{\lambda} \\ 0 & 0 \end{pmatrix} & E_3 &= \sqrt{1-p} \begin{pmatrix} 0 & 0 \\ \sqrt{\lambda} & 0 \end{pmatrix}, \end{aligned} \tag{16}$$

with $p, \lambda \in [0, 1]$. Here, $p + q = 1$, and p and λ are time-dependent functions which are related to the canonical decay rates as follows (Shrikant et al. 2022):

$$\gamma_1(t) = \lambda \dot{p} + \frac{\dot{\lambda} p}{1-\lambda} \quad \gamma_2(t) = \lambda \dot{q} + \frac{\dot{\lambda} q}{1-\lambda}. \tag{17}$$

The Choi state characterizing the generalized amplitude damping (GAD) channel takes the form:

$$\begin{aligned} C_{GAD}(t) = & \frac{(1 + \sqrt{1-\lambda})^2}{4} |\phi^+\rangle\langle\phi^+| + \frac{(1 - \sqrt{1-\lambda})^2}{4} |\phi^-\rangle\langle\phi^-| \\ & + \frac{\lambda}{4} (|\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-| + |\phi^+\rangle\langle\phi^-| \\ & + |\phi^-\rangle\langle\phi^+| + i(1-2p)|\psi^+\rangle\langle\psi^-| - i(1-2p)|\psi^-\rangle\langle\psi^+|). \end{aligned} \tag{18}$$

We observe that the diagonal components of the Choi state are unaffected by p . Hence, our classification between Markovian and non-Markovian behavior relies solely on the damping parameter λ .

4 Methods

In this section, we rigorously define the classification problem under consideration—the differentiation between semi-group, Markovian, and non-Markovian dynamics. Our approach involves the analysis of the evolution of the Choi states across various single-qubit channels.

We first start by introducing key concepts related to time series data and recurrent neural networks (RNNs). Then, we discuss the neural network methodology applied in this work, including details about the optimization procedure for our classification problems and the construction of the dataset.

4.1 Time series and recurrent neural networks

Time series data play a fundamental role in various disciplines, allowing to understand and predict dynamic processes in science, economics, and various other fields. A time series is a collection of observations or measurements ordered chronologically, often at equally spaced intervals. This temporal structure enables the analysis of patterns, trends, and fluctuations over time, providing valuable insights into the dynamics of diverse phenomena (Shumway et al. 2000).

A time series is defined by a sequence of observations x_t , each corresponding to a specific time point t . In the context of discrete time, which is the focus of our investigation, a discrete-time time series is characterized by observations made at distinct, discrete instances. Formally, this involves a set \mathcal{T} representing the discrete times at which

observations are recorded, typically occurring at fixed time intervals (Brockwell and Davis 2002).

Sequential data, like time series, can be effectively processed using recurrent neural networks (RNNs), a class of bidirectional artificial neural networks with recurrent connections. The recurrent connections create a form of memory, allowing to consider information from past time steps when processing the current input.

Two specific variants of RNNs that tackle challenges associated with long-term dependencies and vanishing gradients are long short-term memory (LSTM) and gated recurrent unit (GRU). Further details about these architectures can be found in Appendix 2.

4.2 RNNs for quantum processes classification

In this section, we clarify the application of RNNs in tackling the classification problem of discerning the Markovian nature of quantum processes. The analyzed dataset comprises time series consisting of selected components of Choi states of the quantum process recorded at different time steps. Each entry is associated with a corresponding label indicating whether the process exhibits semi-group, Markovian, or non-Markovian behavior.

The dataset is structured as follows:

$$\mathcal{D} = (\mathcal{X}^{(i)}, y^{(i)})_{i=1}^N, \quad (19)$$

where $\mathcal{X}^{(i)} = [\mathcal{C}^{(i)}(t_0), \dots, \mathcal{C}^{(i)}(T)]$ represents the i -th sample, corresponding to a time series of components of the Choi state with time steps $t \in [t_0, T]$. Additionally, $y^{(i)} \in \{0, 1, 2\}$ denotes the correct class label, where “0” indicates semi-group samples, “1” signifies Markovian samples, and “2” represents non-Markovian samples.

In our simulations, the time series comprises 7 equally spaced time steps ranging from 0 to $T = 3$ s. The dataset comprises 7200 training samples, 900 validation samples, and 900 test samples.

Through the utilization of a series of recurrent and feedforward layers applied to the time series samples, the performance of the network is assessed using accuracy as the figure of merit. Accuracy is defined as the ratio of correct predictions to the total number of predictions. The categorical cross-entropy, used as the loss function, is formulated as follows:

$$\text{CCE} := - \sum_{i=1}^N \sum_{j=1}^3 y_{ij} \log(p_{ij}). \quad (20)$$

Here, N denotes the number of samples to be classified, while 3 represents the number of distinct classes. The true class label $y^{(i)}$ of the i -th sample is one-hot encoded in a

probability distribution $y_i \in \{0, 1\}^3$, where one-hot encoding is a method of representing categorical variables as binary vectors. In this encoding, the true class of the sample is represented by a vector with a 1 in the position of the correct class and 0s in all other positions.

The terms y_{ij} and p_{ij} , both constrained within the range $[0, 1]$, respectively, stand for the actual and model-predicted probabilities for the i -th sample to belong to the j -th class. The negative logarithm of the predicted probability is summed over all samples and classes to compute the categorical cross-entropy loss.

In the next section, we analyze the network efficiency in distinguishing among the classes.

5 Results

We now present the results obtained by using RNNs for time series classification and forecasting. We start discussing the performance of our model in classifying whether a process is a semi-group, Markovian, or non-Markovian. This evaluation is done for dephasing and Pauli channels in arbitrary bases and generalized amplitude damping channels. Then, we show also the network capability in predicting the evolution of several components of the Choi state under both dephasing and Pauli dynamics. Our results demonstrate an optimal accuracy in classification and efficient forecasting capabilities for the analyzed time series data.

Additional information regarding the data generation process is available in Appendix 3.

All simulations are run with TensorFlow (TensorFlow Developers 2023), with the Adam optimizer (Kingma 2014) employed.

5.1 Channel classification

In this section, we present our findings on classifying channels based on the Markovianity of quantum processes. Our model architectures have demonstrated remarkable performance in achieving robust classification accuracy for quantum processes. The neural network architecture employed consists of a GRU with 32 units, augmented by a feedforward neural network comprising 16 neurons, with rectified linear unit (ReLU) as the activation function.¹ The architecture culminates with a softmax layer comprising three neurons.

For these simulations, the learning rate is set to 0.001, the number of epochs to 800, and the batch size to 720. The learning processes exhibit noise; however, stability can be

¹ In this specific setup, feedforward neural networks (FFNNs) achieve comparable accuracies to RNNs. However, RNNs could provide advantages in capturing sequential dependencies and are expected to outperform FFNNs in tasks involving more complex temporal structures, longer time series, or noisy data.

attained by enlarging the dataset (approximately 24,000 samples may suffice) and reducing the learning rate. Importantly, our model demonstrates effectiveness even with a limited number of samples.

Notably, both GRU and LSTM architectures are found to be interchangeable. All the tests are conducted exclusively using the diagonal components within the Bell basis of the Choi states as input.

In Fig. 2, the resulting training and test accuracies of classification tasks involving the analyzed models in Sect. 3 for single-qubit channels—specifically, the dephasing channel in a rotated basis, the Pauli channel, the Pauli channel in a rotated basis and the generalized amplitude damping channel—are illustrated. We also examine the scenario where all channels are included in the training set.

The results presented correspond to a single representative run for each case. The findings have been verified to

be robust across multiple independent runs with varying random initializations of the RNN parameters. In particular, both the accuracy and the convergence behavior of the loss function remain consistent, highlighting the reliability of the approach. We also demonstrate the model's performance when utilizing all components of the Choi matrix, achieving comparable results (see Fig. 2b). Additionally, we analyze two distinct noise scenarios: the first involves Choi states prepared with a fixed fidelity of 0.95 across all samples, while the second presents a more challenging scenario where the fidelity of the Choi states is set to 0.5 for each sample. These scenarios are illustrated in Fig. 2 c and d, respectively. See Appendix 4 for more details about the noise employed. In particular, when $F = 0.25$, the network predictions reduce to random guessing. This is expected, as $F = 0.25$ corresponds to a maximally mixed state, meaning that all information about the channel is lost.

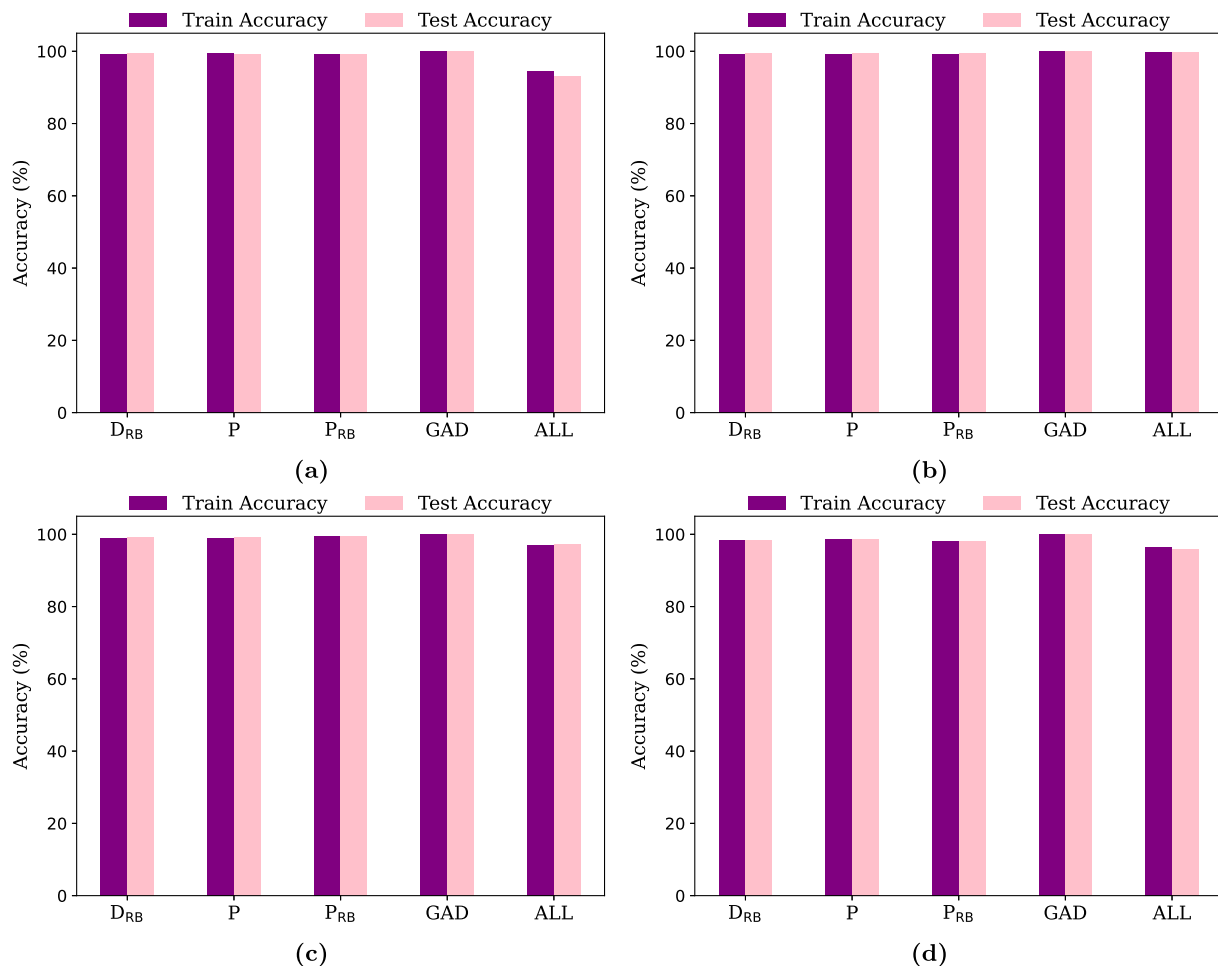


Fig. 2 Train and test accuracies are depicted for various processes in a classification problem: **a** when only the diagonal components of the Choi matrix are used, **b** when all the non-null and unique components of the Choi matrix are used, **c** when the Choi states are prepared with a fixed fidelity $F = 0.95$ across all samples, and **d** when the Choi states

are prepared with fixed fidelity $F = 0.5$. These include dephasing in a rotated basis (D_{RB}), Pauli (P), Pauli in a rotated basis (P_{RB}), generalized amplitude damping (GAD) channels, and all the aforementioned channels collectively labeled as ALL

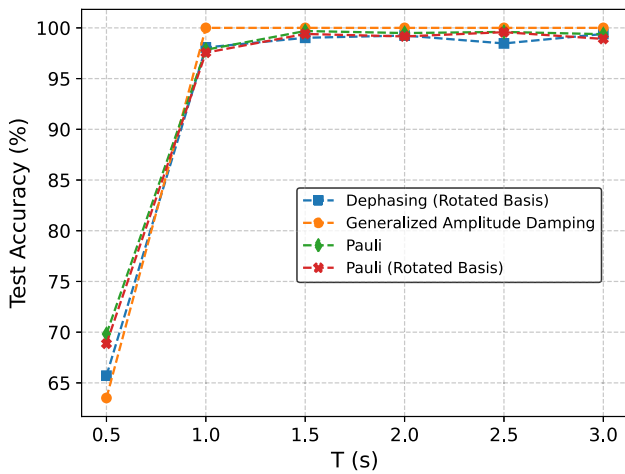


Fig. 3 Correlation between classification accuracy and the length of the time series for dephasing, generalized amplitude damping, Pauli, and Pauli in rotated basis channels, with a time step $\Delta t = 0.5$. The x-axis represents the final time step T of the time series, corresponding to different time series lengths. Each reported result is an average across five distinct runs

Notably, all models examined achieve a test accuracy of at least 95%, indicating no overfitting, i.e., the overspecialization of the model on the training samples.

Time series length We examine how varying time series lengths affect the classification accuracy of our model across all single-qubit channels previously analyzed. The aim is to explore the model sensitivity to the temporal dimension of input data. Remarkably, our findings reveal that, even with a brief time series consisting of only three time steps, our model achieves accuracies exceeding 90%, as illustrated in Fig. 3. Each reported result represents an average across five distinct runs. Observed fluctuations may stem from variations in the training process. The consistency of the hyperparameters of the network (epochs = 800, batch size = 600, learn-

ing rate = 0.001), along with a fixed training cardinality of 6000 samples across all simulations, suggests that a specific combination of these hyperparameters may be more suitable for a particular time series length. These hyperparameters, however, are specific to this simulation and tailored for this particular task.

Generalization properties We evaluate the network’s ability to generalize, specifically its capability to correctly classify processes that were not included in the training set. We conduct four experiments to assess this: (a) training with the standard dephasing channel D and testing on D_{RB} ; (b) training with D and testing on P_{RB} ; (c) training with P and testing on GAD; (d) training with P and testing on P_{RB} .

We present the results of these generalization tests using confusion matrices. A confusion matrix is a table used to describe the performance of a classification model, showing the actual versus predicted classifications. The confusion matrices for these experiments are depicted in Fig. 4.

Notably, in Fig. 4a, the dephasing channels do not achieve satisfactory accuracies for correctly classifying dephasing channels in a rotated basis. This is likely due to the limited information conveyed by the standard dephasing channel. However, in other scenarios, the performance improves significantly. For example, in Fig. 4d, the network demonstrates the ability to correctly classify almost all Pauli channels in a rotated basis when trained with standard Pauli channels, indicating that the classification performance is largely independent of the chosen basis.

These results exhibit slight variability depending on individual runs, which can lead to better or worse outcomes in specific instances. However, in Fig. 4c, the results are highly dependent on individual runs. This dependency arises when the network becomes overly specialized to the specific functions composing the decay rates, significantly impacting performance.

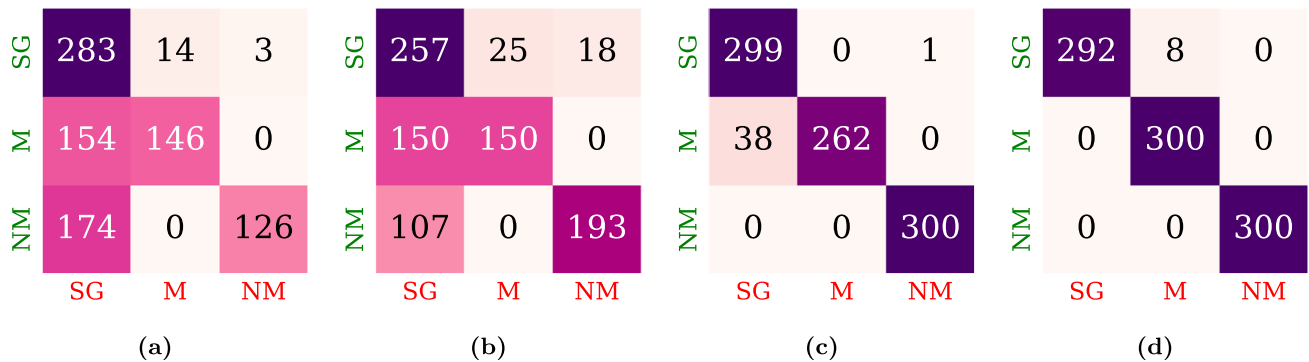


Fig. 4 Classification tasks involving channels not included in the training set. The considered channels are dephasing (D), dephasing in a rotated basis (D_{RB}), Pauli (P), Pauli in a rotated basis (P_{RB}), and generalized amplitude damping (GAD). **a**, **b**, **c**, and **d** depict the performance when using D as train and D_{RB} as test (accuracy 62%), D as train and

P_{RB} as test (accuracy 66%), P as train and GAD as test (accuracy 96%), and P as train and P_{RB} as test (accuracy 99%) channels, respectively. The green labels on the vertical axis represent the actual labels, while the red labels on the horizontal axis represent the predicted labels

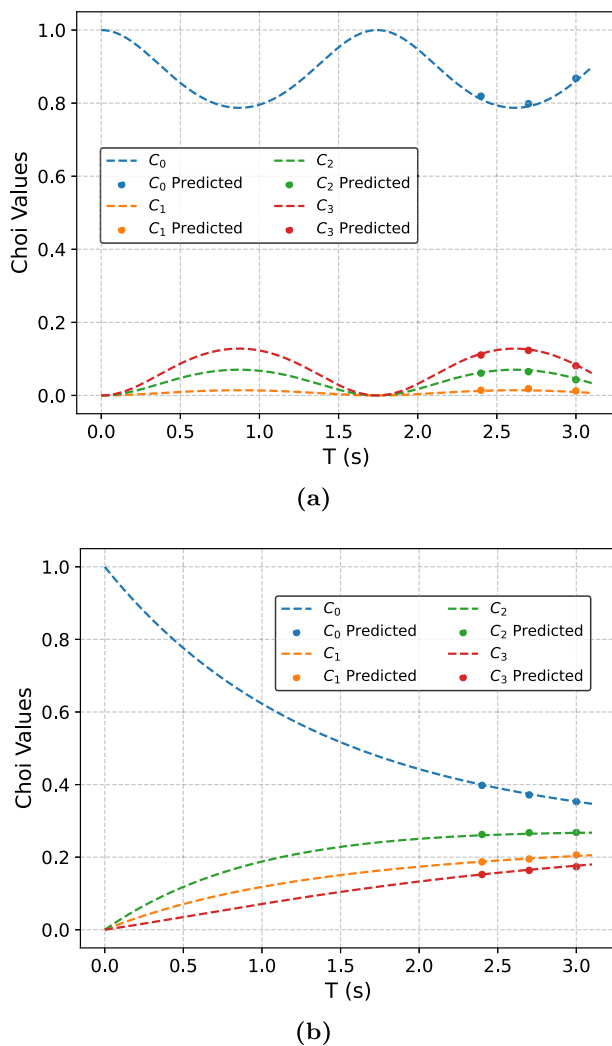


Fig. 5 Time series forecasting for the four diagonal elements C_i of the Choi state in the Bell basis, for a specific test sample exposed to **a** a dephasing channel and **b** a Pauli channel in an arbitrary basis. The plot displays the actual values of C_i with dashed lines, while the predicted points (C_i Predicted) are shown as markers

The results remain consistent when the full Choi matrix is considered. Additionally, to facilitate simulations, each decay rate of every process is characterized by specific functions, differing only in the values of their coefficients, randomly sampled from predefined intervals. This approach may introduce bias when attempting to generalize to other processes.

All simulations were conducted using the same network architecture and hyperparameters as those employed in the simulations shown in Fig. 2.

5.2 Forecasting channel evolution

Forecasting in the context of time series refers to the process of predicting future values or trends based on past observa-

tions. In this section, we focus on predicting the four diagonal components of the Choi state in the Bell basis, corresponding to either a dephasing channel in a rotated basis or a Pauli channel in a rotated basis.

The RNN is trained on historical data, and its performance is assessed through the accurate prediction of future time steps. To provide a comprehensive illustration, we present the forecasting results for two representative samples: one for dephasing and one for Pauli. In each case, eight initial time steps are given, and the model is tasked with predicting the subsequent three steps.

The evaluation of our model demonstrates a remarkably low mean squared error (MSE) of 10^{-5} , indicative of the model's ability to capture the dynamics of the quantum process.

The MSE is computed as $MSE = \frac{1}{n} \sum_{i=1}^n \|\mathbf{C}^{(i)} - \hat{\mathbf{C}}^{(i)}\|^2$, where n is the number of samples, $\mathbf{C}^{(i)}$ is a vector containing the true values of the Choi state components in the prediction time window, and $\hat{\mathbf{C}}^{(i)}$ represents the vector of predicted values of the Choi state components.

The architecture comprises a GRU layer with 64 units, followed by a feedforward layer with 64 units using ReLU as an activation function, and a final layer with 12 units. From Fig. 5, we can see that the prediction matches the actual time series.

6 Conclusion

This study introduces a novel methodology employing recurrent neural networks (RNNs) for the classification and forecasting of Markovian and non-Markovian quantum processes based on time series data derived from Choi states. The model demonstrates exceptional accuracy, surpassing 95% in diverse scenarios involving dephasing, Pauli, generalized amplitude damping channels, and even in the presence of noise. The analysis extends to forecasting the evolution of Choi states, revealing the model's robustness in predicting future time steps with low MSE. The utilization of only the diagonal components within the Bell basis proves effective, highlighting the efficiency of the proposed approach.

The findings hold implications for quantum information theory, quantum control, and quantum thermodynamics, where distinguishing between Markovian and non-Markovian dynamics is of paramount importance.

For future research, exploration into alternative types of data representation beyond Choi states could be pursued. For instance, investigating the feasibility of utilizing Pauli transfer matrices as time series data may provide valuable insights into the dynamics of quantum processes from a different perspective.

Furthermore, future research could explore higher-dimensional systems and examine the interplay between neural

networks and quantum dynamics to tackle the challenges presented by more complex quantum systems and varied scenarios. For instance, upcoming work will focus on designing and utilizing real-world datasets, enabling us to rigorously evaluate the algorithms in practical settings, thereby further validating their performance and applicability in experimental contexts.

This exploration holds the potential to yield valuable insights into the resilience and adaptability of machine learning techniques within practical quantum computing environments.

Appendix 1. Master equation from quantum channel

In the pursuit of characterizing the open dynamics of quantum systems, a fundamental aspect lies in the study of the master equation governing the evolution of the system density matrix.

Our starting point is the master equation expressed as $\dot{\rho}(t) = \mathcal{L}_t[\rho(t)]$, where \mathcal{L}_t is the Liouvillian superoperator. A corresponding set of quantum channels is generally written as $\rho(t) = \mathcal{N}_t(\rho(0)) = \sum_i E_i(t)\rho(0)E_i^\dagger(t)$.

Assuming the invertibility of the map \mathcal{N}_t , one can write $\mathcal{L}_t[\cdot] = \dot{\mathcal{N}}_t\mathcal{N}_t^{-1}[\cdot]$ (Andersson et al. 2007). The operators \mathcal{N}_t and \mathcal{N}_t^{-1} can be computed starting from an alternative representation of the channel \mathcal{N}_t as $K_t = \sum_i E_i(t) \otimes \bar{E}_i(t)$ (Watrous 2018), where $\{E_i(t)\}$ are the Kraus operators. In this way, given a quantum state ρ , we can derive the Lindblad operator as $\dot{K}_t K_t^{-1}(\vec{\rho})$, where $\vec{\rho}$ is the vectorization of the density matrix ρ (Gilchrist et al. 2009).

Appendix 2. Recurrent neural networks

LSTM LSTM consists of a memory cell and a set of gates to regulate information flow (Wang 2017). These gates include the forget gate f_t , responsible for deciding which information from the cell state c_t to discard; the input gate i_t , determining the information to add to the cell state; and the output gate o_t , controlling the information to transmit to the output h_t . The equations governing these operations are

$$\begin{aligned} f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \\ i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \\ o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \\ \tilde{c}_t &= \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \\ c_t &= f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t, \\ h_t &= o_t \cdot \tanh(c_t). \end{aligned} \tag{B1}$$

Here, $\sigma(\cdot)$ represents the sigmoid function, and W_i and b_i , with $i \in \{f, i, o, c\}$, denote the weights and biases specific to each gate. The input is represented by x_t while h_t is the so-called hidden state, that acts as the memory of the network. The operation $[\cdot, \cdot]$ denotes vector concatenation.

GRU Being a simplified variant of LSTM, GRU merges the memory cell and the hidden state. Its fundamental components include the update gate z_t , responsible for governing the extent to which the past hidden state is retained, the reset gate r_t , determining the degree to which the past hidden state is overlooked, and the current memory content h_t , reflecting the updated hidden state. The main gates are (Dey and Salem 2017) :

$$\begin{aligned} h_t &= (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t, \\ \tilde{h}_t &= \tanh(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h), \\ z_t &= \sigma(W_z x_t + U_z h_{t-1} + b_z), \\ r_t &= \sigma(W_r x_t + U_r h_{t-1} + b_r). \end{aligned} \tag{B2}$$

Here, W_i, U_i , and b_i with $i \in \{h, z, r\}$ denote the specific parameters for each gate, and \odot is the Hadamard product.

Appendix 3. Data generation

In this section, we outline the procedure for generating our dataset utilized in training the network. We emphasize that the functions chosen to model the decay rates of the master equation were selected to ensure that the integrated maps are completely positive (CP) at all times.

Dephasing channel Beginning with Eq. 12, we observe that the parameters governing the generation of diverse instances of a dephasing channel in a rotated basis are characterized by the pair (θ, φ) and by $\Gamma(t) = \int_0^t \gamma(s)ds$. Here, $\cos \theta \in [-1, 1]$ and $\varphi \in [0, 2\pi)$ are randomly sampled from a uniform distribution for each individual sample.

To generate various semi-group instances, we sample $\gamma(t) = a$ for each process, where $a \in [0, 1)$. For the generation of Markovian processes, we choose $\gamma(t) = a \tanh t$ or $\gamma(t) = \sin^2(bt)$, with $b \in [2, 4)$. Conversely, for non-Markovian samples, we employ $\gamma(t) = \sin bt$.

Pauli channel The Choi state in Eq. 14 relies on the cumulative functions $\Gamma_k(t) = \int_0^t \gamma_k(s)ds$. To generate semi-group samples, we set $\gamma_k(t) = a_k$, where $a_k \in [0, 1)$. For Markovian samples, we adopt a strategy where $\gamma_k(t) = \sin^2(b_k t)$ or $\gamma_k(t) = a_k \tanh t$, with $b_k \in [2, 4)$ and $a_k \in [0, 1)$. For non-Markovian samples, we use the functions $\gamma_k(t) = \sin(b_k t)$ or consider an alternative scenario characterized by $\gamma_i = \gamma_j = a$ and $\gamma_k(t) = -a \tanh t$, where $a \geq 1$ and i, j, k represent any permutation of the indices 1, 2 and 3.

For samples generated from the Pauli channel in a rotated basis, we assign a unique set of three Euler angles (α, β, γ) to each sample.

GAD channel The Choi state described in Eq. 18 relies on the parameters of the GAD channel, namely p and λ . To distinguish between different classes, we adopt different parameterizations for the decay rates reported in Eq. 18, which are linked to the channel parameters via Eq. 17. For semi-group samples, the solution of the differential equations with constant γ_1 and γ_2 yields $p = \frac{\gamma_1}{\gamma_1 + \gamma_2}$ and $\lambda = 1 - e^{-(\gamma_1 + \gamma_2)t}$. In the case of Markovian samples, p remains constant in $[0, 1]$ while $\lambda(t) = \tanh(bt)$ with $b \in [\frac{1}{2}, 1)$. Finally, for non-Markovian samples, $p \in [0, 1]$ and $\lambda(t) = \frac{1}{2} \sin^2(bt)$ with $b \in [2, 4)$.

Appendix 4. Noise model

In this work, we consider a noise model inspired by the theoretical framework outlined in Macchiavello and Sacchi (2018) and the experimental scenario adopted in Cuevas et al. (2017) and Ciampini et al. (2021). We focus on the case where the dominant source of noise arises from imperfections of the initial Bell state that is used at the input of the quantum channels in order to obtain the time series of the Choi states at the output.

Typically, the ideal input Bell state $|\phi^+\rangle\langle\phi^+|$ is replaced by a noisy Werner state with fidelity $F < 1$, namely

$$\nu_F = F |\phi^+\rangle\langle\phi^+| + \frac{1-F}{3} \sum_{i=1}^3 |\phi_i\rangle\langle\phi_i|, \quad (\text{D1})$$

where $|\phi_1\rangle = |\phi^-\rangle$, $|\phi_2\rangle = |\psi^+\rangle$ and $|\phi_3\rangle = |\psi^-\rangle$ are the remaining Bell states, and $F = \langle\phi^+|\nu_F|\phi^+\rangle$.

Then, in this noisy scenario, the ideal Choi states of Eq. 6 are replaced by

$$\mathcal{C}_\Lambda^F(t, t_0) = (\Lambda(t, t_0) \otimes I) \nu_F. \quad (\text{D2})$$

Correspondingly, the diagonal elements $r_j = \langle\phi_j|\mathcal{C}_\Lambda(t, t_0)|\phi_j\rangle$ of the Choi state in the Bell basis for ideal input are replaced by

$$s_j = \langle\phi_j|\mathcal{C}_\Lambda^F(t, t_0)|\phi_j\rangle = F r_j + \frac{1-F}{3} \sum_{i \neq j} r_i. \quad (\text{D3})$$

The noise model captures a realistic experimental setting where the fidelity of the input state of the channels is less than 1, due to imperfections in the preparation of the Bell state.

We remark that in Cuevas et al. (2017) and Ciampini et al. (2021), the fidelity was experimentally estimated as $F = 0.979 \pm 0.011$.

Author contribution All authors conceived the idea and participated in defining the problem. A.R.M. conducted the simulations and developed the code, under the supervision of M.F.S. and C.M. A.R.M. primarily wrote the manuscript, with multiple critical revisions suggested by M.F.S. and C.M. All authors reviewed and approved the final version of the manuscript for submission.

Funding Open access funding provided by Università degli Studi di Pavia within the CRUI-CARE Agreement. A.R.M. acknowledges support from the PNRR MUR Project PE0000023-NQSTI. M.F.S. acknowledges support from the PRIN MUR Project 2022SW3RPY. C.M. acknowledges support from the National Research Centre for HPC, Big Data and Quantum Computing, PNRR MUR Project CN0000013-ICSC.

Data availability The data and code that support the findings of this study can be given on demand.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adlam E (2020) Entropy 22:1063. <https://doi.org/10.3390/e22091063>
- Andersson E, Cresser JD, Hall MJ (2007) J Mod Opt 54:1695. <https://doi.org/10.1080/09500340701352581>
- Banchi L, Grant E, Rocchetto A, Severini S (2018) New J Phys 20:123030. <https://doi.org/10.1088/1367-2630/aaf749>
- Banerjee S, Banerjee S, Ahmad, (2018) Open quantum systems. Springer. <https://doi.org/10.1007/978-981-13-3182-4>
- Barr J, Zicari G, Ferraro A, Paternostro M (2023) Mach Learn. Sci Technol. <https://doi.org/10.1088/2632-2153/ad2cf1>
- Breuer H-P, Petruccione F (2007) The theory of open quantum systems. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199213900.001.0001>
- Brockwell PJ, Davis RA (2002) Introduction to time series and forecasting. Springer. <https://doi.org/10.1007/978-3-319-29854-2>
- Bylicka B, Chruściński D, Maniscalco S (2014) Sci Rep 4:5720. <https://doi.org/10.1038/srep05720>
- Chai J, Zeng H, Li A, Ngai EW (2021) Mach Learn Appl 6:100134. <https://doi.org/10.1016/j.mlwa.2021.100134>
- Cheong JW, Pradana A, Chew LY (2022) Phys Rev A 106:052410. <https://doi.org/10.1103/PhysRevA.106.052410>

- Choi M-D (1975) *Linear Algebra Appl* 10:285. [https://doi.org/10.1016/0024-3795\(75\)90075-0](https://doi.org/10.1016/0024-3795(75)90075-0)
- Choquehuanca J, de Paula F, Sarandy M (2023) *Phys Rev A* 107:012220. <https://doi.org/10.1088/1751-8121/aba0ee>
- Chruściński D, Siudzińska K (2016) *Phys Rev A* 94:022118. <https://doi.org/10.1103/PhysRevA.94.022118>
- Chruściński D, Macchiavello C, Maniscalco S (2017) *Phys Rev Lett* 118:080404. <https://doi.org/10.1103/PhysRevLett.118.080404>
- Ciampini MA, Cuevas Á, Mataloni P, Macchiavello C, Sacchi MF (2021) *Phys Rev A* 103:062414. <https://doi.org/10.1103/PhysRevA.103.062414>
- Cuevas Á, Proietti M, Ciampini MA, Duranti S, Mataloni P, Sacchi MF, Macchiavello C (2017) *Phys Rev Lett* 119:100502. <https://doi.org/10.1103/PhysRevLett.119.100502>
- Deng J, Lin Y (2022) *Front Comput Intell Syst* 2:81. <https://doi.org/10.54097/fcis.v2i2.4465>
- Dey R, Salem FM (2017) In: 2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS). IEEE, pp 1597–1600. <https://doi.org/10.1109/MWSCAS.2017.8053243>
- Erdman PA, Rolandi A, Abiuso P, Perarnau-Llobet M, Noé F (2023) *Phys Rev Res* 5:L022017. <https://doi.org/10.1103/PhysRevResearch.5.L022017>
- Fanchini FF, Karpat G, Rossatto DZ, Norambuena A, Coto R (2021) *Phys Rev A* 103:022425. <https://doi.org/10.1103/PhysRevA.103.022425>
- Gilchrist A, Terno DR, Wood CJ (2009) [arXiv:0911.2539](https://arxiv.org/abs/0911.2539)
- Gray J, Banchi L, Bayat A, Bose S (2018) *Phys Rev Lett* 121:150503. <https://doi.org/10.1103/PhysRevLett.121.150503>
- Hakoshima H, Matsuzaki Y, Endo S (2021) *Phys Rev A* 103:012611. <https://doi.org/10.1103/PhysRevA.103.012611>
- Hall MJ, Cresser JD, Li L, Andersson E (2014) *Phys Rev A* 89:042120. <https://doi.org/10.1103/PhysRevA.89.042120>
- Harney C, Paternostro M, Pirandola S (2021) *New J Phys* 23:063033. <https://doi.org/10.1088/1367-2630/ac0388>
- Huang W-M, Zhang W-M (2022) *Phys Rev A* 106:032607
- Jamiołkowski A (1972) *Rep Math Phys* 3:275. [https://doi.org/10.1016/0034-4877\(72\)90011-0](https://doi.org/10.1016/0034-4877(72)90011-0)
- Khait I, Carrasquilla J, Segal D (2022) *Phys Rev Res* 4:L012029. <https://doi.org/10.1103/PhysRevResearch.4.L012029>
- Khurana D, Koli A, Khatter K, Singh S (2023) *Multimed Tools Appl* 82:3713. <https://doi.org/10.1007/s11042-022-13428-4>
- Kim C, Park KD, Rhee J-K (2020) *IEEE Access* 8:188853. <https://doi.org/10.1109/ACCESS.2020.3031607>
- Kingma DP (2014) arXiv preprint [arXiv:1412.6980](https://arxiv.org/abs/1412.6980). <https://doi.org/10.48550/arXiv.1412.6980>
- Koutný D, Ginés L, Moczala-Dusanowska M, Höfling S, Schneider C, Predojević A, Ježek M (2023) *Sci Adv* 9:eadd7131. <https://doi.org/10.1126/sciadv.add7131>
- Li L, Hall MJ, Wiseman HM (2018) *Phys Rep* 759:1. <https://doi.org/10.1016/j.physrep.2018.07.001>
- Lidar DA, Bihary Z, Whaley KB (2001) *Chem Phys* 268:35. [https://doi.org/10.1016/S0301-0104\(01\)00330-5](https://doi.org/10.1016/S0301-0104(01)00330-5)
- Lindblad G (1976) *Commun Math Phys* 48:119. <https://doi.org/10.1007/BF01608499>
- Liu J, Lu X-M, Wang X et al (2013) *Phys Rev A* 87:042103. <https://doi.org/10.1103/PhysRevA.87.042103>
- Macchiavello C, Sacchi MF (2018) *Phys Rev A* 97:012303. <https://doi.org/10.1103/PhysRevA.97.012303>
- Mallick B, Mukherjee S, Maity AG, Majumdar A (2024) *Phys Rev A* 109:022247. <https://doi.org/10.1103/PhysRevA.109.022247>
- Martina S, Gherardini S, Caruso F (2023) *Phys Scr* 98:035104. <https://doi.org/10.1088/1402-4896/acb39b>
- Mazza PP, Zietlow D, Carollo F, Andergassen S, Martius G, Lesanovsky I (2021) *Phys Rev Res* 3:023084. <https://doi.org/10.1103/PhysRevResearch.3.023084>
- Merkli M (2020) *Ann Phys* 412:167996. <https://doi.org/10.1016/j.aop.2019.167996>
- Merkli M (2022) *Quantum* 6:616. <https://doi.org/10.22331/q-2022-01-03-616>
- Min B, Ross H, Sulem E, Veysseh APB, Nguyen TH, Sainz O, Agirre E, Heintz I, Roth D (2023) *ACM Comput Surv* 56:1. <https://doi.org/10.1145/3605943>
- Mutti M, De Santi R, Restelli M (2022) In: Chaudhuri K, Jegelka S, Song L, Szepesvari C, Niu G, Sabato S (eds) *Proceedings of the 39th international conference on machine learning*. Proceedings of Machine Learning Research. PMLR, vol 162, pp 16223–16239. <https://proceedings.mlr.press/v162/mutti22a.html>
- Nielsen MA, Chuang IL (2010) *Quantum computation and quantum information*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511976667>
- Palmieri AM, Bianchi F, Paris MG, Benedetti C (2021) *Phys Rev A* 104:052412. <https://doi.org/10.1103/PhysRevA.104.052412>
- Reich DM, Katz N, Koch CP (2015) *Sci Rep* 5:12430. <https://doi.org/10.1038/srep12430>
- Rivas Á, Huelga SF, Plenio MB (2014) *Rep Prog Phys* 77:094001. <https://doi.org/10.1088/0034-4885/77/9/094001>
- Rivas A, Huelga SF (2018) *Open quantum systems*, vol 10. Springer. <https://doi.org/10.1007/978-3-642-23354-8>
- Salehinejad H, Sankar S, Barfett J, Colak E, Valaee S (2017) arXiv preprint [arXiv:1801.01078](https://arxiv.org/abs/1801.01078). <https://doi.org/10.48550/arXiv.1801.01078>
- Schlosshauer M (2019) *Phys Rep* 831:1. <https://doi.org/10.1016/j.physrep.2019.10.00>
- Sgroi P, Palma GM, Paternostro M (2021) *Phys Rev Lett* 126:020601. <https://doi.org/10.1103/PhysRevLett.126.020601>
- Shrikant U, Mandayam P (2023) *Front Quant Sci Technol* 2:1134583. <https://doi.org/10.3389/frqst.2023.1134583>
- Shrikant U, Srikanth R, Banerjee S (2022) arXiv preprint [arXiv:2203.10625](https://arxiv.org/abs/2203.10625)
- Shumway RH, Stoffer DS, Stoffer DS (2000) *Time series analysis and its applications*, vol 3. Springer. <https://link.springer.com/content/pdf/10.1007/978-3-319-52452-8.pdf>
- Takagi R, Endo S, Minagawa S, Gu M (2022) *npj Quant Inf* 8:114. <https://doi.org/10.1038/s41534-022-00618-z>
- Developers TensorFlow (2023). TensorFlow. <https://doi.org/10.5281/zenodo.10126399>
- Utagi S, Srikanth R, Banerjee S (2020) *Sci Rep* 10:15049. <https://doi.org/10.1038/s41598-020-72211-3>
- Wang Y (2017) In: 2017 American control conference (ACC). IEEE, pp 5324–5329. <https://doi.org/10.23919/ACC.2017.7963782>
- Watrous J (2018) *The theory of quantum information*. Cambridge University Press. <https://doi.org/10.1017/9781316848142>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.