npj | quantum information

Article

Published in partnership with The University of New South Wales

https://doi.org/10.1038/s41534-026-01236-9

Time-series forecasting with multiphoton quantum states and integrated photonics

Rosario Di Bartolo[1], Simone Piacentini[2,3], Francesco Ceccarelli[2], Giacomo Corrielli[2], Roberto Osellame[2], Valeria Cimini[1] ✉ & Fabio Sciarrino[1]

Quantum machine learning algorithms have very recently attracted significant attention in photonic platforms. In particular, reconfigurable integrated photonic circuits offer a promising route, thanks to the possibility of implementing adaptive feedback loops, which are an essential ingredient for achieving the necessary nonlinear behavior characteristic of neural networks. Here, we implement a quantum reservoir computing protocol in which information is processed through a reconfigurable linear optical integrated photonic circuit and measured using single-photon detectors. We exploit a multiphoton-based setup for time-series forecasting tasks in a variety of scenarios, where the input signal is encoded in one of the circuit's optical phases, thus modulating the quantum reservoir state. The resulting output probabilities are used to set the feedback phases and, at the end of the computation, are fed to a classical digital layer trained via linear regression to perform predictions. We then focus on the investigation of the role of input photon indistinguishability in the reservoir's capabilities of predicting time-series. We experimentally demonstrate that two-photon indistinguishable input states lead to significantly better performance compared to distinguishable ones. This enhancement arises from the quantum correlations present in indistinguishable states, which enable the system to approximate higher-order nonlinear functions when using comparable physical resources, highlighting the importance of quantum interference and indistinguishability as a resource in photonic quantum reservoir computing.

The recent surge of artificial intelligence, relying on huge neural network (NN) architectures that require high-speed and high-dimensional data processing, is driving renewed interest in alternative hardware platforms beyond conventional microelectronic processors. The core of the NN computation consists of two operations: (1) matrix-vector multiplication, which enables the propagation of information across the network layers, and (2) the application of nonlinear activation functions at each node. Scaling these operations to process increasingly high-dimensional data and complex tasks demands deeper and more expressive models[1]. Yet, traditional digital processors face two major challenges in meeting these demands. The first is computational: the time and memory requirements scale unfavorably with network depth and size, imposing practical limits on achievable performance. The second challenge is linked to the energy consumption and the associated carbon footprint of training and running such large-scale models, making standard processors unsustainable[2]. This is the reason why we are now at a point, often referred to as the electronic bottleneck, which constrains the scalability, speed, and energy efficiency of deep learning models.

In response to these challenges, one promising alternative is neuromorphic computing[3–5], a paradigm introduced in the 1990s by Carver Mead[6] that proposed to emulate the structure and function of the brain using analog circuits to enhance computation efficiency. Although it represented a visionary strategy at the time, it experienced a long period of limited progress after its proposal, largely due to the absence of compelling real-world applications and the lack of a clear incentive to deviate from digital computing, whose circuit components continued to grow following Moore's law. Today, the context has fundamentally changed, catalyzed by the urgent need for novel energy-efficient hardware solutions[7–9] capable of sustaining the rapid evolution of deep NN models.

Two properties are fundamental for these models: *expressivity*, consisting of the ability to approximate complex nonlinear functions, and *memory capacity* to retain and process information across multiple layers or time steps. These two ingredients are critical for solving a wide range of machine learning (ML) tasks, including time-series forecasting[10]. In this path, a new frontier is to explore the adoption of quantum systems to

[1]Dipartimento di Fisica, Sapienza Università di Roma, Roma, Italy. [2]Istituto di Fotonica e Nanotecnologie, Consiglio Nazionale delle Ricerche (IFN-CNR), Milano, Italy. [3]Present address: Quandela, Massy, France. ✉e-mail: valeria.cimini@uniroma1.it

develop more efficient neuromorphic architectures[11], which have access to an enlarged Hilbert space. This expanded state space leads to higher expressivity[12], allowing quantum systems to represent and manipulate complex data structures with fewer physical resources. At the same time, quantum coherence and entanglement can offer mechanisms for encoding correlations and memory in ways that are not readily achievable classically. However, this requires the challenging task of optimizing the quantum system, which presents several practical problems.

Quantum reservoir computing (QRC) offers a practical and promising alternative in the near/medium term[13–17] by restricting the training to a classical readout layer while keeping fixed the quantum system dynamics, thus avoiding the complexities of directly optimizing quantum parameters. This framework has already demonstrated strong potential for addressing nonlinear and temporal learning problems[18], thus holding significant promise for advancing time series forecasting, which is fundamental in fields ranging from finance and climate modeling to natural language analysis. Several physical platforms, including superconducting circuits[19,20] and Rydberg atom arrays[21,22], have already been employed to demonstrate the effectiveness of the QRC paradigm. In particular, photonics systems have been widely adopted lately for QRC, leveraging both continuous-variable squeezed states[23–26] and discrete-variable encodings[27–29]. Recent advances have shown that nonlinearity and memory can be introduced through adaptive feedback schemes[30], where quantum circuits are dynamically modified based on intermediate measurement outcomes, making integrated photonic a strong candidate for further investigation.

In this work, we demonstrate the implementation of a quantum reservoir architecture based on a photonic integrated circuit (PIC) injected with different optical input states. While previous feedback-driven approaches based on integrated circuits have been restricted to the single-photon regime[31], here we investigate time-series forecasting with multiphoton inputs. We focus on assessing the role of quantum correlation in the system's computational capabilities by comparing the performance achieved when the system involves indistinguishable two-photon states and distinguishable photon inputs under otherwise identical conditions. Specifically, we fix the number of physical resources and employ the same classical optimization protocols across all configurations to ensure a fair comparison. The achieved experimental results show that the use of indistinguishable photon states, combined with active feedback dynamics, significantly enhances the expressivity of the reservoir, enabling it to accurately reconstruct higher-order nonlinear functions. As a result, this enhanced nonlinear processing capability does not compromise the fading memory property of the system, which remains essential for effective temporal information processing. Therefore, this improvement translates into superior performance on time-dependent benchmark tasks, including temporal XOR, NARMA sequences, and the prediction of Mackey-Glass chaotic series. On the other hand, we prove that classical correlations alone, present in distinguishable two-photon states, provide little to no advantage compared to quantum correlations.

Results

Photonic quantum reservoir computing: protocol

Reservoir computing is a computational framework inspired by recurrent neural networks (RNNs) in which information is encoded into the dynamics of a fixed nonlinear system, the reservoir. The reservoir maps the inputs into a high-dimensional space, thus enabling complex learning tasks to be solved efficiently simply via linear regression. The reservoir can be a physical system exhibiting quantum behavior, in this case, the approach is known as QRC.

A schematic representation of the protocol we implement for temporal signal processing is illustrated in Fig. 1. In this approach, the training data are sequentially fed into the reservoir as input. The reservoir, treated as a *black box* with random internal connections, dynamically evolves and nonlinearly maps the input features into a higher-dimensional space. The resulting outputs encode temporal patterns and are subsequently processed by a digital linear readout layer, which enables forecasting of future values of the signal.

Our quantum reservoir architecture is based on an integrated quantum photonic platform, injected with either single or multiphoton states, as depicted in Fig. 1b. The time series is mapped into the reservoir dynamics by modulating an optical phase governing the transformation implemented by the circuit on the input photons. In particular, we focus on evaluating the performance of the system by comparing the results obtained when using distinguishable and indistinguishable photon inputs, aiming to assess the impact of quantum correlations on the system's computational capabilities, while the training procedure remains entirely classical.

Whether exploiting photon indistinguishability as a resource for classical ML tasks provides a consistent advantage over classical systems has remained an open question. Recent works have highlighted the potential of indistinguishable photons to enrich the feature space of photonic learning systems and increase the system's computational expressivity[32,33]. However, experimental investigations have been mostly limited to photonic extreme learning machines without feedback, and have reported little or no measurable advantage from photon indistinguishability in low-dimensional systems with only two photons[32].

Given a classical input sequence $\{s_k\}$, we want to predict the time series $\{y_k\}$ generated by the input sequence through an unknown nonlinear or time-varying function. The protocol we implement consists of four key steps: encoding the input into the quantum system, evolving the quantum reservoir, measuring the reservoir output states, and applying a feedback operation conditioned on the measurement outcome.

The first step is to encode the input value $s_k \in [0, 1]$ into the quantum system by modulating the optical phase as $\phi_B = a_{\text{in}} \cdot s_k$, where $a_{\text{in}} \in [-\pi, \pi]$ is an input scaling factor. This phase modulation acts on photons via a phase shifter on the selected optical mode, described by the unitary matrix:

$$U(\phi_B) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{i\phi_B} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (1)$$

At each time step k, the reservoir evolves in response to the encoded input s_k. After the evolution corresponding to each element of the input sequence, the resulting output state is measured using a positive operator-valued measure (POVM) Π_μ, where μ is the number of possible measurement outcomes. To accurately reconstruct the output probabilities, each input in the sequence is sampled multiple times, collecting sufficient statistics over repeated measurements. This process yields an output probability distribution $\mathbf{p}_k(\vec{\mu}) = [p_{k,1}, p_{k,2}, \ldots, p_{k,\mu}]$ associated with the k input.

Finally, the application of the feedback loop plays a crucial role in enriching the computational dynamics of the reservoir by introducing both memory and nonlinearity into the dynamics, allowing the system to adapt based on past outputs. In our implementation, the information derived from selected measurement outcomes $\mathbf{p}_k$ is reinjected into the system during the next cycle $k + 1$. This feedback operation is performed by encoding the selected components of $\mathbf{p}_k$ in the circuit parameters as phase shifts of the form $a_{\text{fb}} \cdot p_k(\mu')$, where $a_{\text{fb}} \in [-\pi, \pi]$ is the feedback scaling weight. This mechanism breaks the unitary evolution typically associated with quantum mechanics and allows the system to process temporal patterns and nonlinear input-output dependencies, while preserving its internal architecture.

For each input s_k of the sequence, the vector of output probabilities $\mathbf{p}_k$ is reconstructed from photon-coincidence events detected between the different outputs of the circuit. These vectors act as high-dimensional feature representations of the system's response to each input of the sequence. Once the probability vectors for each element of the sequence have been collected, they are associated with the corresponding target y_k. At this point, part of the collected data $\{(\mathbf{p}_1, y_1), \ldots, (\mathbf{p}_K, y_K)\}$ of total length K is split into a training set including a sequence of size K_{tr}, used to train the linear regression model, while the remaining $K_{\text{ts}} = K - K_{\text{tr}}$ data is reserved to evaluate its performance on unseen data. We denote model predictions with a tilde. The goal is

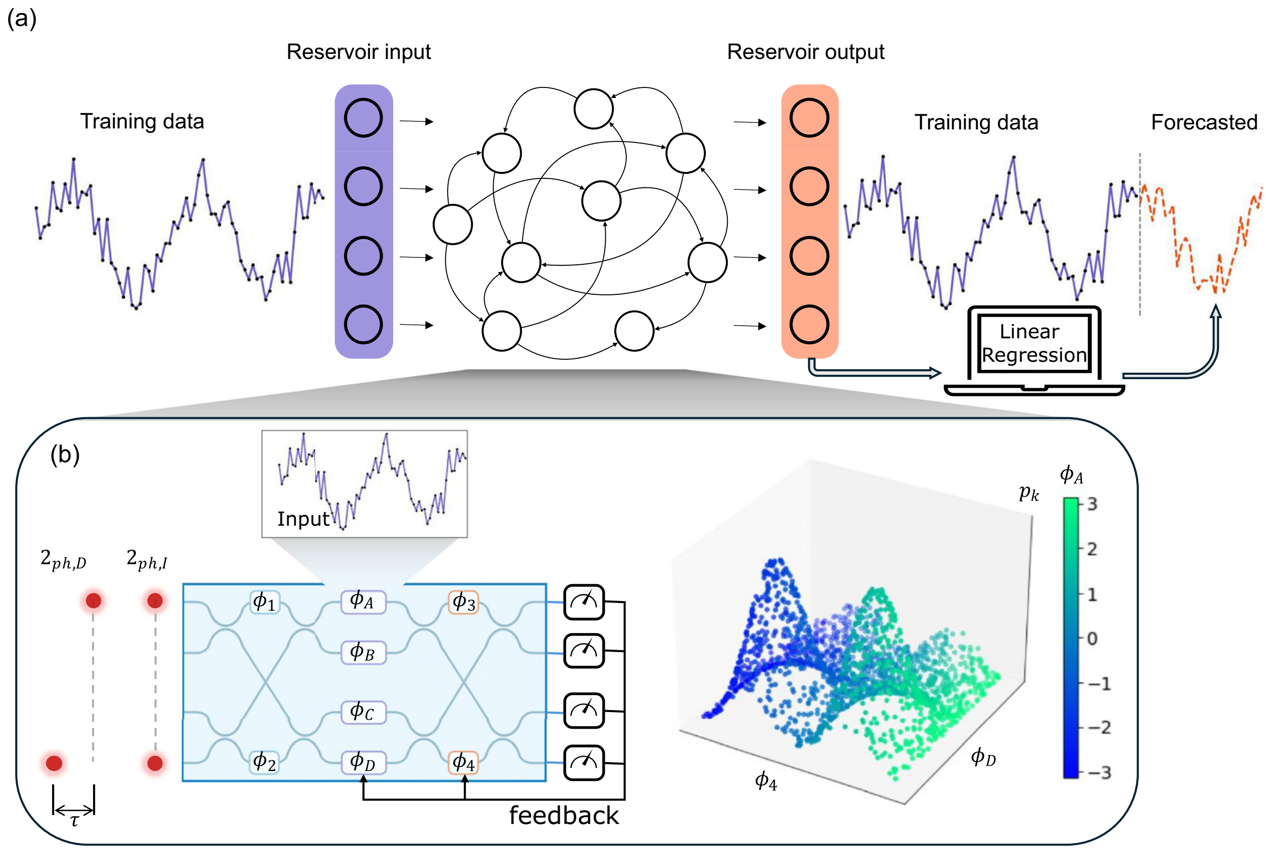

Fig. 1 | Schematic overview of the realized photonic quantum reservoir computing. a Scheme of a classical echo state network (ESN) for time-series forecasting[35]. The input data is injected into a fixed, recurrent reservoir of nonlinear nodes. The reservoir output is linearly combined and trained via linear regression to predict future values. **b** Photonic implementation of a quantum reservoir computing scheme. Two-photon input states, either indistinguishable ($2_{ph,I}$) or distinguishable ($2_{ph,D}$) with a control on the delay τ, are injected into a four-mode reconfigurable photonic integrated platform. This includes the tunable phases ϕ_i (with $i = 1, 2, A, B, C, D, 3, 4$). The input data is encoded into one phase of the central layer. The output statistics are reconstructed using single-photon detectors. A feedback-loop mechanism is based on the reinjection of the past measurement outcomes and is implemented via phase reconfiguration of ϕ_D and ϕ_4. On the right, the plot of an example of a component of the output probability distribution as a function of the internal phases ϕ_A, ϕ_D, and ϕ_4.

to learn a function that, given unseen inputs $\mathbf{p}_{ts}$ from the test set, can accurately predict their corresponding target values $\widetilde{\mathbf{y}}_{ts}$.

We train the digital linear model using *Ridge* regression (see "Methods"), and evaluate its performance by comparing the predicted output vector $\widetilde{\mathbf{y}}_{ts}$ to the true target vector $\mathbf{y}_{ts}$ of the test dataset. In particular, the determination coefficient R^2 is a figure of merit which quantifies their similarity, and is defined as:

$$R^2(\widetilde{\mathbf{y}}_{ts}, \mathbf{y}_{ts}) = \frac{\mathrm{cov}^2(\widetilde{\mathbf{y}}_{ts}, \mathbf{y}_{ts})}{\sigma^2(\widetilde{\mathbf{y}}_{ts})\sigma^2(\mathbf{y}_{ts})}, \tag{2}$$

where cov and σ^2 are the covariance and variance, respectively. This measures the strength of a linear relationship between the two variables. It ranges from 0 to 1, where 0 indicates no linear correlation, while 1 indicates a perfect correlation.

Another metric is the mean-squared error (MSE):

$$\mathrm{MSE} = \frac{1}{K_{ts}} \sum_{k=1}^{K_{ts}} \left(\widetilde{y}_{ts,k} - y_{ts,k}\right)^2, \tag{3}$$

which quantifies how closely the predicted values match the target ones.

Task-independent characterization

To evaluate the computational capabilities of the QRC model, we adopt a set of standard benchmarks and performance metrics from classical ML. These tasks enable both to evaluate the performance of the protocol and to characterize the underlying physical system used as the computational substrate. The two main properties typically investigated are the *short-term memory* and the *expressive power*. The former quantifies the ability of the system to retain and recall information about past inputs, while the latter evaluates the range and complexity of input-output transformations the reservoir can implement. High expressivity indicates a strong capacity to approximate a broad class of nonlinear functions of the inputs.

The analysis of the memory capacity of an ML model involves processing an input sequence randomly sampled from a uniform distribution $s_k \in [0, 1]$, and the target is to reproduce the input value from d cycles earlier, i.e., $y_k = s_{k-d}$. The performance of the model at each delay d is measured using the determination coefficient R_d^2. The total memory capacity is then defined as the sum of these values: $C_\Sigma = \sum_{d=0}^{d_{max}} R_d^2$, where d_{max} is an integer such that $R^2_{d_{max}}$ is approximately zero.

We evaluate the expressivity of the system by feeding an ordered input sequence $s_k \in [0, 1]$ into the reservoir, where each value is mapped to a corresponding optical phase ϕ_B. The readout layer is then trained to approximate nonlinear target functions such as monomials or more complex polynomials. Specifically, the choice of input encoding and the effective dimensionality explored by the reservoir determine its expressive capacity, limiting the set of functions it can represent, typically characterized by a finite set of Fourier frequencies[34].

A common approach to forecasting time series data relies on RNN architectures, since these tasks require both high memory capacity and nonlinear processing capabilities. However, training RNNs can be challenging, especially in neuromorphic implementations. A simplified alternative to address such problems is represented by echo state networks (ESNs)[35] or, more broadly, by reservoir computing. In an ESN, the recurrent component is a fixed, sparsely connected network of nonlinear nodes left untrained, as shown in Fig. 1a. This architecture significantly reduces training complexity while preserving the ability to perform complex time-dependent tasks. However, this is contingent on the optimal exploitation of the system's available degrees of freedom. In particular, the total memory capacity of these architectures is upper-bounded by the number of reservoir nodes N_r, i.e., $C_\Sigma \leq N_r$, and more generally by the rank of the reservoir's correlation matrix $\mathbf{R}$[36].

In QRC, dimensional limitations are particularly critical because the available degrees of freedom are fundamentally constrained by physical resources, such as the number of modes, the number of photons, and the set of accessible measurement outcomes. Furthermore, practical implementations are affected by real-world imperfections, including statistical noise[37,38], losses, and imperfect detection. These factors can further reduce the effective dimensionality of the computational space explored by the system. In this context, we characterize the dimension of the explored space via the Gram matrix of the reservoir state $\mathbf{X}$, where each row represents the state at a given time step. It is computed as $\mathbf{G} = \mathbf{X}^T\mathbf{X}$, and encodes the pairwise inner products between reservoir states across time, offering a quantitative measure of the effective dimensionality explored by the observables[28].

Experimental realization

Our photonic quantum reservoir consists of a hybrid platform made up of a single-photon source, a reconfigurable four-arm integrated interferometer, fabricated via femtosecond laser waveguide writing in glass[39], illustrated in all its details in ref. 40, and single-photon detectors. The source is a BBO crystal designed for Type-II Spontaneous Parametric Down Conversion (SPDC), which is used to generate photon pairs that are subsequently coupled into single-mode fibers (SMFs). The photons are spatially separated, and their indistinguishability, in both time and polarization, is controlled using delay lines and fiber polarization controllers, respectively.

The PIC architecture is shown in Fig. 1b, the device includes only linear optical elements, such as beam splitters (BSs), thermal phase shifters (PSs), and SWAP gates. This architecture allows us to implement the quantum reservoir protocol. In the following, we detail each of the four main steps: state preparation and input encoding, evolution, measurement, and feedback operation.

Preparation and encoding. We evaluate the system performance by exploiting the path degree of freedom for three different input configurations: single-photon ($n_{ph} = 1$), two distinguishable photons ($n_{ph,D} = 2$), and two indistinguishable photons ($n_{ph,I} = 2$) inputs. In the single-photon configuration, the input state $|1000\rangle$ corresponds to injecting a single photon into the first mode of the PIC. The second photon here is used solely as a trigger for post-selected coincidence measurements. The temporal input sequence is encoded into the phase parameter ϕ_B, which modulates the single-photon state as follows:

$$\begin{aligned} |\psi_B\rangle_1 \quad &= \tfrac{1}{2}(|1\,0\,0\,0\rangle - ie^{i\phi_B}|0\,1\,0\,0\rangle \\ &-i|0\,0\,1\,0\rangle - |0\,0\,0\,1\rangle) \end{aligned} \tag{4}$$

By also injecting the second photon into the last mode of the device, we initialize the input state in $|1001\rangle$. In this scenario, we have access to a larger computational space; therefore, we can assess whether additional dimensionality enhances the system's ability to solve the target tasks. In the case of distinguishable photons, the measurement outcomes provide direct access to classical joint probability distributions over pairs of output modes, reflecting correlations that arise from the co-propagation of distinguishable photons through the circuit. These distributions contain additional information compared to single-photon inputs, reflecting classical correlations among the photons' paths. In contrast, when indistinguishable photons are injected, quantum interference leads to measurement statistics that reflect non-classical correlations. These can be exploited to realize richer transformations of the input sequence, thus serving as an additional resource to increase the expressivity and computational capabilities of the reservoir. When injecting indistinguishable photons into the two modes of the PIC, the evolution through the preparation layer allows us to generate two-photon entangled states by leveraging the effect of quantum interference at the BS level. Thus, we investigate the scenarios where the photons are injected either indistinguishably in time or with a controlled temporal delay τ as represented in Fig. 1b, allowing us to prepare in a controlled way separable or entangled two-photon states.

In the encoding layer, the input sequence value s_k is encoded via a phase shifter acting on the phase ϕ_B. Consequently, once the state is prepared and the phase is encoded, the entangled two-photon state becomes:

$$\begin{aligned} |\psi_B\rangle_{2,I} = \quad &- \quad \tfrac{1}{2\sqrt{2}}(|2\,0\,0\,0\rangle + e^{2i\phi_B}|0\,2\,0\,0\rangle + |0\,0\,2\,0\rangle \\ &+ \quad |0\,0\,0\,2\rangle) + \tfrac{1}{2}(|1\,0\,0\,1\rangle - e^{i\phi_B}|0\,1\,1\,0\rangle). \end{aligned} \tag{5}$$

On the other hand, the two-photon distinguishable state takes the form:

$$\begin{aligned} |\psi_B\rangle_{2,D} \quad &= \tfrac{1}{4}\big[\big(|1000\rangle_1 + ie^{i\phi_B}|0100\rangle_1 + i|0010\rangle_1 \\ &-|0001\rangle_1\big) \otimes \big(|1000\rangle_2 + ie^{i\phi_B}|0100\rangle_2 \\ &+|0010\rangle_2 - |0001\rangle_2\big)\big], \end{aligned} \tag{6}$$

where the indices 1 and 2 refer to the two distinguishable photons, which evolve independently across the four spatial modes.

Evolution. At this stage, the system evolves according to a unitary evolution applied to the input modes of the injected photons, depending on the feedback applied to the phases ϕ_D and ϕ_4. This evolution is realized by the PIC and is given by the unitary operator:

$$\mathcal{U}_{\phi_4} \otimes \mathcal{U}_{\phi_D} = \frac{1}{2}\begin{pmatrix} e^{i(\phi_D+\phi_4)} & ie^{i\phi_4} & i & -1 \\ ie^{i(\phi_D+\phi_4)} & -e^{i\phi_4} & 1 & i \\ ie^{i\phi_D} & 1 & -1 & i \\ -e^{i\phi_D} & i & i & 1 \end{pmatrix}. \tag{7}$$

This unitary dynamics, which incorporates the input-dependent feedback encoded in the phases, drives the internal quantum state of the reservoir and encodes temporal correlations across input sequences. Further details about the evolution of the states injected into the PIC are provided in the Supplementary Information S1.

Measurement. The probability distribution of the circuit's output states is measured using avalanche photodiodes (APDs). For each experimental configuration, we collect an average total number N_{tot} of coincidence counts, which are used to reconstruct the output probability vector $\mathbf{p}_k(\vec{\mu})$. Specifically, the same average counts have been chosen for all the different optical inputs to ensure consistent statistical resolution and a fair comparison, thus granting that differences in performance arise from the system's dynamics rather than variations in measurement statistics. To reliably reconstruct all the possible output configurations and distinguish the events in which the two photons exit through different modes from those in which both photons are detected in the same mode, we employ fiber beam splitters at each output port. This detection scheme yields 10 distinct outcomes when two-photon input states are injected, and 4 possible outcomes when single-photon inputs are used.

Feedback operation. The measurement performed at the time step k will condition the encoding and the measurement of the next step $k+1$, since a part of the distribution $\mathbf{p}_k(\vec{\mu})$ is fed back into the reservoir in order to implement the temporal feedback loop. Specifically, the feedback applied to the phase ϕ_D during the step k is determined by the measurement outcome corresponding to a selected component μ' of the probability reconstructed at the previous time step, i.e., $p_{k-1}(\mu')$. On the other hand, the feedback applied to ϕ_4 is set based on a probability component that incorporates the system evolution in two steps earlier, i.e., $p_{k-2}(\mu'')$. After acquiring the entire sequence of measurements, the model is trained using Ridge regression, which maps the recorded output probabilities to the target values of the computational task.

Experimental results

We start by choosing the configuration of phase shifters that maximizes the rank of the system's Gram matrix. This selection defines the transformation used to encode the input sequence and the ones of the two feedback loops applied during the protocol. Consequently, the integrated device is calibrated to have full control over the selected phase-shifters.

Once the transformation is fixed, task-specific optimization can be performed by tuning the weights that map both the input sequence and the feedback signals onto the corresponding optical phases. This corresponds to adjusting the input weight coefficient a_{in}, as well as the two feedback weights $a_{fb,D}$ and $a_{fb,4}$, which control the influence of the feedback on the optical phases ϕ_D and ϕ_4, respectively. These parameters are treated as hyperparameters; however, they are optimized once for a given task and then kept fixed for the entire run. To perform this optimization efficiently, we employ the Python `Optuna` package, a hyperparameter optimization framework based on Bayesian optimization algorithms[41]. In addition to tuning the weights, we also optimize the choice of which components μ' and μ'' of the output probability vector are used to drive the two feedback loops. Indeed, depending on the input sequence, the informativeness of the output modes changes, making this choice task-dependent. These hyperparameter settings permit the system to adapt to the specific task demand: some tasks benefit from greater memory capacity to capture temporal correlations, while others require enhanced nonlinearity to improve the separation of complex input patterns[42]. Further details are provided in the Supplementary Information S1, Supplementary Tabs. S1-2 and Supplementary Figs. S4 and S5. After this task-level calibration, the experiment proceeds with all the parameters fixed; the only quantity that changes from step to step during reservoir operation is the input phase and the feedback phases, which are updated depending on the previous output probabilities. The only model fitted from data is the final classical linear readout. Finally, to quantify the statistical uncertainty of our experimental results, we propagate photon-counting shot noise through the full analysis pipeline using a Monte Carlo resampling procedure. Starting from the experimentally recorded detection counts used to estimate the output probability distributions, we generate synthetic datasets by Poisson-resampling each count. This reproduces the dominant noise contribution in our experiment, namely, finite counting statistics. For each task, we generate 100 resampled datasets and recompute the complete performance evaluation for each realization.

We characterize our system performance in the regimes of single-photon input states ($n_{\mathrm{ph}} = 1$) and two-photon input states. In particular, we consider both distinguishable ($n_{\mathrm{ph},D} = 2$) and indistinguishable ($n_{\mathrm{ph},I} = 2$) two-photon inputs.

Starting with the study of short-term memory via the linear memory capacity task, we encode a randomly and uniformly sampled sequence of values $s_k \in [0, 1]$ into the reservoir. The resulting reconstructed output probability vector is then used to train the linear regression model. The results of the linear memory capacity task are reported in Fig. 2a. In this analysis, we compare the performance of the QRC system when exploiting the dynamics of the aforementioned different optical input states. These results highlight that the presence of feedback dynamics is the main factor that gives memory to the system. In contrast, quantum correlations alone do not appear to play a decisive role in this task, as the memory capacity is essentially identical for two-photon states in both distinguishable and indistinguishable configurations. Notably, when feedback is active, the QRC with a two-photon input state achieves slightly better performance compared to the single-photon case, as detailed in the Supplementary Information S2 and Supplementary Fig. S14.

The expressivity of the physical reservoir is evaluated by testing the ability of the QRC system to reconstruct monomial functions $f_n(x) = x^n$, with $x \in [0, 1]$ and n ranging from 2 to 13. This benchmark is commonly used to quantify the nonlinear computational capacity of a reservoir, since higher-order monomials require more complex transformations of the input. The MSE in the monomial reconstruction on the test set is reported as a function of the exponent order n in Fig. 2b for the different reservoir configurations investigated. We observe in this scenario that the configuration employing indistinguishable photon inputs in combination with feedback dynamics consistently achieves superior performance compared to other scenarios, resulting in lower MSE values, showing that quantum correlations can serve as a valuable resource for enhancing the system's expressivity. Such an enhancement is evident not only in improved prediction accuracy but also in the learning efficiency, which can be quantified either in terms of required training epochs for convergence[43] or by the number of training data to achieve a certain performance threshold. Specifically, as we show in the Supplementary Information S2 and Supplementary Figs. S6–S12, the configuration using two indistinguishable photons reaches performance saturation faster than other input configurations, indicating that quantum interference enables the reservoir to construct richer feature representations with fewer training examples. Importantly, simply having access to joint classical probability distributions leads only to a modest improvement in terms of expressivity. This is evident in the configuration with two distinguishable photons, which slightly outperform the single-photon case, despite offering more degrees of freedom. These classical regimes tend to saturate at the same expressivity limit. In contrast, the use of indistinguishable photons provides a more evident enhancement in this nonlinear learning task.

The expressivity of the QRC is further tested by considering as a second benchmark the reconstruction of nonlinear polynomial functions of the form $f_N(x) = \sum_{n=1}^{N}(-1)^n x^n$. Using the same dataset as in the monomial case, this task evaluates the reservoir's ability to capture a richer combination of nonlinear terms, offering a stronger test of its computational capabilities. The MSE in reconstructing these polynomial functions for increasing integer values of $N \in \{1, \ldots, 7\}$ is plotted in Fig. 2c. As in the monomial case, the configuration with indistinguishable photons achieves the lowest MSE, indicating superior expressivity. In contrast, configurations using distinguishable photons, as well as the single-photon input, exhibit significantly worse performance as N increases. Other effects of the photon indistinguishability in terms of the expressivity of the photonic platform are provided in the Supplementary Information S1 and Figs. S1–S3.

Experimental machine learning benchmarks

The computational capabilities of the photonic QRC platform are finally evaluated on standard ML tasks. In particular, we benchmark its performance on tasks widely adopted in reservoir computing literature[23,26,30,31,44]. These include the temporal XOR, the nonlinear autoregressive moving average family (NARMA-N), and the Mackey-Glass (MG) time-series forecasting task. Each of them is designed to probe different aspects of the system's computational performance. In particular, these testbeds allow us to analyze how the physical platform processes information and to investigate the role of quantum correlations as a computational resource when applied to real classical datasets.

The temporal XOR is a binary task where the target output is $y_k = s_k \bigoplus s_{k-d}$, with an input sequence of binary values $s_k \in \{0, 1\}$. This task requires the system to store and nonlinearly combine temporally delayed inputs. Since the performance of the model is evaluated using the accuracy on a test sequence of new data, the raw outputs of the linear regression are rounded in a successive step to the binary values 0 and 1.

The achieved results are shown in Fig. 3a, b, where the accuracy performance changes varying the delay $d \in \{1, \ldots, 4\}$. Specifically, the bar plot in Fig. 3a shows the accuracy for the different delays d, for the different input states: $n_{\mathrm{ph}} = 1$, $n_{\mathrm{ph},D} = 2$ and $n_{\mathrm{ph},I} = 2$. As expected, the accuracy decreases with increasing delay, reflecting the growing memory demand of the task. For $d = 1$ and $d = 2$, all the configurations perform nearly perfectly, with

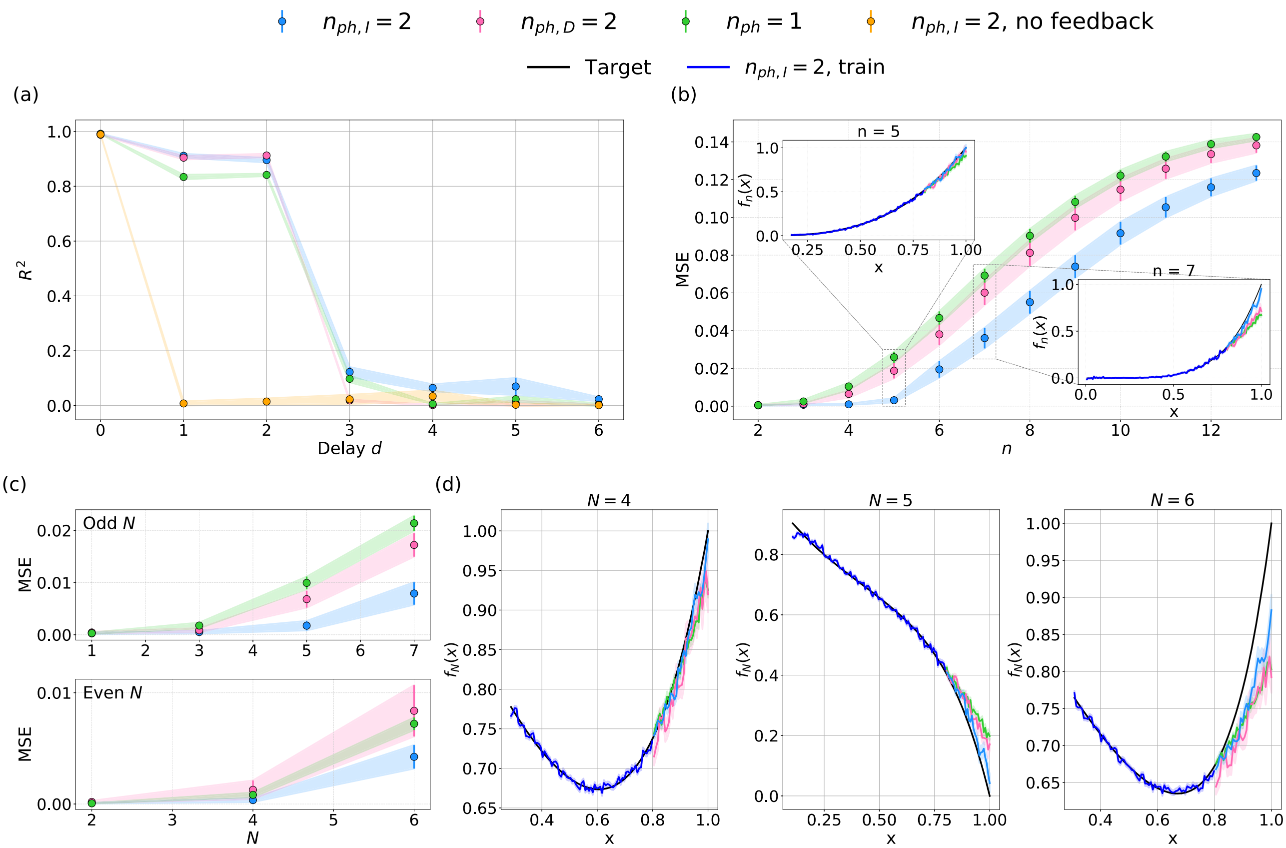

Fig. 2 | Characterization of the quantum reservoir computing model: short-term memory and expressivity. We report the quantum reservoir performance achieved for different photonic input states: two indistinguishable photons ($n_{\mathrm{ph},I} = 2$, light blue dots), two distinguishable photons ($n_{\mathrm{ph},D} = 2$, pink dots), and a single photon ($n_{\mathrm{ph}} = 1$, green dots). **a** The short-term memory capacity is measured via the coefficient of determination R_d^2 as a function of the delay $d \in [0, 6]$. We report the results obtained when injecting into the device two indistinguishable photons but without feedback dynamics (yellow dots). The dataset contains 497 points in total. **b** The expressivity of the QRC model is measured by looking at the reservoir performance, in terms of the mean squared error (MSE) on the test set, in reconstructing $f_n(x) = x^n$ for $n = 2, \ldots, 13$. Insets display the predicted outputs for $n = 5$ and $n = 7$ for different photon-input configurations compared to the target (black lines). It is also plotted the corresponding training output for $n_{\mathrm{ph},I} = 2$ (dark blue line). The datasets contain 150 points in total. **c** The expressivity of the model is further explored through the approximation of polynomial functions defined as $f_N(x) = \sum_{n=1}^{N}(-1)^n x^n$, with $N = 1, \ldots, 7$. The MSE is plotted as a function of N. **d** Three examples of polynomial function approximations for selected values of $N = 4$, 5, and 6. In panels (**b**) and (**d**), the lighter solid lines (light blue, pink and green) correspond to the test-set predictions for $n_{ph,I} = 2$ inputs, $n_{ph,D} = 2$ inputs, and $n_{\mathrm{ph}} = 1$ input states, respectively. Error bars and shaded areas in the plot are obtained from Monte Carlo Poisson resampling of the experimental photon-count distributions (100 realizations) and represent one standard deviation of the resulting performance distribution.

accuracies approaching 1. However, as the delay increases to $d = 3$, the accuracy of the configuration with $n_{\mathrm{ph},I} = 2$ continues to learn effectively, while the one with either $n_{\mathrm{ph}} = 1$ or $n_{\mathrm{ph},D} = 2$ drops significantly, approaching the random guess level of 50%. Figure 3b shows the output for the input state configurations with $n_{\mathrm{ph},I} = 2$ and $n_{\mathrm{ph},D} = 2$, considering $d = 1$ and $d = 3$. In each subplot, the predicted output y_k is plotted over the 60 time steps representing the test set. At $d = 1$, both systems track the target sequence closely, while for $d = 3$, although the predictions become noisier and deviate more often from the target, the configuration with $n_{\mathrm{ph},I} = 2$ still reaches an accuracy of 0.70 ± 0.03. In contrast, the configuration with $n_{\mathrm{ph},D} = 2$ is no longer able to reconstruct the output, resulting in an accuracy equal to a random guess. In contrast to the short-term memory study, these results show that exploiting quantum correlations in the input state enables the system to better handle more complex tasks. This is evidenced by the marked improvement in performance observed for the input state $n_{\mathrm{ph},I} = 2$ compared to $n_{\mathrm{ph},D} = 2$ configuration. We note that the behavior observed in the single-photon experimental case (in particular, the improvement from $d = 1$ to $d = 2$) originates from the global optimization procedure, which was performed to maximize the overall performance across all delay tasks rather than each delay independently; a task-specific optimization would maximize the performance at each delay, generally yielding different optimal configurations for different delays. As further evidence, in the ideal numerical simulation reported in Fig. S15(c) of the Supplementary Information, where this global experimental optimization constraint is absent, performance is indeed monotonic.

Next, the indistinguishability advantage is investigated via the NARMA sequence task, which tests both short- and long-term memory combined with nonlinear transformations. The input sequence is randomly and uniformly sampled in [0, 1], and the target is:

$$y_k = \alpha y_{k-1} + \beta y_{k-1} \sum_{j=k-N}^{k-1} y_j + \gamma\left(u_{k-1}^3 + u_{k-1}^5\right) + \delta, \qquad (8)$$

where $u_k = \mu + \nu s_k$ with $\mu = 0$ and $\nu = 0.2$, while the other parameters are $\alpha = 0.3$, $\beta = 0.05$, $\gamma = 50$ and $\delta = 0.1$. The evaluation of the task is illustrated in Fig. 3c, where performance is studied via the MSE as a function of the NARMA order N[45]. Specifically, Eq. (8) is a modified version of the usual NARMA equation present in the literature and is used to emphasize the nonlinear processing capability of the system together with short-term memory requirements. This choice is motivated by the memory capacity analysis of our platform, which depends on the feedback architecture, in particular, by the number of active feedback phases. The adapted task, therefore, remains genuinely dynamical and requires a finite amount of memory, but it is designed to be more sensitive to nonlinear transformations of the input. The results show a similar behavior to the one obtained for the temporal XOR

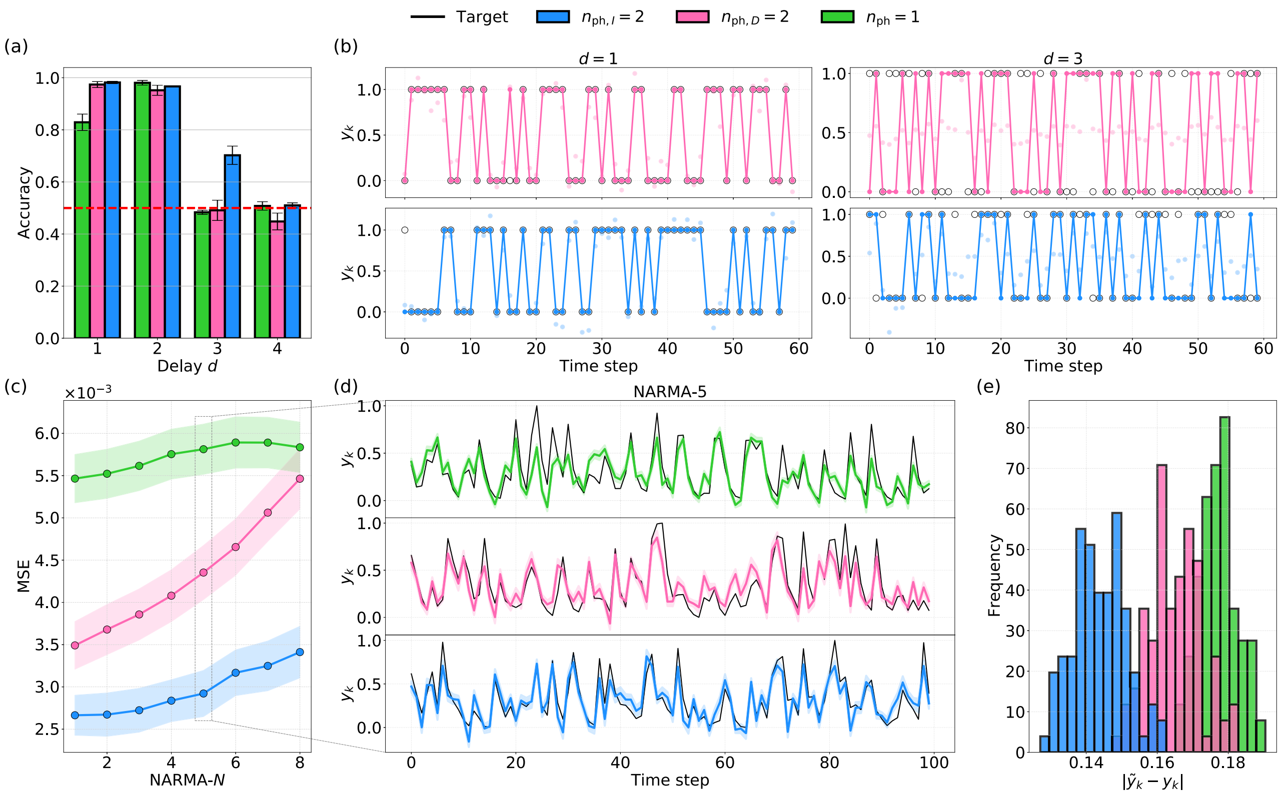

Fig. 3 | Quantum reservoir performance on machine learning benchmarks: temporal XOR and NARMA-*N*. The tasks are evaluated for different input configurations: two indistinguishable photons ($n_{ph,I} = 2$, light blue), two distinguishable photons ($n_{ph,D} = 2$, pink), and single photons ($n_{ph} = 1$, green). **a** The temporal XOR is defined as $s_k \bigoplus s_{k-d}$, where $s_k \in \{0, 1\}$ is the input sequence, and d is the temporal delay. The task is evaluated using classification accuracy for delays $d = 1, \ldots, 4$. The datasets contain 300 points in total. **b** The output predictions are compared with the targets for both $n_{ph,I} = 2$ and $n_{ph,D} = 2$ configurations, and for $d = 1$ and $d = 3$. While both perform well at $d = 1$, only the $n_{ph,I} = 2$ configuration successfully addresses the increased memory and nonlinearity demands at $d = 3$. **c** Performance on the test set for the NARMA-N task[45]. Also, this testbed shows the enhancement due to quantum correlations; indeed, the configuration with $n_{ph,I} = 2$ always outperforms the one with $n_{ph,D} = 2$ and $n_{ph} = 1$. The datasets contain 500 points in total. **d** The test with respect to the target for the three configurations investigated is plotted for the function NARMA-5. Both the predicted and the target values y_k are normalized to ensure a fair comparison. **e** The different normalized mean absolute errors between the predicted and the target values are compared in a histogram. Shaded areas in the plot are obtained from Monte Carlo Poisson resampling of the experimental photon-count distributions (100 realizations) and represent one standard deviation of the resulting performance distribution.

task, indeed the system with $n_{ph,I} = 2$ outperforms both the $n_{ph} = 1$ and $n_{ph,D} = 2$ cases across all the values of $N \in \{1, \ldots, 8\}$. The comparison for NARMA-5 is further emphasized in Fig. 3d, showing how the configuration with indistinguishable photon tracks the target more closely than the distinguishable case. This supports the point that the presence of quantum correlations expands the accessible computational space in a meaningful way, enabling the reservoir to capture and reproduce complex temporal dependencies, especially when nonlinear dynamics are involved. Finally, the results in Fig. 3e summarizes this behavior through the mean absolute error among the different input states.

The last task examined is the forecasting capabilities of our QRC model for the Mackey-Glass time delay differential equation, described by complex continuous-time chaotic dynamics:

$$\mathring{s}(t) = \frac{\alpha s(t-\tau)}{1 + s(t-\tau)^{\beta}} - \gamma s(t), \tag{9}$$

where the parameters are fixed to $\alpha = 0.2$, $\beta = 10$, $\gamma = 0.1$, and $\tau = 17$. The latter is responsible for the chaotic attractor dynamics shown in Fig. 4a. Then, the differential equation is numerically solved assuming the initial condition $s(0) = 1.2$, a constant past history $s(t) = 1.2$ for $t \in [-\tau, 0]$, and setting the integration time step to $\Delta t = 1$. The resulting sequence is sampled at each integration step, so that one update $s(t) \rightarrow s(t + \Delta t)$ corresponds to one reservoir cycle $k \rightarrow k + 1$. The scalar time series is then min-max normalized to obtain values $s_k \in [0, 1]$, which are encoded into the phase of the photonic circuit. The aim of this task is to perform a time-series forecasting that generates predictions t_f cycles into the future, i.e., trying to predict the target values $y_k = s_{k+t_f}$. In Fig. 4b, c, we evaluate the ability of the model to learn the Mackey-Glass time series, comparing the forecasting for $t_f = 3$ across the two configurations: $n_{ph,D} = 2$ and $n_{ph,I} = 2$. While both scenarios involve two-input photons, the output spectra of the optical network show different numbers of frequency components. This difference arises from the total number of indistinguishable input photons, which affects the capacity of the model to represent complex signal structures[46]. Despite this difference, both configurations succeed in capturing the fundamental frequency of the target signal. This is enabled by the direct encoding of the time-series and by the optimization of the input scaling hyperparameter a_{in}, which helps the output of the network to match its frequency. However, the configuration with two indistinguishable input photons also in this scenario slightly outperforms the reservoir with separable states in reconstructing the signal amplitude, showing an enhanced expressivity[34], thus expanding the set of functions accessible to the linear regression model. Additional details are provided in the Supplementary Information S2 and Supplementary Figs. S13–S15, together with the scaling of the platform effect Supplementary Figs. S16–S20.

Discussion

The ambition to develop quantum neural networks has attracted growing interest recently. A widely adopted approach relies on variational strategies[47], where the parametrized quantum circuit is trained through a

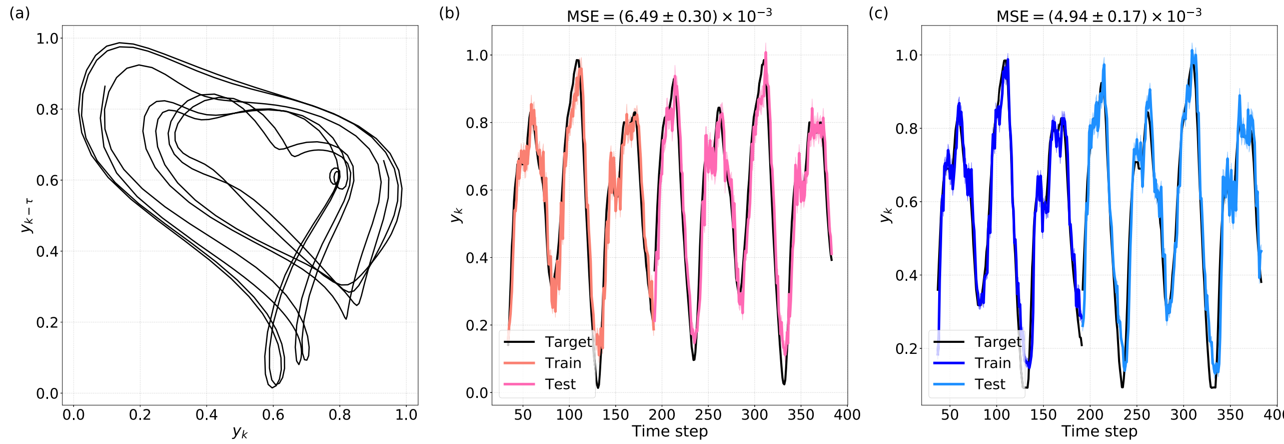

Fig. 4 | Mackey-Glass time series forecasting. a Chaotic attractor of the Mackey-Glass (MG) time-delay differential equation, simulated with parameters $\alpha = 0.2$, $\beta = 10$, $\gamma = 0.1$, and delay $\tau = 17$, which induces chaotic dynamics. The differential equation is numerically integrated, then time-discretized, normalized, and directly encoded as input to the quantum reservoir system. **b**, **c** Forecasting of future time steps with prediction horizon $t_f = 3$, comparing the target signal (black) with the predicted outputs during both training and testing phases. Two photon-input configurations are considered: **b** injection of two distinguishable photons ($n_{\mathrm{ph},D} = 2$, shown in salmon for training and pink for test), and **c** injection of two indistinguishable photons ($n_{\mathrm{ph},I} = 2$, shown in dark blue for training and light blue for test). The achieved MSE is reported in the title of the corresponding plot. Each dataset contains 390 points. Shaded areas in the plot are obtained from Monte Carlo Poisson resampling of the experimental photon-count distributions (100 realizations) and represent one standard deviation of the resulting performance distribution.

hybrid quantum-classical feedback loop. While successful in a range of applications[48–51], such methods face intrinsic scalability challenges, including the Barren plateau problem[52], where gradients vanish exponentially with system size. This makes this procedure insufficient or infeasible when dealing with complex and high-dimensional systems that require extensive gradient optimization. These limitations motivate the exploration of alternative models such as quantum reservoir computing. At the same time, it is important to note that QRC does not automatically bypass all scalability bottlenecks. Recent works have shown that, in certain QRC settings, particularly for increasingly large and highly scrambling reservoirs, the measured reservoir features can exhibit exponential concentration with system size, implying an unfavorable sampling complexity and potentially degraded learnability at scale[53,54]. This highlights a nontrivial trade-off between increasing expressivity and avoiding regimes where readout observables become too strongly concentrated, while overly restricting the dynamics may reduce expressivity and lead to classical simulability.

In this work, we have experimentally implemented a photonic quantum reservoir protocol for temporal sequences forecasting using a reconfigurable integrated photonic platform, necessary for implementing real-time feedback loops. By injecting into the reservoir single-photon, two-photon distinguishable, and two-photon indistinguishable states, encoded in the path degree of freedom, we have systematically investigated how quantum correlations influence computational performance.

The successful forecasting of time-series data generally relies on two key capabilities of the computing platform: its capacity to memorize information about past inputs, to keep track of temporal correlations, and the ability to model nonlinear input-output relations. Our study demonstrates that the first capability can be engineered in integrated photonics systems through adaptive feedback mechanisms, with performance mainly governed by the number of active feedback loops and depending on the size of the accessible configuration space. These elements define the reservoir's memory. The second capability, linked to the expressivity of the system, can be improved by the use of multiphoton quantum states. When quantum correlations are introduced, exploiting photon indistinguishability, the system gains a clear advantage in reconstructing higher-order nonlinear functions when fixing the number of physical resources. This enhancement directly translates into task-dependent improvements in accuracy on classical time series prediction benchmarks characterized by strong nonlinearities. The observed performance gain can be attributed to the combination of multiphoton inputs and adaptive dynamical feedback, which increases the reservoir's sensitivity to variations in the parameters encoding of the input sequence. Such increased sensitivity can be interpreted in terms of higher Fisher Information associated with indistinguishable two-photon states, which makes the output probabilities more responsive to changes in the encoded data. This has been largely proven in prior adaptive phase estimation experiments carried out on the same platform[40,51,55], and investigating this connection further represents a promising direction for future research.

Additionally, we performed numerical simulations, which further demonstrate that increasing the number of photons and the size of the photonic circuit can, in principle, enhance memory capacity and expressivity. However, the advantages associated with enlarging the Hilbert space are strongly task-dependent and can be masked by saturation effects when the benchmark does not require additional degrees of freedom. In particular, for relatively simple function-reconstruction tests such as the monomial and polynomial benchmarks considered here, performance improves only up to the point where the reservoir's effective feature space and feedback-induced dynamics become sufficient, after which different scaled configurations converge to comparable accuracy. For this reason, larger photonic reservoirs should be evaluated on more demanding learning tasks and protocols, which require simultaneously long-range temporal structure and higher-order nonlinear transformations so that the additional expressivity provided by multiphoton indistinguishability and the richer memory dynamics available in larger circuits can translate into measurable computational gains.

Overall, our results provide experimental evidence that quantum correlations and multiphoton states enrich the structure of the reservoir's dynamics in a meaningful way, learning more effectively temporal dependencies when implementing adaptive feedback loops. Notably, this enhancement emerges already when employing two-photon states in a four-mode device, demonstrating the possibility of having an advantage even when operating in a relatively low-dimensional space.

Methods

Digital linear layer

In quantum reservoir computing, a digital linear layer is responsible for the training of the learning model. This layer takes as input the high-dimensional output features of the model, **X**, and uses linear regression to

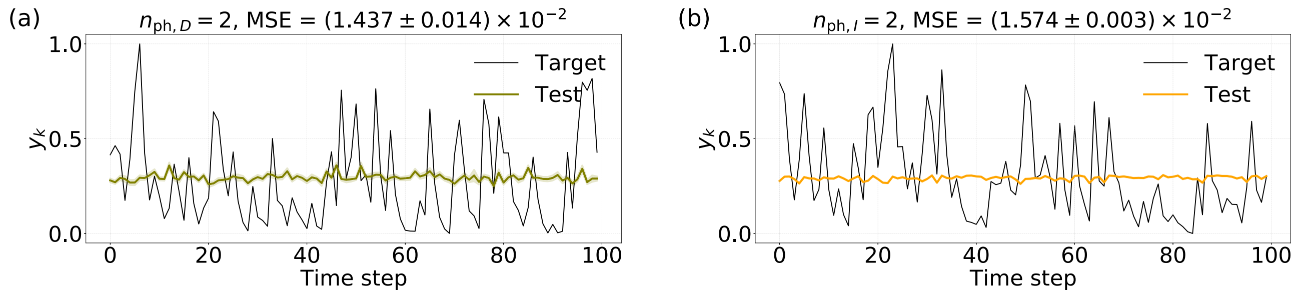

Fig. 5 | Performance of NARMA-5 task using reservoir without feedback-loop mechanism. **a** Output prediction (dark green) versus target (black) for the configuration with $n_{\mathrm{ph},D} = 2$. **b** Output prediction (orange) versus target for the configuration with $n_{\mathrm{ph},I} = 2$. In both cases, the datasets are reshuffled to eliminate any temporal correlations, thus preventing the reservoir from retaining memory of past inputs. The resulting mismatch confirms that memory is essential for solving time-dependent tasks. Both the test and the target outputs y_k are min-max normalized to the range [0, 1] to ensure a fair comparison between $n_{\mathrm{ph},D} = 2$ and $n_{\mathrm{ph},I} = 2$. Shaded areas in the plot are obtained from Monte Carlo Poisson resampling of the experimental photon-count distributions (100 realizations) and represent one standard deviation of the resulting performance distribution.

adjust the weights of a matrix W and approximate the target $\mathbf{y}$. Specifically, we use the linear model Ridge from the Python library *sklearn* to perform linear regression efficiently. In this framework, where the prediction function is defined as $f(\mathbf{p}_k, W) = \mathbf{p}_k^T W$ with W being the vector of trainable weights, the training process consists in minimizing a regularized least-squares loss function:

$$\min_W \sum_{k=1}^{K} \left(f(\mathbf{p}_k, W) - y_k\right)^2 + \alpha ||W||^2, \tag{10}$$

with α the regularization strength and $||W||^2$ the squared l_2-norm of the weights. Using the compact notation $\mathbf{X} = (\mathbf{p}_1, \ldots, \mathbf{p}_K)^T$ for the data matrix, and $\mathbf{y} = (y_1, \ldots, y_K)^T$ for the target vector, the least-squares error becomes $|\mathbf{X}W - \mathbf{y}|^2$. When the regularization parameter α goes to zero, the solution of such a minimization is $W = \mathbf{X}^{+}\mathbf{y}$, where $\mathbf{X}^{+}$ is the Moore-Penrose pseudoinverse of $\mathbf{X}$, and, if the inverse exists, $\mathbf{X}^{+} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$. For $\alpha \neq 0$ it becomes:

$$W = (\mathbf{X}^T\mathbf{X} + \alpha I)^{-1}\mathbf{X}^T y, \tag{11}$$

where I is the identity matrix. Specifically, the role of the regularization term is to prevent overfitting by penalizing large weight values, with its strength controlled by the hyperparameter $\alpha \in [0, \infty)$.

Another hyperparameter considered in the training procedure is the *washout*, which refers to the initial number of time steps during which the output of the reservoir is discarded and not used for training. This is done because, at the beginning of a sequence $\{s_k\}$, the reservoir state is more influenced by the initial conditions than by the input-driven dynamics. Discarding these initial steps exploits the fading memory property of the reservoir and allows the system to evolve into a dynamic regime that is independent of its initial state. Consequently, the training for each task is described in sec. "Experimental results" are carried out by optimizing both α and the *washout* using Optuna. The search ranges for the optimization are set to $\alpha \in [10^{-25}, 10^{-1}]$ and *washout* $\in$ [3, 50]. The α range is chosen empirically to focus on regions associated with better performance, while the *washout* range is selected based on the short-term memory analysis shown in Fig. 2a, ensuring that the amount of training data is not excessively reduced.

Additional details on the feedback

We explore two feedback configurations. In the first configuration, the feedback loops applied to the phases ϕ_D and ϕ_4 depend only on the immediate past output, specifically at time step $k-1$. In the second configuration, the feedback incorporates a longer temporal history: the feedback applied to ϕ_D still depends on the output at time step $k-1$, while the feedback applied to ϕ_4 is based on the output at time step $k-2$. We have observed that, in the nonlinear regime, such as in the reconstruction of nonlinear functions, the expressivity of the system benefits from a feedback loop that relies only on the most recent step $k-1$. However, when the task places greater demands on temporal information processing, that is, when the model requires memory, performance improves by including information from earlier steps. In such cases, extending the feedback to incorporate additional past outputs enhances the system's ability to capture temporal dependencies.

We further investigate the role of the feedback loop in the reservoir dynamics by evaluating performance when it is removed. As already shown in Fig. 2a, removing the feedback loop in a task that requires memory completely cancels the system's memory capabilities. Here, we extend this analysis to a task requiring both expressivity and memory, that is, the NARMA tasks. To this end, we use the same experimental datasets described in sec. "Experimental machine learning benchmarks," but apply a reshuffling of the input sequences. The prediction results obtained using two reservoir configurations on the reshuffled datasets are shown in Fig. 5. In both cases, since the absence of any temporal structure, the reservoir is no longer able to reproduce the nonlinear dynamics required by the NARMA tasks, resulting in no agreement with the target signals. These results highlight the crucial role of feedback and memory in enabling the reservoir to model temporal functions.

Data availability

All data generated and/or analyzed during the current study are not publicly available due to the large volume of raw experimental data, but are available from the corresponding author on reasonable request.

Code availability

The code used is available from the corresponding author upon reasonable request.

Received: 29 September 2025; Accepted: 2 April 2026;
Published online: 17 April 2026

References

1. Raghu, M., Poole, B., Kleinberg, J., Ganguli, S. & Sohl-Dickstein, J. On the expressive power of deep neural networks. In *International Conference on Machine Learning*. 2847–2854 (PMLR, 2017).
2. Patterson, D. et al. Carbon emissions and large neural network training. Preprint at https://arxiv.org/abs/2104.10350 (2021).
3. Marković, D., Mizrahi, A., Querlioz, D. & Grollier, J. Physics for neuromorphic computing. *Nat. Rev. Phys.* **2**, 499–510 (2020).
4. Schuman, C. D. et al. Opportunities for neuromorphic computing algorithms and applications. *Nat. Comput. Sci.* **2**, 10–19 (2022).
5. Mehonic, A. et al. Roadmap to neuromorphic computing with emerging technologies. *APL Mater.* **12** https://arxiv.org/abs/2407.02353 (2024).

6. Mead, C. Neuromorphic electronic systems. *Proc. IEEE* **78**, 1629–1636 (1990).
7. Pal, A. et al. An ultra energy-efficient hardware platform for neuromorphic computing enabled by 2D-TMD tunnel-FETs. *Nat. Commun.* **15**, 3392 (2024).
8. Göltz, J. et al. Fast and energy-efficient neuromorphic deep learning with first-spike times. *Nat. Mach. Intell.* **3**, 823–835 (2021).
9. Xu, H. et al. A low-power vertical dual-gate neurotransistor with short-term memory for high energy-efficient neuromorphic computing. *Nat. Commun.* **14**, 6385 (2023).
10. Du, C. et al. Reservoir computing using dynamic memristors for temporal information processing. *Nat. Commun.* **8**, 2204 (2017).
11. Spagnolo, M. et al. Experimental photonic quantum memristor. *Nat. Photon.* **16**, 318–323 (2022).
12. Schuld, M. & Killoran, N. Quantum machine learning in feature Hilbert spaces. *Phys. Rev. Lett.* **122**, 040504 (2019).
13. Fujii, K. & Nakajima, K. in *Reservoir Computing: Theory, Physical Implementations, and Applications.* 423–450 (eds. Nakajima, K., Fischer, I.) (Springer, Singapore, 2021).
14. Mujal, P. et al. Opportunities in quantum reservoir computing and extreme learning machines. *Adv. Quant. Technol.* **4**, 2100027 (2021).
15. Ghosh, S., Nakajima, K., Krisnanda, T., Fujii, K. & Liew, T. C. Quantum neuromorphic computing with reservoir computing networks. *Adv. Quant. Technol.* **4**, 2100053 (2021).
16. Innocenti, L. et al. Potential and limitations of quantum extreme learning machines. *Commun. Phys.* **6**, 118 (2023).
17. Xiong, W. et al. On fundamental aspects of quantum extreme learning machines. *Quant. Mach. Intell.* **7**, 20 (2025).
18. Fujii, K. & Nakajima, K. Harnessing disordered-ensemble quantum dynamics for machine learning. *Phys. Rev. Appl.* **8**, 024030 (2017).
19. Senanian, A. et al. Microwave signal processing using an analog quantum reservoir computer. *Nat. Commun.* **15**, 7490 (2024).
20. Yasuda, T. et al. Quantum reservoir computing with repeated measurements on superconducting devices. Preprint at https://arxiv.org/abs/2310.06706 (2023).
21. Bravo, R. A., Najafi, K., Gao, X. & Yelin, S. F. Quantum reservoir computing using arrays of Rydberg atoms. *PRX Quantum* **3**, 030325 (2022).
22. Kornjača, M. et al. Large-scale quantum reservoir learning with an analog quantum computer. Preprint at https://arxiv.org/abs/2407.02553 (2024).
23. Paparelle, I. et al. Experimental memory control in continuous-variable optical quantum reservoir computing. *Nature Photonics*. 1-8 https://doi.org/10.1038/s41566-026-01880-9 (2026).
24. Cimini, V. et al. Large-scale quantum reservoir computing using a Gaussian boson sampler. Preprint at https://arxiv.org/abs/2505.13695 (2025).
25. García-Beni, J., Giorgi, G. L., Soriano, M. C. & Zambrini, R. Scalable photonic platform for real-time quantum reservoir computing. *Phys. Rev. Appl.* **20**, 014051 (2023).
26. García-Beni, J., Giorgi, G. L., Soriano, M. C. & Zambrini, R. Squeezing as a resource for time series processing in quantum reservoir computing. *Opt. Express* **32**, 6733–6747 (2024).
27. Zia, D. et al. Quantum extreme learning machines for photonic entanglement witnessing. *Sci. Adv.* **11**, https://doi.org/10.1126/sciadv.ady7987 (2025).
28. Nerenberg, S., Neill, O. D., Marcucci, G. & Faccio, D. Photon number-resolving quantum reservoir computing. *Opt. Quantum* **3**, 201–210 (2025).
29. Suprano, A. et al. Experimental property reconstruction in a photonic quantum extreme learning machine. *Phys. Rev. Lett.* **132**, 160802 (2024).
30. Kobayashi, K., Fujii, K. & Yamamoto, N. Feedback-driven quantum reservoir computing for time-series analysis. *PRX Quantum* **5**, 040325 (2024).
31. Selimović, M. et al. Experimental neuromorphic computing based on quantum memristor. Preprint at https://arxiv.org/abs/2504.18694 (2025).
32. Joly, M. et al. Harnessing photon indistinguishability in quantum extreme learning machines. Preprint at https://arxiv.org/abs/2505.11238 (2025).
33. Yin, Z. et al. Experimental quantum-enhanced kernel-based machine learning on a photonic processor. *Nat. Photonics* https://doi.org/10.1038/s41566-025-01682-5 (2025).
34. Schuld, M., Sweke, R. & Meyer, J. J. Effect of data encoding on the expressive power of variational quantum-machine-learning models. *Phys. Rev. A* **103**, 032430 (2021).
35. Jaeger, H. The "echo state" approach to analysing and training recurrent neural networks-with an erratum note. *13 GMD Technical Report 148* (German National Research Center for Information Technology, Bonn, Germany, 2001).
36. Jaeger, H. Short term memory in echo state networks. *Fraunhofer Institute for Autonomous Intelligent Systems Technical Report 152* (GMD Forschungszentrum Informationstechnik, 2002).
37. Khan, S. A., Hu, F., Angelatos, G. & Türeci, H. E. Physical reservoir computing using finitely-sampled quantum systems. Preprint at https://arxiv.org/abs/2110.13849 (2021).
38. Hu, F. & Gao, X. Generalization error in quantum machine learning in the presence of sampling noise. Preprint at https://arxiv.org/abs/2410.14654 (2024).
39. Corrielli, G., Crespi, A. & Osellame, R. Femtosecond laser micromachining for integrated quantum photonics. *Nanophotonics* **10**, 3789–3812 (2021).
40. Valeri, M. et al. Experimental multiparameter quantum metrology in adaptive regime. *Phys. Rev. Res.* **5**, 013138 (2023).
41. Akiba, T., Sano, S., Yanase, T., Ohta, T. & Koyama, M. Optuna: a next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Association for computing machinery, 2019).
42. Stepney, S. Physical reservoir computing: a tutorial. *Nat. Comput.* **23**, 665–685 (2024).
43. Siemaszko, M., Buraczewski, A., Le Saux, B. & Stobińska, M. Rapid training of quantum recurrent neural networks. *Quant. Mach. Intell.* **5**, 31 (2023).
44. García-Beni, J. et al. Quantum machine learning via continuous-variable cluster states and teleportation. *EPJ Quant. Tech.* **12**, 63 (2025).
45. Atiya, A. F. & Parlos, A. G. New results on recurrent network training: unifying the algorithms and accelerating convergence. *IEEE Trans. Neural Netw.* **11**, 697–709 (2000).
46. Gan, B. Y., Leykam, D. & Angelakis, D. G. Fock state-enhanced expressivity of quantum machine learning models. *EPJ Quantum Technol.* **9**, 16 (2022).
47. Cerezo, M. et al. Variational quantum algorithms. *Nat. Rev. Phys.* **3**, 625–644 (2021).
48. Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 4213 (2014).
49. Agresti, I. et al. Demonstration of hardware efficient photonic variational quantum algorithm. Preprint at https://arxiv.org/abs/2408.10339 (2024).
50. Hoch, F. et al. Variational approach to photonic quantum circuits via the parameter shift rule. *Phys. Rev. Res.* **7**, 023227 (2025).
51. Cimini, V. et al. Variational quantum algorithm for experimental photonic multiparameter estimation. *npj Quant. Inf.* **10**, 26 (2024).
52. McClean, J. R., Boixo, S., Smelyanskiy, V. N., Babbush, R. & Neven, H. Barren plateaus in quantum neural network training landscapes. *Nat. Commun.* **9**, 4812 (2018).
53. Sannia, A., Giorgi, G. L. & Zambrini, R. Exponential concentration and symmetries in quantum reservoir computing. Preprint at http://arxiv.org/abs/2505.10062 (2025).
54. Xiong, W. et al. Role of scrambling and noise in temporal information processing with quantum systems. Preprint at https://arxiv.org/abs/2505.10080 (2025).

55. Cimini, V. et al. Deep reinforcement learning for quantum multiparameter estimation. *Adv. Photonics* **5**, 016005 (2023).

Acknowledgements

This work is supported by the MUR PNRR project Spoke 4 and Spoke 7 (Grant No. PE0000023-NQSTI), the ERC Advanced Grant QU-BOSS (QUantum advantage via non-linear BOSon Sampling, Grant No. 884676), and the European Union's Horizon Europe research and innovation program under the EPIQUE Project (Grant No. 101135288). V.C. and F.S. acknowledge support from MUR FARE Ricerca in Italia QU-DICE (Grant No. R20TRHTSPA).

Author contributions

V.C. and F.S. conceived the experiment. R. Di B., V.C. and F.S. carried out the experiment and performed the data analysis. S.P., F.C., G.C. and R.O. fabricated the integrated photonic device and performed its initial characterization. All the authors discussed the results and contributed to the writing of the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41534-026-01236-9.

Correspondence and requests for materials should be addressed to Valeria Cimini.