Scalable Photonic Platform for Real-Time Quantum Reservoir Computing

Jorge García-Beni, Gian Luca Giorgi, Miguel C. Soriano, and Roberta Zambrini*
Instituto de Física Interdisciplinar y Sistemas Complejos (IFISC), UIB–CSIC UIB Campus, Palma de Mallorca E-07122, Spain

(Received 28 July 2022; revised 5 April 2023; accepted 6 July 2023; published 24 July 2023)

Quantum reservoir computing (QRC) exploits the information-processing capabilities of quantum systems to solve nontrivial temporal tasks, improving over their classical counterparts. Recent progress has shown the potential of QRC exploiting the enlarged Hilbert space, but real-time processing and the achievement of a quantum advantage with efficient use of resources are prominent challenges towards viable experimental realizations. In this work, we propose a photonic platform suitable for real-time QRC based on a physical ensemble of reservoirs in the form of identical optical pulses recirculating through a closed loop. While ideal operation achieves maximum capacities, statistical noise is shown to undermine any quantum improvement. We propose a strategy to overcome this limitation and sustain the QRC performance when the size of the system is scaled up. The protocol is conceived for experimental implementations to be viable with current technology.

DOI: 10.1103/PhysRevApplied.20.014051

I. INTRODUCTION

Quantum photonic technologies are currently being exploited for quantum communications, quantum computing, and quantum information processing [1,2], for their speed-of-light propagation, ultrafast operations and gates, and weak interactions with the environment even at room temperature. In the context of computation, photonic quantum computing for boson sampling has shown a time advantage of 14 orders of magnitude over today's classical supercomputers [3,4]. In measurement-based one-way quantum computing in *continuous-variable* (CV) regimes cluster states up to one million modes have been reported using time multiplexing [5] and reconfigurable cluster states achieved with frequency multiplexing [6]. Variational eigensolvers implemented in photonic quantum processors [7] have achieved much more efficient use of quantum resources than alternative algorithms, such as quantum phase estimation. Quantum states of light have also been used in Ising machines, proving more efficient than current algorithms in certain scenarios [8,9] and complementing the results obtained for alternative classical approaches [10,11]. In this work, we propose a quantum photonic approach to time series processing designing and addressing the scalability, memory and performance for time-series prediction of an optical setup in the quantum reservoir computing (QRC) framework.

Reservoir computing (RC) is a paradigm of machine learning in which the information-processing capabilities of dynamical systems are exploited for solving temporal tasks, with real-world examples such as the prediction of monthly electricity production [12], and financial [13,14] or water-level forecasting [15,16]. In addition, RC can also be applied to solve static tasks, such as the classification of phonemes [17] or the detection of human finger movements from EEG data [18]. In practice, RC can solve these information-processing tasks without the need for an external memory thanks to the fading memory present in the internal state of the reservoir itself [19]. Since RC exploits generic dynamical systems for computing, the concept of RC has been successfully transferred to physical substrates [20], with the prominent example of high-speed photonic and optoelectronic implementations [21–24]. Ultimately, RC has been generalized to the quantum regime in order to benefit from the large number of degrees of freedom available in quantum systems [25,26]. In order to experimentally achieve time-series processing with superior performance in quantum reservoir computing with respect to classical approaches, several challenges need to be addressed, identifying the most promising applications, efficient platform designs and dealing with quantum measurement retaining improvements over classical analogues [25,27,28].

With respect to classical reservoir computing, where single measurements on the reservoir produce the relevant information at the output layer, when moving into the quantum realm, one usually extracts the expectation values of observables at the output, from large ensembles

*roberta@ifisc.uib-csic.es

of experiment copies. In the pioneering experimental implementation of quantum reservoir computing in *noisy intermediate-scale quantum* (NISQ) platforms [29], this was achieved by repeating the input processing sequence several times, being this an obstacle towards viable real-time technological realizations. Our goal is to show a strategy to move to temporal signal processing, operating in a continuous way and without buffering inputs in external memories. The approach can be adapted to different photonic platforms and takes advantage of light-speed propagation and fast operation to monitor the reservoir processor. Optical sources producing high repetition pulses and optical fibers allow one to design an ensemble of identical reservoirs inside a closed loop, which removes the necessity for an external classical memory. The reservoir signal is continuously driven by the external inputs and monitored through a beam splitter and homodyne detection, thus obtaining expected values of observables without external buffers. We are restricting the analysis here to vacuum Gaussian states, as we know they can provide universal RC [30]. These reservoir states can be engineered as complex networks in the frequency domain [6,31–33].

After introducing the photonic platform and main tools for quantum reservoir computing in Sec. II, we address its memory capabilities (Sec. III) both in the ideal case of an infinite ensemble (Sec. III A) where statistical errors vanish and in a realistic scenario of a finite ensemble assessing the limitations of statistical noise (Sec. III B). The analysis of noise detrimental effects in the resolution of past inputs allows identification of strategies to improve the performance when the reservoir size is scaled up (Sec. III C). We also address the performance of our proposal for chaotic time-series prediction (Sec. III D).

II. PHOTONIC PLATFORM

A. Platform description

Before introducing the photonic setup for QRC, let us start by recalling the main features of reservoir computing for time-series processing. RC schemes [15] consist of three main layers: the input, the reservoir, and the output. First, an input signal, typically belonging to a time series, is injected into the reservoir, a dynamical system (often a recurrent neural network) that performs a complex nonlinear transformation to the injected data. Then, in the readout layer, a certain number of reservoir observables are measured and their combination (usually linear) is optimized to match the desired target, depending on the temporal task, like, e.g., linear memory or chaotic series prediction (see Appendices A and B for further details).

The RC paradigm has been recently extended to quantum reservoirs [25,34–36] showing a quantum improvement due to the enlarged Hilbert space and therefore to the increased processing capability. Nevertheless, this approach also presents several challenges to be viable experimentally. In order to act as an online time-series processor, the reservoir needs to continuously receive the input data and produce output extracted for the desired task. The relevant information for the readout layer is generally encoded in the expected values of the observables and due to the stochastic nature of quantum measurements, several detections are needed. Furthermore, the act of measuring yields backaction into the system that may also negatively affect the performance of the reservoir. Two strategies to overcome this last issue in qubit platforms have been recently proposed in Ref. [28] using weak measurements and partial sequence repetition (rewinding). Otherwise one needs to restart the protocol for each input injection, buffering the input sequence as in Ref. [29]. In the following, we propose a photonic platform design for QRC offering fast and easily scalable operation accounting for measurements. We have conceived the protocol to be experimentally feasible with state-of-the-art technologies.

The proposed design is based on continuous variables and revolves around the use of optical pulses, whose individual dynamics along the feedback loop defines the reservoir computer [see scheme in Fig. 1(a)]. Indeed the reservoir is a traveling pulse, containing N modes whose interaction is mediated by a nonlinear ($\chi^{(2)}$) crystal and is modeled as a complex network [bottom inset in Fig. 1(a)]. The input is encoded in each ancilla pulse [see top inset in Fig. 1(a)] coupled with the reservoir through a beam splitter (BS) and is prepared in a product state of N Gaussian squeezed vacuum states. We consider their squeezing angles as the classical inputs as this guarantees a good QRC performance [30]; alternatively, the ancilla state could be directly treated as a quantum input itself to be processed, providing quantum (instead of classical) information [25,37]. The reservoir pulse travels in a closed feedback loop, that can be enabled by an optical fiber. This feedback allows information from previous inputs to be retained, an essential requirement for temporal series processing. At the other output arm of the BS, the reflected input and transmitted reservoir signals are instead detected, after interacting with a nonlinear crystal. We are considering for this purpose a *homodyne detector* [HD in Fig. 1(a)], common in CV quantum optical experiments whose bandwidth continuous improvements allow for much faster detections [38,39].

A crucial feature of the setup presented in Fig. 1 is the possibility to delay several reservoir pulses enabling to run the experiment with M copies of the pulses. This provides the needed ensemble to effectively realize quantum measurements without restarting the protocol by reinjecting

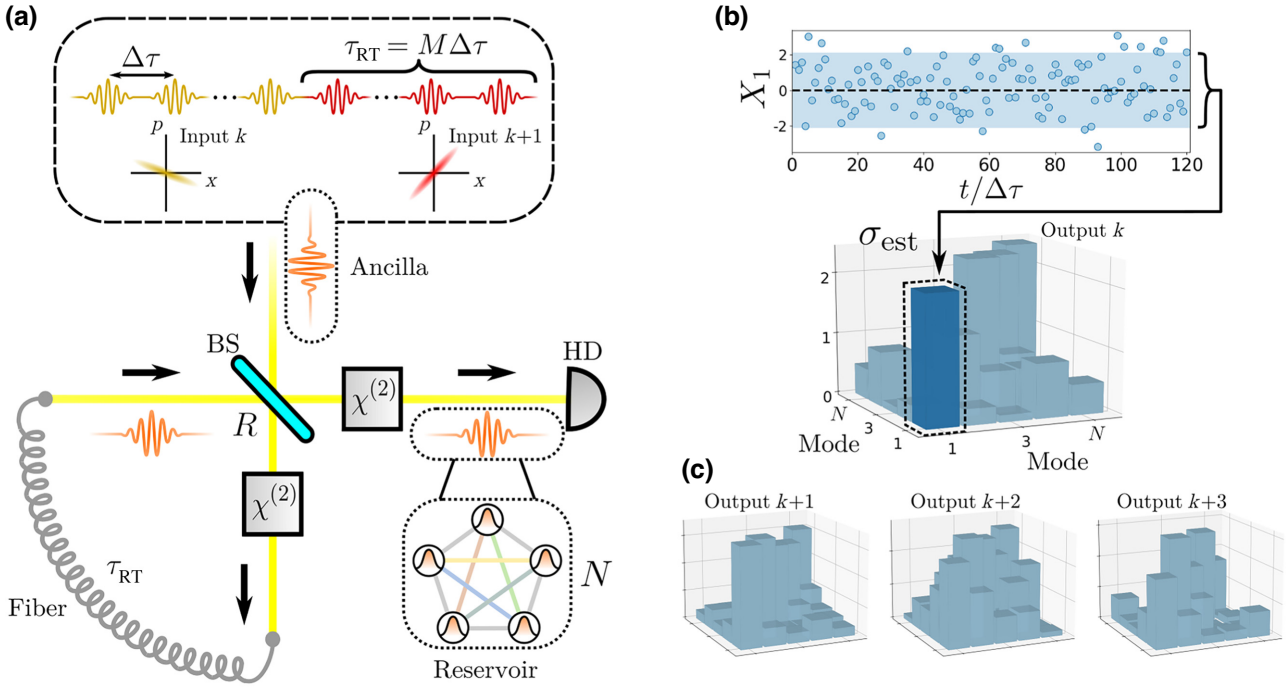

FIG. 1. Scheme of the process: (a) schematical drawing of the whole platform; (b) example of a series of the measured x quadrature of the first mode for each pulse (horizontal axis) and a 3D representation of the estimated covariance matrix for the kth round trip (covariance matrix representations for the following round trips are depicted below it). In this case we have taken $N = 5$, $M = 5000$, and $R = 0.75$. The horizontal axes represent the position of each term in σ_{est}. (c) Sequence of estimated σ_{est} for three consecutive inputs.

the input sequence. Indeed, the same input is encoded in M ancilla states before being updated [upper inset in Fig. 1(a)] and every pulse interacts with a matching reservoir pulse at a given round trip. This creates a physical ensemble of M copies of the same reservoir in real time, whose expected values are obtained via homodyne detection. High optical repetition rates and fast processing determine the potential of this scheme for sequential input processing while the signal is continuously injected. Beyond the reservoir size (the number of modes N in each pulse) and the ensemble size (M input repetitions), other crucial parameters determining the reservoir features are the BS reflectivity R (we assume partial transmission, i.e., $R \neq 0, 1$) and the time interval between pulses $\Delta\tau$ (the time interval that a whole round trip lasts is $\tau_{\text{RT}} = M\Delta\tau$). At the kth round trip, the injected pulses are squeezed vacuum states that have the input s_k encoded in their squeezing angle. We focus on CV Gaussian states and then model the effect of each element through the *covariance matrix* of the external pulses (see Appendix C for details). In particular, the BS couples the feedback and the external signal, generally entangling the feedback loop pulses with the ones traveling to the detector. The $\chi^{(2)}$ crystals induce squeezing and correlations between the modes in each pulse and are modeled by two (in principle different even if we omit a distinguishing label) Hamiltonians $\hat{H}_{\chi^{(2)}} = \hat{H}_{\text{free}} + \hat{H}_{\text{int}}$ with

$$\hat{H}_{\text{free}} = \sum_{i=1}^{N} \omega_i \hat{a}_i^\dagger \hat{a}_i$$
$$\hat{H}_{\text{int}} = \sum_{j>i} \left(g_{ij} \hat{a}_i^\dagger \hat{a}_j + i h_{ij} \hat{a}_i^\dagger \hat{a}_j^\dagger + \text{h.c.} \right) \qquad (1)$$

where $\hat{a}_i$ ($\hat{a}_i^\dagger$) is the annihilation (creation) operator of mode i; ω_i is its frequency and parameters g_{ij} and h_{ij} depend on the second-order nonlinearities of the crystals and are not tuned but rather assumed to be random real numbers and different for each one of the two crystals. We assume that these crystals have a high enough bandwidth to keep the entering pulses independent of one another, and only couple the modes inside each pulse. The time each pulse spends inside a $\chi^{(2)}$ material is labeled as Δt (equal for both crystals).

For the measurement scheme we propose multimode homodyne detection: we measure, at the same time, the same quadrature operator for every mode. That is, for every incoming pulse the detector measures all x quadratures $\hat{x}_i = \hat{a}_i + \hat{a}_i^\dagger$ $(i = 1, \ldots, N)$, obtaining a set of stochastic outcomes, X_i, defining a quantum trajectory [40,41]. Thus, the measurement output of the mth pulse at the kth round trip is given by the vector $\mathbf{X}^{(k,m)} =$

$(X_1^{(k,m)},\ldots,X_N^{(k,m)})^\top$ and, from these quadrature measurements, we can access the expected values of the covariance matrix, fully characterizing the squeezed vacuum states at the output for our readout layer. That is, for each round trip (and thus, each input in the sequence) we extract $N(N+1)/2$ outputs for the readout layer $O_{ij}^{(k)} = \langle X_i^{(k)} X_j^{(k)}\rangle_M - \langle X_i^{(k)}\rangle_M \langle X_j^{(k)}\rangle_M$ with $j \geq i = 1,\ldots,N$ and with $\langle\cdot\rangle_M$ corresponding to the ensemble average over M pulses. These outputs give the *estimated* covariance matrix $\left[\sigma_{\text{est}}^{(k)}\right]_{ij} = O_{ij}^{(k)}$ at each kth round trip, in which the input s_k is injected. While here we restrict to the N-position quadrature (and homodyne detection), one could also consider momenta (i.e., $2N$ quadratures) through heterodyne detection. Higher-order moments could also be considered as features for the readout layer to improve the performance in some temporal tasks (see Sec. III D). For Gaussian states, these higher-order moments are functions of the first- and second-order ones. Here we restrict our analysis to the covariance matrix for the linear and nonlinear memory. In Fig. 1(b), a visual representation of this readout protocol is shown. The outcome measurements of the M sequence of $X_1^{(k,m)}$ at an unspecified step k are shown and the black dashed line is the mean, vanishing because we are working with vacuum states. The (blue shaded) standard deviation of X_1 provides $[\sigma_{\text{est}}]_{1,1}$ shown below, and, considering all N modes, one estimates the covariance matrix. The covariance matrix at each time step k constitutes the output layer [Fig. 1(c)], optimized by a simple linear regression (see Appendix A) to achieve the best performance in the desired task. We notice that a measurement in one of the beam-splitter outputs generally influences the conditional state of the other beam (being these generally entangled), but we consider here only ensemble-averaged quantities for Gaussian states when these expected values correspond to the unconditional quantities.

The design proposed in Fig. 1(a) is reminiscent of other photonic approaches for RC based on time-delay feedback and time multiplexing [21,24]. In analogy to these works, the feedback loop provides the mechanism for fading memory in the optical implementation. The time multiplexing instead has a different role: in our platform it has the distinctive function to provide the ensemble for quantum measurements of the covariance matrix, while in classical approaches it serves as a way to increase the dimensionality of the reservoir. The reservoir size in the current proposal depends instead on the number of (frequency or spatial) modes contained in each temporal pulse. Let us remark that the main novelties of our proposal are both the feedback loop, providing memory to the quantum system, and the physical ensemble stored in the optical fiber. This allows us to compute averages in real time. Our protocol can be applied to any quantum state, even not Gaussian, but the Gaussian states are suitable for a deeper analytical treatment based on the symplectic formalism (see Appendix C 2).

B. Feasibility of the photonic design

In this section, we discuss the viability of experimental implementations of our proposal with state-of-the-art technology. The main aspects of interest are the engineering of the input ancilla states, the generation of complex networks using $\chi^{(2)}$ nonlinearities, the multimode detection of the output pulses (measuring all the modes at once) and the losses due to the fiber propagation.

Input states' generation: in the frequency domain, squeezed states can be deterministically generated at high rates for a wide spectrum band. Specifically, over 10^8 pulses per second containing up to 21 squeezed spectral modes each have been recently attained [42]. The ability to encode the inputs in the squeezing phase of each pulse is also required for the feasibility of our platform. Highly accurate and versatile phase-setting devices have been demonstrated using programmable phase shifters in Xanadu's Borealis platform [43]. Accurately changing the squeezing levels and phases of pulses at time intervals below 100 ns has also been attained recently by continuously modulating the pump light of a waveguide optical parametric amplifier [44].

Optical networks: network topologies similar to the ones considered in Eq. (1) are experimentally attainable in the frequency domain [6,31,32,42,45]. Reconfigurability in the spectral network shapes can also be reached by changing the measurement basis of the homodyne detection [6,42].

Homodyne detection: being able to measure several frequency modes at once is another key requirement for the real-time processing capabilities of our platform. Multimode homodyne detection of several frequency bands has been demonstrated in experimental setups [46,47].

Fiber losses: retaining the ensemble pulses inside a long optical fiber without great losses is another relevant aspect to tackle for the experimental implementation of our real-time processing proposal. Low-loss fiber delay lines have been successfully implemented in several CV optical setups [5,8,9,43,48], ranging from single pulse-delay lines [5,48] to over 10^5 recirculating pulses in a 5-km fiber [9]. So fiber losses do not constrain the viability of our platform with state-of-the-art technology.

In summary, the key aspects of our suggested platform have already been demonstrated experimentally. The main challenge ahead is to combine them in a common realization.

III. RESULTS

In this section, we present a detailed analysis of the information-processing capabilities of the quantum

photonic platform for time-series processing. RC systems require the ability to retain the memory of previous signal inputs, the capacity to reproduce nonlinear functions of the inputs and high dimensionality, which requires that the readout observables are linearly independent functions of the inputs. For the evaluation of the reservoir memory, we numerically compute the *linear capacity* of the photonic QRC, which measures how accurately can the system reproduce inputs in the past. It is a normalized output-target correlation that ranges from 0 to 1. The higher the value of the linear capacity, the better the target input is reproduced by the reservoir. To test both the nonlinearity as well as the high dimensionality, we make use of the total *information-processing capacity* (IPC) [49], which generalizes the linear capacity memory to nonlinear contributions (see Appendix B for a detailed explanation). It quantifies the expressivity of a wide variety of dynamical systems and has recently been used in quantum settings [30,50]. The IPC ranges from 0 to the number of output observables (upper bound). When this upper bound is reached, it means that all the readout observables are linearly independent. In our case, as we are using the covariance matrix of the measured x quadratures, the number of observables ($\langle \hat{x}_i \hat{x}_j \rangle$) is equal to $N(N+1)/2$ (N denoting the reservoir size, or the number of modes in each pulse).

In the following, we consider the memory capacity of the proposed photonic QRC both in ideal conditions and when explicitly accounting for statistical noise. In particular, we consider the N mode frequencies to be equal and scale time so that $\omega_i = 1 \quad \forall i$ in Eq. (1). This is a reasonable approximation when considering modes in the frequency domain, as usually in experiments with frequency combs the detuning is much smaller than the coupling strengths between modes [6,45]. The coupling parameters g_{ij} and h_{ij} have been chosen from a uniform distribution in the intervals $[\langle g \rangle - \Delta g, \langle g \rangle + \Delta g]$ and $[\langle h \rangle - \Delta h, \langle h \rangle + \Delta h]$, respectively, where $\langle g \rangle = 0.2$, $\langle h \rangle = 0.3$, and $\Delta g = \Delta h = 0.1$, while the interaction time (Δt) inside each crystal is set to one. These coupling values have been chosen to ensure the stability of the system, that is, the energy of every pulse traveling through each crystal remains bounded for any interaction time values. This is not a necessary condition, but it ensures that the fading memory condition is fulfilled at every realization (see Appendix F for details). The model in Eq. (1) allows us to describe a broad scenario of possible implementations including squeezing and hopping. We have ensured that this Hamiltonian coupling choice does not produce squeezing levels far above what is experimentally attainable. Concretely, the average squeezing has values well within what was produced in the multimode setting [6], with at most 3 dB of squeezing. Furthermore, none of our simulated crystals produce squeezing larger than 6.5 dB, which is feasible in frequency multiplexed quantum optics [6]. We also notice that the frequencies ω_i in Eq. (1) can be considered as detunings with respect to some reference (lab) frequency.

The external inputs s_k are encoded in the ancilla as squeezing angles, providing a nonlinear input encoding [30]; each one of the M pulses in the ancilla train is a squeezed vacuum state with covariance matrix equal to $\sigma_{\text{anc}}^{(k)} = \bigoplus_{i=1}^{N} \sigma_{\text{sq}}^{(k)}$, where $\sigma_{\text{sq}}^{(k)}$ is the covariance matrix of a single mode squeezed state. In particular, squeezing strength is set to $r_k = 1$ while the squeezing angle encoding is $\phi_k = 3\pi s_k/4$ (further details in Appendix C). It was recently reported how the encoding choice affects the degree distribution of the IPC [30] and our choice provides a balanced contribution of the linear terms and the nonlinear ones. Different angle encodings would provide different linear to nonlinear contributions (see Appendix D). The input parameters, even when random, are set to lead to squeezing levels currently viable with state-of-the-art technology, which is around 15 dB [51], equivalent to $r_k \simeq 1.7$.

A. Ideal case

In the ideal case of an infinite number of pulses, the covariance can be obtained as the limit

$$\sigma_{\text{ideal}}^{(k)} = \lim_{M \to \infty} \sigma_{\text{est}}^{(k)}. \tag{2}$$

While not experimentally attainable, this ideal covariance matrix sets the performance of our reservoir in the absence of statistical noise, providing an insight to quantify the effect of this noise in realistic (finite ensemble) scenarios. Furthermore, the ideal performances we present here can be compared with most results in QRC in the literature [25], generally not accounting for experimental limitations.

As a consequence of the linearity of our photonic platform, by inspecting the inputs' dependence of σ_{ideal} (see Appendix E), we observe that inputs at different times contribute additively to the covariance matrix, so we can decompose the covariance as

$$\sigma_{\text{ideal}}^{(k)} = \sum_{d=0}^{\infty} \gamma_d^{(k)}, \tag{3}$$

where the matrices $\gamma_d^{(k)}$ are functions of a single input with delay d. Hence, if we are in the kth round trip, the covariance is a sum of nonlinear functions $\gamma_d^{(k)}$ of the previous inputs s_{k-d} and $\gamma_d^{(k)} : \Re \to \Re^N \times \Re^N$ turns each input into an $N \times N$ positive symmetric matrix (N denoting the size of the reservoir). The reflectivity R of the BS is found to be the physical parameter determining the relative size of

these terms (see Appendix E)

$$\gamma_d^{(k)} \propto (1-R)^2 R^{d-1}, \quad \text{for } d \geq 1, \tag{4}$$

with an exponential decay R^{d-1} in the delay for the average magnitude of the observables. Indeed, each time a feedback pulse gets reflected by the BS in the fiber loop, it is scaled by R. If the signal recirculating in the loop is weak (small reflectivity R), a much faster decay in the magnitude of delayed matrices $\gamma_d^{(k)}$ is observed (when the delay d increases). The BS reflectivity is then expected to have a strong influence on the fading memory in this setup. As for the term $(1-R)^2$ in Eq. (4), it is due to the repeated transmission through the BS of each delayed term: first as an ancilla signal when it enters the loop and then when it leaves the loop towards the detector. Then, on the one hand, increasing the value of the reflectivity R reduces the amount of feedback light reaching the detector; on the other hand, it also increases the memory retention of the reservoir, as the magnitude decay of the delayed matrices in Eq. (3) is slower.

We now present the numerical results for the performance of the QRC in the ideal limit, starting with the *linear capacity*, which brings information about the ability of our system to reproduce the encoded inputs at different times in the past (linear memory). In Fig. 2(a), the linear capacity is shown as a function of the delay, d, of the target input. By increasing the number of modes (from $N = 8$ to $N = 10$) and consequently the size of the output layer, the system can reproduce more delayed inputs, extending the memory. In contrast, the capacity for low delays does not change. Interestingly, tuning the BS reflectivity R alters the shape of the linear capacity curve. Consistently with Eq. (4), for higher values of the BS reflectivity R, further terms into the past can be resolved, although the curve starts to descend earlier from its maximum value, achieved with smaller reflectivity.

The total IPC for different values of the reservoir size N is shown in Fig. 2(b) as a bar plot, where each bar is split into different linear and nonlinear memory contributions (identified by their polynomial degree). Concretely, the linear contributions [blue bars in Fig. 2(b)] correspond to areas below the lines in Fig. 2(a). The dashed blue line—representing the normalized capacity equal to one—shows that the capacity saturates its maximum theoretical value $N(N+1)/2$ for every N, corresponding to the number of output degrees of freedom (terms of the covariance matrix). Actually, our QRC platform is found to display both fading memory and echo state property, which are required for good RC performance (see proof in Appendix F). This explains why the normalized IPC is saturated, as the readout observables are linearly independent and the above properties are fulfilled [49]. Furthermore, this occurs for every value of the reflectivity R that has been tested [e.g., we show the total IPC for both $R = 0.9$ and $R = 0.75$ in Fig. 2(b)]. The degree distribution shown in Fig. 2(b), displaying significant contributions up to cubic degree, also does not change significantly with R, being mainly determined by the encoding choice (see Appendix D for details).

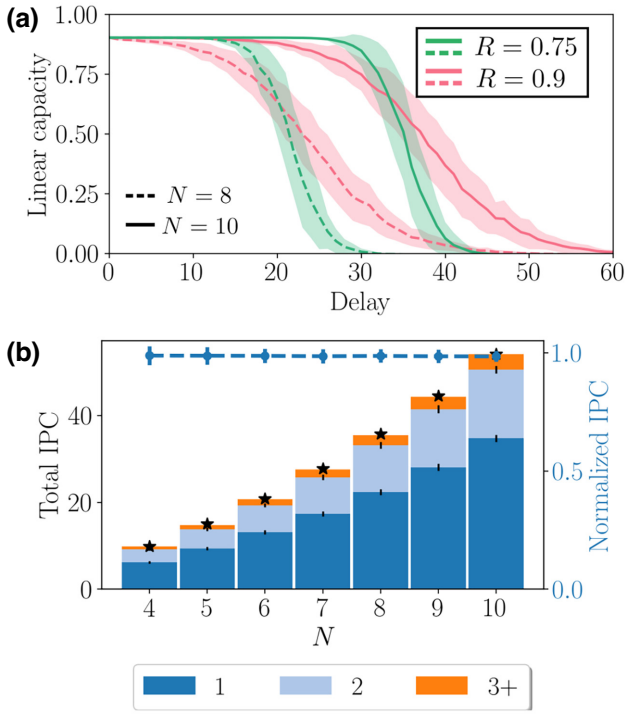

FIG. 2. Capacity in the ideal case: (a) linear capacity as a function of the delay of the input for different values of the BS reflectivity R (color) and different values of the reservoir size N (line); (b) IPC as a function of N for $R = 0.9$. The bars follow the vertical axis at the left, which shows the total absolute capacity. Each bar is also split into different degree contributions. The discontinuous blue line follows the vertical axis at the right, which depicts the normalized IPC. Black stars follow the left axis and correspond to the value of the total IPC for $R = 0.75$, as a function of N. Every plot is the result of averaging over 100 realizations with different random networks in the coupling crystals and different input strings (this average has been performed in every figure of the paper).

B. Finite measurement ensemble

In this subsection, a more realistic scenario is considered, when the physical ensemble is limited by the finite number of pulses, M, inside the loop. Then, any estimation of the observables is affected by statistical errors, whose average magnitude depends on M. Additive noise in the readout layer has been considered in Gaussian CV models for RC with coherent states [52], and for nontemporal tasks [53]. Here we explore statistical noise due to a physical ensemble of reservoirs for temporal tasks and for QRC with squeezed vacuum states. Hence the performance based on estimated covariance elements will be

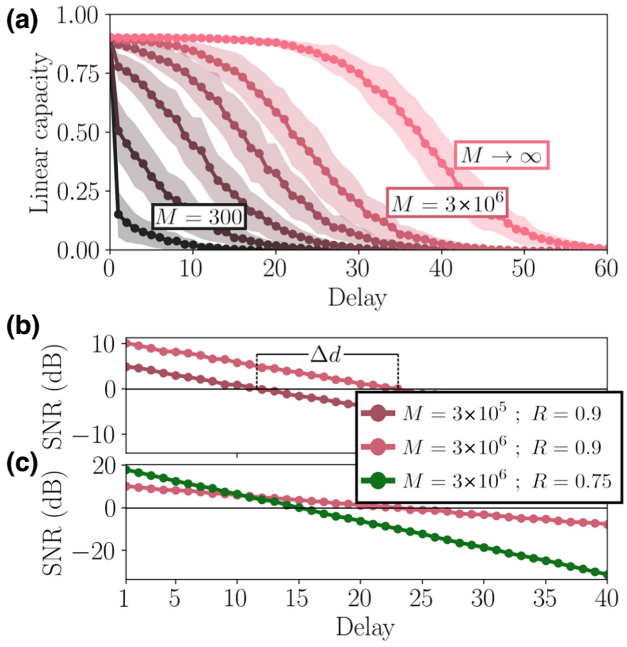

FIG. 3. (a) Linear capacity as a function of the delay for different values of the ensemble size M ranging in powers of 10 from 300 to 3×10^6, as well as the curve from the ideal case ($M \to \infty$). In all curves the reservoir size is equal to $N = 10$ and the reflectivity is $R = 0.9$; (b),(c) SNR (in decibels) for the delayed terms γ_d as a function of the delay (d). In (b) this is shown for fixed $R = 0.9$ and different values of M and in (c) it is shown for fixed M and two values of R. In both figures $N = 10$.

quantified and used to determine how the resources scale in this platform when the reservoir size increases. To simulate the ensemble, we compute the conditional evolution and measurement outcome of every realization (as shown in Appendix C 2) and perform the averages among them. In Fig. 3(a), the linear capacity as a function of the delay is shown in the ideal case and for different values of the ensemble size M, while the reservoir size N and the reflectivity R are kept fixed. The number of delayed inputs a 10-node reservoir can ideally reproduce with good capacity is high (approximately 30 injection steps). The collapse of the linear capacity for finite samples is clearly seen in Fig. 3(a) for finite ensemble sizes M. Even though we show a progressive improvement up to 3×10^6 fiber pulses (with a memory of up to 10 further delays times when increasing the ensemble size by factors of 10), we also see that the ability to reproduce inputs further into the past of the ideal system is hard to reach. The origin of this performance scaling can be traced back to the exponential decay of the delayed terms, $\gamma_d^{(k)}$, in Eq. (4) and is quantified in the following.

In order to be able to reproduce a given input term with delay d, the ratio between its corresponding term γ_d in the ideal covariance and the statistical error should be large enough on average. As the magnitude of γ_d decays exponentially with d, it becomes harder to keep the statistical noise below it, actually requiring an exponential increase of pulses with the delay to resolve further inputs in the past. Scaling with the reservoir size N will also be severely affected by statistical errors as the root of performance improvements by increasing N is in the ability of the system to reproduce further delayed terms (see Sec. III C).

The performance presented in Fig. 3(a) can be quantified distinguishing in the measured observables, σ_{est}, the ideal case contribution, σ_{ideal}, and an added stochastic noise term, ξ_M, which depends on the number of pulses circulating the fiber, M:

$$\sigma_{\text{est}}^{(k)} = \sigma_{\text{ideal}}^{(k)} + \xi_M^{(k)}. \tag{5}$$

The additive stochastic noise term for large ensembles has a standard deviation $\text{std}\left(\xi_M^{(k)}\right) \propto M^{-1/2}$. In contrast, the ideal covariance terms $\gamma_d^{(k)}$ decay exponentially with the delay d [Eq. (4)]. The delay resolution will then be determined by the number of terms $\gamma_d^{(k)}$ whose magnitude is greater than the noise magnitude. How much do we need to increment the ensemble size M in order to maintain equivalent signal resolution at larger delays? This can be determined by constraining the signal-to-noise ratios (SNRs), as shown in Appendix G 1, and leads to an exponential factor in the increment in the number of the pulses

$$\left\langle \left| \frac{\gamma_d^{(k)}}{\xi_M^{(k)}} \right| \right\rangle_{\mathbf{s},\hat{H}} = \left\langle \left| \frac{\gamma_{d+\Delta d}^{(k)}}{\xi_{M'}^{(k)}} \right| \right\rangle_{\mathbf{s},\hat{H}} \longrightarrow \frac{M'}{M} = R^{-2\Delta d}, \tag{6}$$

where $\langle\cdot\rangle_{\mathbf{s},\hat{H}}$ stands for the average among realizations of the reservoir [Eq. (1)] and random input strings. $\gamma_d^{(k)}\xi_M^{(k)-1}$ is a $(N \times N)$ matrix whose elements are $\left[\gamma_d^{(k)}\xi_M^{(k)-1}\right]_{ij} \equiv \left[\gamma_d^{(k)}\right]_{ij}\left[\xi_M^{(k)}\right]_{ij}^{-1}$. This ratio matrix represents the SNR of $\gamma_d^{(k)}$. The SNR constraint [Eq. (6)] establishes that good visibility is obtained by incrementing the number of pulses by a factor exponential in the delay, Δd. Equivalently, the delay resolution enhancement Δd is a logarithmic function of the ensemble size ratio M'/M [see Eq. (G6)]. Consistently with the ideal case Eq. (4), the dependence on the reflectivity through the scaling factor, $R^{-2\Delta d}$, is reduced by increasing the reflectivity R, implying a more convenient implementation (requiring fewer measurements) for large reflectivity BS.

Some implications of Eqs. (4) and (6) are illustrated in Figs. 3(b) and 3(c) where the SNRs are plotted as functions of the delay, d, in decibels. We observe that they are straight lines with height equal to $\langle|\gamma_1^{(k)}\xi_M^{(k)-1}|\rangle_{\mathbf{s},\hat{H}}$ (at delay 1) and slope dependent only on the BS reflectivity, which is due to the R^{d-1} dependency in Eq. (4).

In Fig. 3(b), we see that the effect of incrementing the ensemble size is a uniform improvement of the SNR at each delay, corresponding to a shift equal to $5\log_{10}\left(M'/M\right)$ (see Appendix G 2). In Fig. 3(c) instead, the reflectivity R is changed while keeping M fixed. This R change alters both the SNR of $\gamma_1^{(k)}$ (height) and the decay of the SNR for further d (slope). A change from R to R' yields a difference of the SNR of $\gamma_1^{(k)}$ of

$$\left\langle\left|\frac{\gamma_1^{(k)}(R')}{\xi_M^{(k)}}\right|\right\rangle_{\mathbf{s},\hat{H}} - \left\langle\left|\frac{\gamma_1^{(k)}(R)}{\xi_M^{(k)}}\right|\right\rangle_{\mathbf{s},\hat{H}} = 20\log_{10}\left[\frac{1-R'}{1-R}\right]. \tag{7}$$

Hence for small delays, a large reflectivity corresponds to a reduced feedback light measured in the detector and can be detrimental, while for larger delays the situation is reversed. Increasing the value of the reflectivity R translates into an enhancement of memory retention improving the reservoir performance. Indeed more delayed terms remain above the noise threshold [zero line for the SNR in Fig. 3(c)].

C. Strategy to improve size scaling performance

A main advantage of quantum with respect to classical RC is the possibility to access a large Hilbert space [25,34,54]. Our aim is to test this potential when including quantum measurement, addressing how statistical noise affects the performance of the system when increasing the size of the reservoir, N. We have seen that in the ideal case, the total IPC grows quadratically with N [Fig. 2(b)] due to the delay depth enhancement [Fig. 2(a)]. In the finite ensemble case, however, the IPC displays a suboptimal growth with the reservoir size [blue bars in Fig. 4(a)]. Indeed the resolution of high delay inputs becomes increasingly demanding as the ideal delayed terms decrease exponentially with the delay d, Eq. (4). It follows that in order to maintain a constant normalized IPC when increasing the reservoir size N one needs to improve the measurement precision, by increasing the ensemble size M. Our aim is to quantify how to scale the resources (here the number of feedback loop pulses M), to maintain a good IPC for larger reservoirs, as in the ideal case. In principle, if the delay depth is a function $d(N) \sim \alpha N^2$ (in the ideal case) we would require (in the finite ensemble case) the ensemble size M to be an exponential function of the network size N of the form proportional to $R^{-2\alpha N^2}$ for the delay resolution to also improve quadratically with N. In the following, we show the performance of a less demanding use of resources, with the ensemble size M scaling with a polynomial, instead of an exponential, factor in the reservoir size N but still allowing to resolve longer delays.

From our previous discussion [Fig. 3(c)], it can be inferred that reservoirs with a smaller size, N, will exhibit

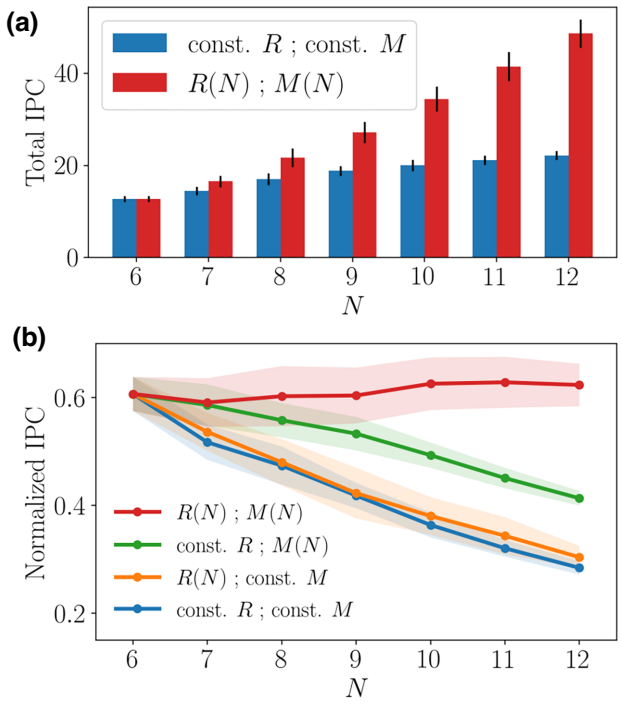

FIG. 4. (a) Total IPC as a function of the reservoir size N for the case of constant reflectivity $R = 0.72$ and ensemble size $M = 1.4 \times 10^5$ (blue) and the case where $R(N) = 1 - 10N^{-2}$ and $M(N) = 3N^6$. (b) Normalized IPC as a function of N for four different cases: when both R and M are constant (blue), when M is constant and R varies with N (orange), when R is constant and M varies with N (green) and when both R and M vary with N (red). The constant values of the reflectivity and the ensemble size are $R = 0.72$ and $M = 1.4 \times 10^5$, while the N-variable ones follow the functions $R(N) = 1 - 10N^{-2}$ and $M(N) = 3N^6$.

shorter memory and therefore achieve better performance for smaller values of the reflectivity R. In contrast, as N increases, making R larger would be beneficial, improving the resolution in accessing higher delayed input information. This suggests a strategy to improve the performance and the scaling of our reservoirs by tuning the reflectivity of the BS, which is typically an accessible parameter in experimental setups. We propose to take both the reflectivity R and the ensemble size M as functions of the reservoir size N, $R \equiv R(N)$ and $M \equiv M(N)$, in order to sustain a high normalized IPC when increasing N. We find that a good resolution of delayed terms can be maintained for larger reservoirs when considering

$$R(N) = 1 - \frac{\mathcal{C}}{N^2} \quad \text{and} \quad M(N) \propto N^8, \tag{8}$$

in which $\mathcal{C}$ is an arbitrary constant (see derivation in Appendix G 3). We notice that the convenience of the quadratic dependence in the reflectivity follows from the scaling of the delay depth with the system size. Therefore for non-Gaussian states, where the output layer could grow

faster than quadratically with N, a different scaling could be needed.

The condition in Eq. (8) ensures a quadratic scaling of the delay resolution with the network size, N. It thus may also guarantee a similar scaling with N of the total IPC. Actually, in the parameter range explored here, even numerical results limited to a less demanding number of pulses, choosing the scaling of the ensemble size $M(N) \propto N^6$ but for the reflectivity $R(N) = 1 - 10N^{-2}$, succeed in displaying a quadratic growth of the IPC [red bars in Fig. 4(a)]. In other words, we achieve the ideal performance scaling and a sharp improvement with respect to the suboptimal scaling for the reflectivity and the ensemble size kept constant [blue bars in Fig. 4(a)]. In Fig. 4(a), we compare the total IPC scaling with the reservoir size N for the case of constant ensemble size M and reflectivity R (blue bars) and the case in which both M and R scale with N (red bars). We set a 0.6 target normalized IPC obtained with $M(N = 6) = 1.4 \times 10^5$ measurements for the smallest considered reservoir ($N = 6$). Of course a higher target could be set with a larger ensemble $M(N = 6)$. A quadratic capacity scaling is achieved with the mentioned growth of reflectivity and ensemble size with N [red bars in Fig. 4(a)]. In order to assess the respective influences of increasing R and M, in Fig. 4(b), we compare the normalized capacities of the two previous cases (blue and red curves, respectively), with two more scenarios, when one of these parameters is kept constant. Both the increase of the reflectivity and the ensemble size are found to play a key role in achieving the best performance. These results illustrate that this photonic quantum platform exhibits the three main ingredients for QRC, namely memory, nonlinearity and high dimensionality (quantum advantage), in a realistic scenario.

D. Performance for a time-series prediction task

In this section, we analyze the performance of our reservoir in forecasting a chaotic time series. The concrete task is to predict the next step in the series, so the target function is $\bar{y}(s_k) = s_{k+1}$. We use, as common in the literature, the so-called Santa Fe dataset of experimental measurements of a chaotic laser [55,56]. The angle encoding chosen to perform the task is $\phi_k = 3\pi s_k/4$ as it provides high linear memory, as well as nonlinear memory [as shown in Fig. 2(b) and Appendix E]. For the output layer, we consider the covariance matrix and also fourth-order moments: $\left\{\left\langle X_i^2 X_j^2\right\rangle_M, \left\langle X_i^3 X_j\right\rangle_M\right\}_{j\geq i}$ for $i,j = 1,\ldots,N$ (the k superscripts, denoting the round trip, have been omitted for clarity). These observables have been added to avoid underfitting the data and, as a technical note, provide functions of crossed input terms of the form $s_{k-d}s_{k-d'}$, which are relevant for performing this prediction task. Different input encodings on Gaussian states than the ones used in this paper have been shown to provide crossed input functions in their second-order moments as well [52].

The dataset analyzed in this work contains a total of 4000 input points, which we divide into three consecutive sequences: the wash-out steps (of length L_m), the training steps (of length L) and the testing steps (of length L'). The length of the training step sequence is always fixed to be $L = 3000$, while the wash out and testing length would depend on the choice of the BS reflectivity R. The wash-out sequence length is set to guarantee the echo state property, and from Eq. (4) we require $R^{L_m} < 10^{-8}$ (so that $L_m \simeq 27$ for $R = 0.5$ and $L_m \simeq 175$ for $R = 0.9$). Given that the higher the reflectivity, the higher the memory retention of the reservoir, the number of wash-out steps to forget the initial conditions increases with R. The testing phase is done with the remaining data after the wash-out and the training phases.

In Fig. 5(a), we compare a sequence of signal values belonging to the testing phase with the reservoir predictions, both in the ideal case and in the finite ensemble case (with ensemble size $M = 10^6$). They provide accurate predictions, although the finite ensemble case shows a higher error when the oscillation amplitudes change abruptly. In Fig. 5(b) the *normalized mean square error* (NMSE), defined in Eq. (A5), is plotted as a function of the

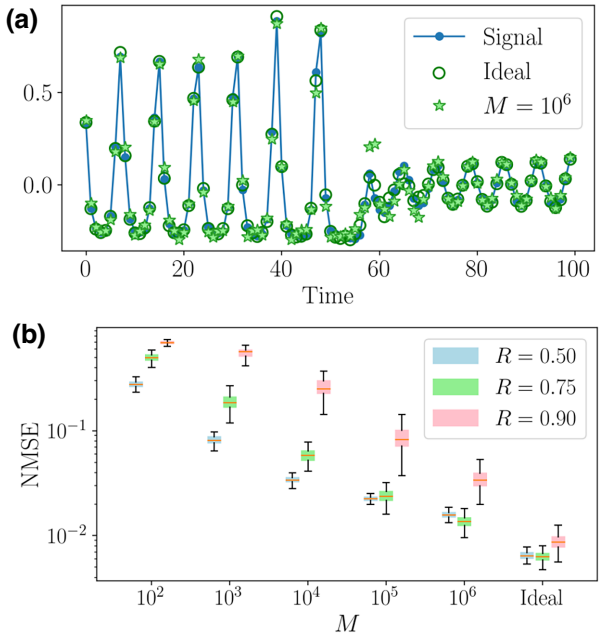

FIG. 5. Time-series prediction of chaotic signal: (a) prediction of chaotic signal (blue lines and dots) using a reservoir of $N = 12$ modes and reflectivity $R = 0.75$ both in the ideal case (green circles) and the finite ensemble case with size $M = 10^6$ (green stars). (b) Box plot of the normalized mean squared error (NMSE) as a function of the size of the ensemble M for different values of R. For a given value of M, the boxes for each reflectivity are split in the x axis to avoid overlapping. In all cases the number of modes inside each pulse was $N = 12$.

ensemble size M (including the ideal limit $M \to \infty$) for different values of the BS reflectivity R. The performance consistently improves when the ensemble size increases, reaching state-of-the-art performance (with NMSE $\sim 10^{-2}$ [57–59]) for high ensemble sizes.

In the ideal scenario, errors below the 10^{-2} threshold are reached with a relatively small reservoir size ($N = 12$). As shown in Fig. 5(b), for relatively small measurement ensembles, the best performance is achieved with $R = 0.5$, while for higher values of M the optimal reflectivity is $R = 0.75$. As we have seen in Sec. III B, for smaller values of R the SNR of the nearest past inputs (with a small delay) is greater [see Fig. 3(c)]. For very small ensembles, and thus a very high statistical noise, these setups with small reflectivity become more robust. When the ensemble size increases, a high SNR of further delayed inputs (higher memory) becomes more relevant, and thus the case of $R = 0.75$ improves its performance in comparison to the $R = 0.5$ case. Still, forecasting and memory performance are often varying differently [60]. It seems that, although the $R = 0.9$ case provides the most amount of memory, it is not so relevant for this forecasting task, and a higher SNR for small and intermediate delayed inputs [the case of $R = 0.75$ in Fig. 5(b)] has the best performance.

IV. DISCUSSION AND OUTLOOK

Optical platforms have been successful in classical RC [61] and for their features, such as fast processing rates and low decoherence, are also promising candidates for the first quantum implementations [25]. In ideal conditions, photonic platforms for QRC have been predicted to achieve optimal performance, displaying a quantum improvement, in particular the access to an enlarged Hilbert space and operation with low signals [30,62]. Still, to implement successfully and in real-time temporal tasks with a QRC advantage, open challenges need to be overcome [28,29,62,63], namely the limited experimental precision when estimating the output layer, as well as reducing the needed resources, also avoiding the use of external memories. Here, we have proposed an optical platform suitable for real-time QRC based on a physical ensemble of reservoirs, as independent pulses recirculating inside an optical fiber at each input injection.

The proposed setup displays the needed features for RC, such as fading memory and, in the limit of an infinite ensemble, it achieves optimal performance, with quadratic scaling of the IPC with the reservoir size (number of modes inside each pulse) for *vacuum* Gaussian states. This ideal limit is degraded by the effect of statistical noise and, as we increase the size of the reservoir N, the number of pulses in the fiber to sustain this quantum performance needs to be boosted. The strategy proposed here to overcome this limitation in large Hilbert space allows sustaining a quadratic scaling of the IPC, with efficient use of resources, thanks to an increase of the beam-splitter reflectivity when the reservoir size is scaled up. Indeed, being the dynamics of our platform linear, we can get analytical insights and find a bound factor for the number of pulses that is polynomial in the reservoir size. Beyond the performance in terms of memory, we have also checked that our platform can be used for nontrivial temporal tasks such as chaotic series prediction.

Optical devices required to implement our proposal in a laboratory are available with state-of-the-art photonic technology and the proposed architecture for on-line time-series processing provides both a tunable fading memory and the possibility to realize large measurement ensembles. With respect to recent experiments in NISQ circuits [29], our approach overcomes the need for external storage of input and repeating the experiment, just requiring the photonic QRC to be faster than the input rate to achieve real-time processing. For nontemporal classification tasks, a general dynamical framework applicable in circuit QED platforms has been reported in Ref. [63], while integrated photonic circuits [62] have been proposed using a quantum memristor. Our approach can be extended also to these platforms and going beyond Gaussian states.

Our work paves the way to real-time time-series processing with QRC, with state-of-the-art photonics technology and displaying a scalable quantum improvement leveraging limited resources.

ACKNOWLEDGMENTS

We acknowledge the Spanish State Research Agency, through the Severo Ochoa and María de Maeztu Program for Centers and Units of Excellence in R&D (MDM-2017-0711 and CEX2021-001164-M funded by the MCIN/AEI/10.13039/501100011033) and through the QUARESC project (PID2019-109094GB-C21 and -C22/ AEI / 10.13039/501100011033). We also acknowledge funding by CAIB through the QUAREC project (PRD2018/47). G.L.G. is funded by the Spanish MEFP/MiU and co-funded by the University of the Balearic Islands through the Beatriz Galindo program (BG20/00085). The CSIC Interdisciplinary Thematic Platform (PTI) on Quantum Technologies in Spain is also acknowledged. J.G-B. is funded by the Conselleria d'Educació, Universitat i Recerca of the Government of the Balearic Islands with grant code FPI/036/2020.

APPENDIX A: RESERVOIR COMPUTING

Reservoir computing (RC) is a ML framework that takes advantage of complex dynamical systems to perform learning tasks with a low-cost training protocol. It is especially suited for temporal tasks regarding time series. Every RC scheme can be separated into three steps and layers: the input layer, the reservoir dynamics and the readout layer. For the input layer, a vector $\mathbf{s}_k$ from the time-series

signal, $\{\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_L\}$, is encoded and fed into the dynamical system, also called *reservoir*, degrees of freedom. While in general each $\mathbf{s}_k$ can be a vector, throughout our work we take it as a scalar (the same applies to the output elements). The length of the list, L, denotes the number of training samples, or training steps, to perform the learning protocol. After each introduction of an input, the reservoir degrees of freedom, $\mathbf{v}_{k-1}$, evolve in time through a nonlinear mapping for a fixed time Δt (reservoir dynamics). The reservoir observables and degrees of freedom, $\mathbf{v}_k$, are measured and taken as outputs. A linear mapping is then applied to these observables to yield an output vector $\mathbf{y}_k$ (readout layer), which is built as a function of the weight parameters that are updated in the training stage. After the last step, the protocol continues for the following input, $\mathbf{s}_{k+1}$.

We can write the dynamical mapping of the reservoir degrees of freedom as

$$\mathbf{v}_k = \mathcal{H}\left(\mathbf{v}_{k-1}, s_k\right) \quad ; \quad k \in \mathbb{Z}, \tag{A1}$$

in which we need $\mathcal{H}$ to be a nonlinear mapping of the input s_k. This mapping will remain fixed throughout the whole protocol. Following standard RC practices, we take the output layer function to be a linear function of the reservoir observables

$$\begin{aligned} y_k &= w_0 + \mathbf{w}^\top \mathbf{v}_k \\ &= \left(1, \mathbf{v}_k^\top\right) \begin{pmatrix} w_0 \\ \mathbf{w}, \end{pmatrix} \end{aligned} \tag{A2}$$

with the $(D_R + 1)$-dimensional vector $\mathbf{W} = \left(w_0, \mathbf{w}^\top\right)^\top$ containing the training weights (we add a constant bias term w_0). For a given set of L inputs that we feed to the reservoir, we can define the following matrices:

$$V = \begin{pmatrix} 1 & \mathbf{v}_1^\top \\ 1 & \mathbf{v}_2^\top \\ \vdots & \vdots \\ 1 & \mathbf{v}_L^\top \end{pmatrix} \quad ; \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_L \end{pmatrix}, \tag{A3}$$

so that the following equation holds:

$$\mathbf{y} = V\mathbf{W}. \tag{A4}$$

After choosing a given target function that we want our reservoir to reproduce, $\bar{\mathbf{y}}$, we want to find the weights that minimize the normalized mean-square error (NMSE) of the predicted $\mathbf{y}$ and the target

$$\mathrm{NMSE}_L(\mathbf{y}, \bar{\mathbf{y}}) = \frac{\sum_{k=1}^{L} (y_k - \bar{y}_k)^2}{\sum_{k=1}^{L} \bar{y}_k^2}. \tag{A5}$$

The optimal set of weights for this condition to hold are the ones obtained through the following procedure [64]:

$$\mathbf{W}_{\mathrm{opt}} = V^{\mathrm{MP}} \bar{\mathbf{y}}, \tag{A6}$$

where $V^{\mathrm{MP}} = \left(V^\top V\right)^{-1} V^\top$ is the Moore-Penrose inverse of V. The higher the value of L the more precise our estimation of the optimal weights will usually be.

APPENDIX B: LINEAR CAPACITY AND INFORMATION-PROCESSING CAPACITY

Throughout this paper, we test the performance of our (already trained) reservoirs using the *capacity* to reproduce a function of the inputs. We consider a reservoir that has already finished the learning protocol to reproduce a given target function $\bar{y}$. We thus take a given string of L' inputs $\mathbf{s} = (s_1, s_2, \ldots, s_{L'})$ and a vector $\bar{\mathbf{y}} = (\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_{L'})$ with components $\bar{y}_k = \bar{y}(\mathbf{s}, k)$. The length of the additional sequence of inputs and target function vectors, L', denotes the number of testing steps right after the learning protocol. Then, the capacity of our reservoir to reproduce the given target function $\bar{y}$ is

$$C_{\bar{\mathbf{y}}} = \max\left[0, 1 - \min_{\mathbf{W}} \mathrm{NMSE}_{L'}(\mathbf{y}, \bar{\mathbf{y}})\right], \tag{B1}$$

where $\mathbf{y}$ is the output vector of the reservoir, as defined in Eq. (A4), after the training of the weights has been performed. To check the memory of the system we use the *linear capacity*, in which we set the target functions to be $\bar{y}_d(s_k) = s_{k-d}$, where the parameter d denotes the delay of the input. As a quantitative measure of how well our reservoir performs in general, we use the *information processing capacity* (IPC) or total capacity [49]. The main idea is to compute the capacity of all the different orthogonal functions that our reservoir can approximate. By choosing a complete, orthogonal family of functions, $\{\bar{y}\}$, the IPC can be estimated as

$$\mathrm{IPC} = \sum_{\bar{y}} C_{\bar{\mathbf{y}}}. \tag{B2}$$

In our simulations, we have chosen the family of functions to be

$$\bar{y}_D^{(d)}(s_k) = \mathcal{P}_D(s_{k-d}), \tag{B3}$$

where the function $\mathcal{P}_D$ is the normalized Legendre polynomial of degree D. As we cannot sum over an infinite number of delays and degrees, we have chosen $d \in [0, d_{\max}]$

and $D \in [1, D_{\max}]$. For every simulation in this article, it has sufficed to take $d_{\max} = 75$ and $D_{\max} = 5$.

Fading memory is a necessary property for any dynamical system to work as a reservoir computer [65]. In our platform, short-term memory is ensured by the BS coupling (proof on Appendix F). In fact, as it was shown in Eq. (4), the magnitude of the delayed terms $\gamma_d^{(k)}$ dropped as R^d (d denoting the delay and R denoting the BS reflectivity). For a large enough delay, the reservoir would not be able to resolve that delayed input information. As we start our reservoir in a completely random state, we first introduce a list of L_m wash-out inputs, before starting the learning protocol, just to ensure the initial conditions have been forgotten by the reservoir. We make L_m large enough to have the prefactor going below the numerical precision. This is equivalent to fulfilling the following condition:

$$R^{L_m} < 10^{-16}, \tag{B4}$$

so the number of wash-out steps depends on R. For our estimations of the linear capacity and the IPC on Sec. III, the training steps have been chosen to be $L = 10^4$ and the testing steps $L' = 5000$.

APPENDIX C: GAUSSIAN STATES

The evolution of an N-mode Gaussian quantum state $\hat{\rho}$ is completely determined by the dynamics of its displacement vector $\mathbf{r}$ of dimension $2N$ and its positive symmetric covariance matrix σ of dimension $2N \times 2N$ [66,67]. Each element of $\mathbf{r}$ (σ) denotes the mean value (covariance) of each quadrature for each mode. As we inject vacuum states, the displacements of the external signal are null vectors. The covariance matrix of the ancilla state is the composition of N independent and identical squeezed states. We can thus write its covariance matrix at round trip k as $\sigma_{\text{anc}}^{(k)} = \bigoplus_{i=1}^{N} \sigma_{\text{sq}}(s_k)$, in which $\sigma_{\text{sq}}(s_k)$ is the covariance matrix of each single external mode. It is a nonlinear function of the kth input, s_k, which can be expanded as

$$\sigma_{\text{sq}}(s_k) = \begin{pmatrix} c_+(s_k) & z(s_k) \\ z(s_k) & c_-(s_k) \end{pmatrix}, \tag{C1}$$

for the following functions:

$$c_\pm(s_k) = \cosh(2r_k) \pm \cos(\phi_k) \sinh(2r_k) \tag{C2}$$

$$z(s_k) = \sin(\phi_k) \sinh(2r_k). \tag{C3}$$

The two main parameters to tune are the squeezing strength, r_k, and the squeezing angle, ϕ_k. As commented at the beginning of Sec. III, we use the following encoding: $r_k = 1$ and $\phi_k = 3\pi s_k/4$.

1. Gaussian measurements on multipartite systems

In this section, we provide an analytical treatment to describe the effect of quantum measurements on the quadrature operators. We start by defining a composite Gaussian state, which we can separate into subsystems A and B. We can define the displacement vector and covariance matrix of the total Gaussian state as

$$\mathbf{r} = \begin{pmatrix} \mathbf{r}_A \\ \mathbf{r}_B \end{pmatrix}$$

$$\Gamma = \begin{pmatrix} \sigma_A & \sigma_{AB} \\ \sigma_{AB} & \sigma_B \end{pmatrix}, \tag{C4}$$

where $\mathbf{r}_A$ ($\mathbf{r}_B$) and σ_A (σ_B) are the displacement vector and covariance matrix of subsystem A (B), while σ_{AB} denotes the correlations between subsystem A and B. We now consider measurements that can be performed on subsystem B so that the conditional state of subsystem A remains Gaussian. For that we consider the family of general-dyne measurements [67,68], in which homodyne detection is included. Thus, measuring the subsystem B with a random outcome, $\mathbf{r}_{\text{out}}^B$, gives rise to the following conditional state of A:

$$\mathbf{r}_A' = \mathbf{r}_A + \sigma_{AB} (\sigma_B + \sigma_m)^{-1} \left(\mathbf{r}_{\text{out}}^B - \mathbf{r}_B\right) \tag{C5}$$

$$\sigma_A' = \sigma_A - \sigma_{AB} (\sigma_B + \sigma_m)^{-1} \sigma_{AB}^\top, \tag{C6}$$

where σ_m is a positive symmetric matrix that depends on the kind of general-dyne measurement that is being performed. The measurement outcome is drawn from a multivariate Gaussian distribution with covariance matrix equal to $\sigma_B + \sigma_m$ and mean vector equal to $\mathbf{r}_B$. Thus, we can write

$$\mathbf{r}_{\text{out}}^B = \mathbf{r}_B + \sqrt{\sigma_B + \sigma_m}\,\mathbf{u}(\mathbf{0}, I), \tag{C7}$$

where $\mathbf{u}(\mathbf{0}, I)$ is a random vector drawn from a Gaussian distribution with 0 mean and covariance matrix equal to the identity, I. With this, we can rewrite Eq. (C5) as

$$\mathbf{r}_A' = \mathbf{r}_A + \sigma_{AB} (\sigma_B + \sigma_m)^{-1/2} \mathbf{u}(\mathbf{0}, I), \tag{C8}$$

which describes how measuring subsystem B has affected the state in A.

For the specific case of homodyne detection of the x quadratures of each mode, we have [66,67,69]

$$\sigma_m = \lim_{z \to \infty} \bigoplus_{i=1}^{N} \text{diag}\left(z^{-2}, z^2\right). \tag{C9}$$

In this case, the term of Eqs. (C5) and (C6) $(\sigma_B + \sigma_m)^{-1}$ in the limit of z tending to infinity, tends towards the

following expression:

$$\lim_{z\to\infty}\left[\sigma_B+\bigoplus_{i=1}^{N}\mathrm{diag}\left(z^{-2},z^2\right)\right]^{-1}=(\Pi\sigma_B\Pi)^{\mathrm{MP}}, \quad \text{(C10)}$$

where $\Pi=\bigoplus_{i=1}^{N}\begin{pmatrix}1&0\\0&0\end{pmatrix}$ and MP stands for the Moore-Penrose inverse, [69]. The outcome vector from the homodyne measurement can be modeled from Eq. (C7), using the σ_m matrix from Eq. (C9). However, special care has to be taken in this case, as in the limit of $z\to\infty$ we are introducing a divergent variance in the p-quadrature degrees of freedom. The shape that the inverse matrix from Eq. (C10) takes in this limit ensures convergence of Eqs. (C5) and (C6), as the p-quadrature degrees of freedom of the measurement outcome do not play a role in determining the conditional state of A. In practice, to obtain the outcome observables of the homodyne detection, we trace out the diverging degrees of freedom and, thus, consider only the x quadratures. The resulting probability distribution of the outcome becomes

$$p(\mathbf{x}^B_{\mathrm{out}})=\frac{\exp\left\{-\left(\mathbf{x}^B_{\mathrm{out}}-\mathbf{x}_B\right)^{\top}\sigma^{-1}_{B,\mathbf{x}}\left(\mathbf{x}^B_{\mathrm{out}}-\mathbf{x}_B\right)\right\}}{\pi^N\sqrt{\mathrm{Det}(\sigma_{B,\mathbf{x}})}}, \quad \text{(C11)}$$

where $\sigma_{B,\mathbf{x}}=\mathrm{Tr}_{\mathbf{p}}(\sigma_B)$ and $\mathbf{x}_B=\mathrm{Tr}_{\mathbf{p}}(\mathbf{r}_B)$ (the degrees of freedom of the p quadratures have been traced out). The notation $\mathrm{Tr}_{\mathbf{p}}[\bullet]$ does not actually stand for the usual partial trace of a matrix, it is just a way of writing we are tracing out the components of the covariance matrix and first-moment vector, which contain information of the p quadratures of every mode. For a generic $2N$-dimensional covariance matrix, σ, tracing out these components would yield a N-dimensional matrix with components:

$$\left[\mathrm{Tr}_{\mathbf{p}}(\sigma)\right]_{ij}=\langle\hat{x}_i\hat{x}_j\rangle-\langle\hat{x}_i\rangle\langle\hat{x}_j\rangle, \quad \text{(C12)}$$

where the mean values stand for the quantum expected values of the observables for a given quantum state. This is consistent with the fact that, in experiments, homodyne detection only yields an outcome of one quadrature for each mode.

In our platform, we can easily identify the two subsystems. Before the BS coupling, we have the pulse coming out from the fiber and the external ancilla pulse, which are both independent of one another. After the BS coupling, we have the pulse that belongs to the fiber path as one subsystem and the one that travels to the homodyne detector (HD) as another. In this last case, there is generally entanglement between them (appearing as correlation terms in the σ_{AB} matrix). As the x quadratures of each pulse that reaches the detector is measured, the relative pulse (going through the fiber path) is conditioned to the measurement outcomes as detailed in Eqs. (C5) and (C6). The observables averaged over this ensemble of conditional states resemble the unconditional evolution, as it is expected in the case of Gaussian states [68]. Backaction effects in other quantum substrates would generally affect averages performed over conditional ensembles, so the unconditional evolution is not completely obtained through averaging. In the context of QRC this can negatively affect the performance, as it was shown in the case of qubits [28].

2. Round-trip dynamics of the reservoir

In this section, we describe the dynamical evolution of a single pulse at each round trip. As a physical ensemble of pulses is found in the fiber, the resulting equations will hold true for every pulse. At round trip k, every pulse coming out from the fiber will eventually couple to an external ancilla state through the BS. The quantum state of the whole system prior to the BS coupling can be described with the following displacement vector and covariance matrix:

$$\mathbf{r}_0^{(k)}=\begin{pmatrix}\mathbf{r}_R^{(k)}\\\mathbf{0}\end{pmatrix} \quad \text{(C13)}$$

$$\Gamma_0^{(k)}=\begin{pmatrix}\sigma_R^{(k)}&0\\0&\sigma_{\mathrm{anc}}^{(k)}\end{pmatrix}, \quad \text{(C14)}$$

where we write it as a composite system in which $\mathbf{r}_R$ and σ_R are the displacements and covariance matrix of the pulse coming out from the fiber and σ_{anc} is the covariance matrix of the ancilla pulse; the subindex 0 is introduced to denote the initial state in which the ancilla pulse and the reservoir pulse have not yet arrived to the BS. We remark that the ancilla pulse has a null displacement vector, as it is a vacuum state. Also, as both pulses are initially independent, the off-diagonal matrices of Γ_0 are null. As there are N modes inside each pulse, with two quadratures each, the displacement vectors of each subsystem are $2N$ dimensional and their covariance matrices are $2N\times 2N$ matrices. In total, $\mathbf{r}_0$ is a $4N$-dimensional vector and Γ_0 is a $4N\times 4N$ matrix. The action of the BS can be written in the following matrix form:

$$B_R=\begin{pmatrix}\sqrt{R}\,I_{2N}&\sqrt{T}\,I_{2N}\\-\sqrt{T}\,I_{2N}&\sqrt{R}\,I_{2N}\end{pmatrix}, \quad \text{(C15)}$$

where I_{2N} is the $2N\times 2N$ identity matrix. We recall that R (T) is the reflectivity (transmissivity) of the BS. In turn, the symplectic matrix that describes the evolution inside both nonlinear crystals is the following:

$$S(\Delta t)=\begin{pmatrix}S_1(\Delta t)&0\\0&S_2(\Delta t)\end{pmatrix}, \quad \text{(C16)}$$

in which S_1 (S_2) describes the evolution inside the crystal placed at the feedback fiber (detector) path. These matrices describe the evolution of the quadrature operators under the action of quadratic Hamiltonians such as the ones we are considering in Eq. (1). Concretely, the matrices $S_1(\Delta t)$ and $S_2(\Delta t)$ are analog to the generic unitary operator $\hat{U}(\Delta t) = \exp\left[-i\hat{H}_{\chi^{(2)}}\Delta t\right]$ for each nonlinear crystal. They act on the quadrature operators in phase space [66,67]. In our discussion of the parameters at the beginning of Sec. III, it was mentioned that the materials were not allowed to produce squeezing levels far beyond 15 dB. The squeezing produced by the materials can be computed by the Bloch-Messiah decomposition of matrices S_1 and S_2 [70,71]. We write the symplectic matrix of the whole process (BS + nonlinear crystals) as follows:

$$S'(\Delta t) = S(\Delta t)B_R. \tag{C17}$$

After both pulses have come out from the nonlinear media, their state parameters have evolved as follows:

$$\mathbf{r}' = S'(\Delta t)\mathbf{r}_0, \tag{C18}$$

$$\Gamma' = S'(\Delta t)\Gamma_0 S'(\Delta t)^\top, \tag{C19}$$

where the k labeling has been omitted for clarity. We can now split both $\mathbf{r}'$ and Γ' into two subsystems: one for the pulse that is being reinjected into the fiber and another one for the pulse that travels to the detector. We thus label with the "fiber" ("HD") subscript to the parameters of the pulse that is reinjected in the fiber (traveling to the detector). So the resulting displacement vector and covariance matrix from Eqs. (C18) and (C19) can be written as

$$\mathbf{r}' = \begin{pmatrix} \mathbf{r}_{\text{fiber}} \\ \mathbf{r}_{\text{HD}} \end{pmatrix}, \tag{C20}$$

$$\Gamma' = \begin{pmatrix} \sigma_{\text{fiber}} & \sigma_{\text{corr}} \\ \sigma_{\text{corr}}^\top & \sigma_{\text{HD}} \end{pmatrix}, \tag{C21}$$

in which σ_{corr} is a $2N \times 2N$ matrix containing the correlations between the fiber pulse and the detected pulse. It is equivalent to σ_{AB} in Eq. (C4). As we saw in Sec. C 1, and equivalent to the result in Eq. (C7) the measurement outcome of the measured pulse quadratures is the following:

$$\mathbf{r}_{\text{out}}^{(k)} = \mathbf{r}_{\text{HD}}^{(k)} + \sqrt{\sigma_{\text{HD}}^{(k)} + \sigma_m}\,\mathbf{u}^{(k)}, \tag{C22}$$

where, again, $\mathbf{u}^{(k)}$ is a random $2N$-dimensional vector whose components are drawn from a normal distribution with zero mean a variance equal to one. The feedback pulse coming out from the fiber in the following round trip will have the following parameters:

$$\mathbf{r}_R^{(k+1)} = \mathbf{r}_{\text{fiber}}^{(k)} + \sigma_{\text{corr}}^{(k)}\left[\sigma_{\text{HD}}^{(k)} + \sigma_m\right]^{-1/2}\mathbf{u}^{(k)}, \tag{C23}$$

$$\sigma_R^{(k+1)} = \sigma_{\text{fiber}}^{(k)} - \sigma_{\text{corr}}^{(k)}\left[\sigma_{\text{HD}}^{(k)} + \sigma_m\right]^{-1}\left[\sigma_{\text{corr}}^{(k)}\right]^\top, \tag{C24}$$

in which the state of the pulse is conditioned to the outcome measurement from the homodyne detector [see Eqs. (C5) and (C6)].

3. Recursive equations

In this subsection, we work from Eqs. (C23) and (C24) to obtain expressions for the relevant output parameters as functions of the input history. These expressions are going to be useful to obtain the ideal case observables from Eq. (3). We start with the output displacements. From Eq. (C18) we have the following simple relations:

$$\mathbf{r}_{\text{fiber}}^{(k)} = \sqrt{R}S_1(\Delta t)\mathbf{r}_R^{(k)} \tag{C25}$$

$$\mathbf{r}_{\text{HD}}^{(k)} = -\sqrt{T}S_2(\Delta t)\mathbf{r}_R^{(k)}. \tag{C26}$$

We can substitute Eq. (C23) into Eqs. (C25) and (C26). Thus, by recursion, it yields the result

$$\mathbf{r}_{\text{fiber}}^{(k)} = \sum_{d=1}^{k} R^{d/2}S_1^d(\Delta t)\sigma_{\text{corr}}^{(k-d)} \times \left[\sigma_{\text{HD}}^{(k-d)} + \sigma_m\right]^{-1/2}\mathbf{u}^{(k-d)} \tag{C27}$$

$$\mathbf{r}_{\text{HD}}^{(k)} = -TS_2(\Delta t)\sum_{d=1}^{k} R^{(d-1)/2}S_1^d(\Delta t)\sigma_{\text{corr}}^{(k-d)}, \times \left[\sigma_{\text{HD}}^{(k-d)} + \sigma_m\right]^{-1/2}\mathbf{u}^{(k-d)}. \tag{C28}$$

Deviations from the origin (vacuum) in the reservoir pulse are originated by the stochastic displacements that the recurrent backaction produces. The same procedure can be performed with the covariance matrices. From Eq. C19, these relations follow

$$\sigma_{\text{fiber}}^{(k)} = S_1(\Delta t)\left[R\sigma_R^{(k)} + T\sigma_{\text{anc}}^{(k)}\right]S_1(\Delta t)^\top, \tag{C29}$$

$$\sigma_{\text{HD}}^{(k)} = S_2(\Delta t)\left[T\sigma_R^{(k)} + R\sigma_{\text{anc}}^{(k)}\right]S_2(\Delta t)^\top, \tag{C30}$$

$$\sigma_{\text{corr}}^{(k)} = \sqrt{RT}S_1(\Delta t)\left[\sigma_R^{(k)} - \sigma_{\text{anc}}^{(k)}\right]S_2(\Delta t)^\top. \tag{C31}$$

From the equation for σ_{fiber}, we note that similar recursion equations can be obtained by substituting it in Eq. (C19). By recursion we can, again, get the expression of the covariance matrix in Eq. (C30) as functions of the input history. We write only the expression for σ_{HD} below, as

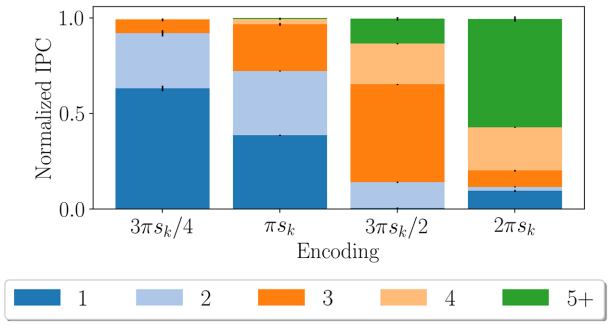

FIG. 6. Encoding choice effect on the IPC: bar plot of the normalized IPC as a function of different angle encodings $\phi_k = \beta s_k$. Each bar is split into different degree contributions. In all cases the reservoir size is $N = 10$ and the reflectivity is $R = 0.5$.

it is the only one that is necessary for the mathematical derivations in the next section:

$$\sigma_{\mathrm{HD}}^{(k)} = RS_2\sigma_{\mathrm{anc}}^{(k)}S_2^{\top} + TS_2\left[\sum_{d=1}^{k} R^{d-1}S_1^{d-1}\left[TS_1\sigma_{\mathrm{anc}}^{(k-d)}S_1^{\top} - \sigma_{\mathrm{corr}}^{(k-d)}\left[\sigma_{\mathrm{HD}}^{(k-d)} + \sigma_m\right]^{-1}\left(\sigma_{\mathrm{corr}}^{(k-d)}\right)^{\top}\right]\left(S_1^{d-1}\right)^{\top}\right]S_2^{\top}. \tag{C32}$$

The Δt dependency of S_1 and S_2 has been omitted for clarity.

APPENDIX D: INPUT ENCODING AND NONLINEARITY

In this brief section we are going to elaborate on the degree distribution of the IPC as a function of the encoding choice, which is known to influence the nonlinearity [30,72,73]. We recall that the input ancilla states are single-mode squeezed vacuum states with the input signal encoded in their covariance matrices as in Eqs. (C1)–(C3). Concretely, in the main text we have used a squeezing angle encoding setting $r_k = 1$ and $s_k = 3\pi s_k/4$. The input could also be encoded in the squeezing strength r_k, yielding a different degree distribution [30,72]. In this Appendix we focus our attention on angle encodings with $r_k = 1$ and $\phi_k = \beta s_k$, where $\beta > 0$. In Eqs. (C2) and (C3) the input nonlinearity comes from the functions $\sin(\phi_k)$ and $\cos(\phi_k)$. We now write the Taylor expansion of these functions as

$$\sin(\beta s_k) = \beta s_k + \frac{\beta^3 s_k^3}{6} + \mathcal{O}(\beta^5), \tag{D1}$$

$$\cos(\beta s_k) = 1 - \frac{\beta^2 s_k^2}{2} + \mathcal{O}(\beta^4). \tag{D2}$$

And so we actually see that the smaller the value of β, the closer $c_{\pm}(s_k)$ ($z(s_k)$) are from being linear functions of s_k and s_k^3 (s_k^2). So for smaller values of β the linear and low nonlinear contributions to the IPC will be greater. On the other hand, as we increase the value of β, the high nonlinear terms will be more relevant. This is visualized in Fig. 6.

APPENDIX E: ENSEMBLE AVERAGES AND IDEAL CASE LIMIT

In this section, we derive the expression for the covariance matrix of the output signal in Eq. (C22) for an infinite ensemble of pulses, so that we get rid of any statistical error. It is relevant to remark that, as we are not considering any filter, every pulse in the ensemble is conditioned to a measurement outcome history throughout a certain number of round trips. As we see, by averaging over an ensemble of conditioned realizations with no filters, the unconditional state arises. As it was already commented in Sec. II, this is an already known feature of conditional Gaussian states under general-dyne monitoring, [68].

To denote the ensemble pulses we add a label m, which ranges from 1 to M (the total number of pulses in the ensemble). In that manner, we take Eq. (C22) and add the ensemble label, so the measurement outcome of the mth pulse is

$$\mathbf{r}_{\mathrm{out}}^{(k,m)} = \mathbf{r}_{\mathrm{HD}}^{(k,m)} + \sqrt{\sigma_{\mathrm{HD}}^{(k)} + \sigma_m}\,\mathbf{u}^{(k,m)}. \tag{E1}$$

In Fig. 1(b) the outcome measurement of the first term of $\mathbf{r}_{\mathrm{out}}^{(k,m)}$ for some number of pulses in the ensemble was shown. As we discussed at the end of Sec. C 1, for homodyne detection of the x quadratures the readout vector has N terms (instead of the $2N$ terms that $\mathbf{r}_{\mathrm{out}}^{(k,m)}$ in a more general scenario). For this section, we arrive at the final expressions in the general case and, then, take the homodyne limit. In Sec. II, we introduced the ensemble averages as

$$\left\langle A^{(k)}\right\rangle_M = \frac{1}{M}\sum_{m=1}^{M} A^{(k,m)} \tag{E2}$$

for any generic measured observable, A. We use the following notation to denote the average limit of an infinite number of ensemble realization: $\left\langle A^{(k)}\right\rangle \equiv \lim_{M\to\infty}\left\langle A^{(k)}\right\rangle_M$. Having this defined, we consider the expected value of the covariance of the output signal from Eq. (E1) and obtain

$$\left\langle \mathbf{r}_{\mathrm{out}}^{(k)}\mathbf{r}_{\mathrm{out}}^{(k)\top}\right\rangle = \left\langle \mathbf{r}_{\mathrm{HD}}^{(k)}\mathbf{r}_{\mathrm{HD}}^{(k)\top}\right\rangle + \sigma_{\mathrm{HD}}^{(k)} + \sigma_m, \tag{E3}$$

in which we have taken advantage of the fact that the terms of every vector $\mathbf{u}^{(k,m)}$ are taken from a one-dimensional normal distribution with zero mean and variance equal to one. If we substitute $\mathbf{r}_{\mathrm{HD}}^{(k,m)}$ with its expression from Eq. (C28), perform the average over infinite pulses, and

also substitute $\sigma_{\mathrm{HD}}^{(k)}$ from its expression in Eq. (C32), we can derive the resulting expression

$$\left\langle \mathbf{r}_{\mathrm{out}}^{(k)} \mathbf{r}_{\mathrm{out}}^{(k)\top} \right\rangle = RS_2 \sigma_{\mathrm{anc}}^{(k)} S_2^{\top} + \sigma_m + T^2 S_2 \left[\sum_{d=1}^{k-1} R^{d-1} S_1^d \sigma_{\mathrm{anc}}^{(k-d)} \left(S_1^d\right)^{\top} \right] S_2^{\top}. \quad \text{(E4)}$$

We already note that these second moments resemble the unconditional dynamics, as there are no terms depending on $\sigma_{\mathrm{corr}}^{(k-d)} \left[\sigma_{\mathrm{HD}}^{(k-d)} + \sigma_m\right]^{-1} \sigma_{\mathrm{corr}}^{(k-d)\top}$, which determined the conditional evolution of the feedback pulses, as seen in Eqs. (C23) and (C24). Although the pulses that reach the detector are not, in general, vacuum states (due to the conditional displacements), we can easily see that $\left\langle \mathbf{r}_{\mathrm{out}}^{(k)} \right\rangle = \mathbf{0}$ (back to the unconditional scenario). In that case, the second moments' expression that we derived in Eq. (E4) is equal to the covariance matrix of the output signal, that is, $\left\langle \mathbf{r}_{\mathrm{out}}^{(k)} \mathbf{r}_{\mathrm{out}}^{(k)\top} \right\rangle = \mathrm{cov}\left(\mathbf{r}_{\mathrm{out}}^{(k)}\right)$. We rename $\mathrm{cov}\left(\mathbf{r}_{\mathrm{out}}^{(k)}\right)$ to $\sigma_{\mathrm{ideal}}^{(k)}$, yielding

$$\sigma_{\mathrm{ideal}}^{(k)} = \sum_{d=0}^{\infty} \gamma_d^{(k)} + \sigma_m, \quad \text{(E5)}$$

in which each $\gamma_d^{(k)}$ is the term including $\sigma_{\mathrm{anc}}^{(k-d)}$ from Eq. (E4). Expression (E5) is similar to Eq. (3). The only difference lies in the fact that in Eq. (3) the covariance corresponds to the position quadratures block as we have considered homodyne detection of the x quadratures to obtain $\sigma_{\mathrm{ideal}}^{(k)}$. This is equivalent to tracing out the p-quadrature degrees of freedom from every term in Eq. (E5), taking into account that $\mathrm{Tr}_{\mathbf{p}}(\sigma_m) = 0$ (for this measurement scheme). This yields the expression

$$\begin{aligned} \gamma_0^{(k)} &= R\mathrm{Tr}_{\mathbf{p}}\left\{S_2 \sigma_{\mathrm{anc}}^{(k)} S_2^{\top}\right\} \\ \gamma_d^{(k)} &= (1-R)^2 R^{d-1} \mathrm{Tr}_{\mathbf{p}}\left\{S_2 S_1^d \sigma_{\mathrm{anc}}^{(k-d)} \left(S_1^d\right)^{\top} S_2^{\top}\right\} \\ &\quad \text{for } d \geq 1, \end{aligned} \quad \text{(E6)}$$

in the case of homodyne detection of the x quadratures. We have replaced the transmissivity T by $1-R$ to make the relation between Eqs. (E6) and (4) more evident. It can be noted from Eq. (E6) that the dependence on d is not only found on R^{d-1}, but also inside the trace (in the terms S_1^d). However, we observe numerically that averaging among different Hamiltonians and different input strings removes every dependency on d apart from the one in R^{d-1}.

APPENDIX F: FADING MEMORY CONDITION AND ECHO STATE PROPERTY

In this section, we briefly state how both the fading memory condition and the echo-state property [15, 65], which are necessary conditions for a functional RC platform, are fulfilled by our platform. We begin from the theorem formulated in Ref. [30] for linear Gaussian dynamics. We have a linear Gaussian system whose quadrature operators evolve at the kth time step such as

$$\hat{\mathbf{r}}_R^{(k+1)} = A\hat{\mathbf{r}}_R^{(k)} + B\hat{\mathbf{r}}_{\mathrm{anc}}^{(k)}, \quad \text{(F1)}$$

where $\hat{\mathbf{r}}_R^{(k)}$ is the quadrature operator vector of the reservoir and $\hat{\mathbf{r}}_{\mathrm{anc}}^{(k)}$ the one of the ancilla input, both at time step k. Then both the echo-state property and the fading memory condition are fulfilled if $\rho[A] < 1$, being $\rho[\bullet]$ the *spectral radius* of a matrix. In our case, the symplectic matrix determining the dynamics of our platform is the one from Eq. (C17), which can be explicitly written in the form

$$S'(\Delta t) = \begin{pmatrix} \sqrt{R} S_1(\Delta t) & -\sqrt{1-R} S_1(\Delta t) \\ \sqrt{1-R} S_2(\Delta t) & \sqrt{R} S_2(\Delta t) \end{pmatrix}. \quad \text{(F2)}$$

In our platform, it can be clearly seen how $A \equiv \sqrt{R} S_1(\Delta t)$, so we must have $\rho\left[\sqrt{R} S_1(\Delta t)\right] = \sqrt{R}\rho[S_1(\Delta t)] < 1$. It can be shown that $\rho[S_1(\Delta t)] = 1$. The proof is carried out as follows: from the time evolution of the quadrature operators in the Heisenberg picture under a quadratic Hamiltonian [as on Eq. (1)] it can be shown that the resulting symplectic transformation can be written as $S_1(\Delta t) = \exp(\Omega H_1 \Delta t)$, where H_1 is a $2N \times 2N$ symmetric matrix and $\Omega = \bigoplus_{i=1}^{N} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. We have considered only stable Hamiltonians in our simulations, so $H_1 > 0$ (positive definite). In this scenario, it can be shown that the eigenvalues of ΩH_1 are purely imaginary [67], which means that $\rho[\exp(\Omega H_1 \Delta t)] = 1$. We thus have that in our platform $\rho(A) = \sqrt{R}$, which is always < 1, and thus both the fading memory condition and the echo-state property are fulfilled.

APPENDIX G: RESOLUTION ANALYSIS

In this section, we analyze in detail the main mathematical relations concerning the SNR of the delayed terms, $\gamma_d^{(k)}$. We thus show how we have obtained the main theoretical results from Sec. III B. We start from the delayed SNR matrix, $\left|\gamma_d^{(k)}/\xi_M^{(k)}\right|$, that was introduced in Sec. III B. If we average the SNR among realizations of the input story and the Hamiltonians, we observe that

$$\left\langle \left| \frac{\gamma_d^{(k)}}{\xi_M^{(k)}} \right| \right\rangle_{\mathbf{s},\hat{H}} = R^{d-1} \left\langle \left| \frac{\gamma_1^{(k)}}{\xi_M^{(k)}} \right| \right\rangle_{\mathbf{s},\hat{H}}. \quad \text{(G1)}$$

Equation (G1) is a key numerical observation and will be useful in the following arguments. For the rest of the section, the subindex $\mathbf{s}, \hat{H}$ and the label (k) are removed for simplicity. That is, when averaging over realizations

the magnitude of the ideal delayed terms differs only on a factor equal to R^{d-1}. In Eq. (G1) the γ_1 term can be further expanded as

$$\left\langle \left| \frac{\gamma_1}{\xi_M} \right| \right\rangle = \sqrt{M}(1-R)^2 \mathcal{C}_{\mathrm{SNR}}. \tag{G2}$$

In Eq. (G2) the term $\sqrt{M}$ comes from the dependency of M of the noise, while the term $(1-R)^2$ denotes the transmissivity dependence from Eq. (4) that has been already commented. The term $\mathcal{C}_{\mathrm{SNR}}$ is neither a function of M, R nor d, but can be dependent on all the other parameters (including N). From Eqs. (G1) and (G2) we are going to derive the main theoretical equations regarding the SNR.

1. Derivation of Eq. (6)

In this subsection, we derive an expression that accounts for how much we have to scale the number of measurements M to be able to resolve Δd steps in the past. We start by considering that we can already, with M measurements, properly resolve inputs up to a certain delay d. That is, the SNR of $\gamma_d^{(k)}$ for a statistical noise $\xi_M^{(k)}$ is high enough. We now impose the SNR of $\gamma_{d+\Delta d}$ to be equal to the one of $\gamma_d^{(k)}$. This would require a higher number of measurements, which we call M'. This equality can be written as follows:

$$\left\langle \left| \frac{\gamma_d}{\xi_M} \right| \right\rangle = \left\langle \left| \frac{\gamma_{d+\Delta d}}{\xi_{M'}} \right| \right\rangle. \tag{G3}$$

If we now substitute from Eqs. (G1) and (G2), we get

$$\sqrt{M} R^d = \sqrt{M'} R^{d+\Delta d}. \tag{G4}$$

By reordering and simplifying the equation above we end up with the following result:

$$\frac{M'}{M} = R^{-2\Delta d}, \tag{G5}$$

which is the one shown in Eq. (6). We can isolate the term Δd from Eq. (G5) to yield the following result:

$$\Delta d = \log_R \sqrt{\frac{M}{M'}}. \tag{G6}$$

This relation relates the measurement increase, M'/M, with the resolution enhancement, Δd. We note that a linear increase in the delay resolution yields an exponential increase in the number of measurements.

2. SNR main equations

The SNR lines shown in dBs in Figs. 3(b) and 3(c) have a simple mathematical representation. From Eq. (G1) we can infer that

$$10\log_{10}\left\langle \left| \frac{\gamma_d}{\xi_M} \right| \right\rangle = 10\log_{10}\left\langle \left| \frac{\gamma_1}{\xi_M} \right| \right\rangle + 10(d-1)\log_{10} R, \tag{G7}$$

which is just a linear equation of d. In visual representations, such as in Figs. 3(b) and 3(c), the slope corresponds to $10\log_{10} R$ and the height to $10\log_{10}\langle|\gamma_1 \xi_M^{-1}|\rangle$. Using Eq. (G2), we can further expand the SNR of $\gamma_1^{(k)}$ inside the logarithm as

$$10\log_{10}\left\langle \left| \frac{\gamma_1}{\xi_M} \right| \right\rangle = 10\left[\frac{1}{2}\log_{10} M + 2\log_{10}(1-R) + \log_{10}\mathcal{C}_{\mathrm{SNR}}\right]. \tag{G8}$$

The SNR thus has a simple behavior when changing the parameters M and R. For instance, changing the number of measurements from M to M' yields a difference on the SNR of $\gamma_1^{(k)}$ of

$$\left\langle \left| \frac{\gamma_1}{\xi_{M'}} \right| \right\rangle - \left\langle \left| \frac{\gamma_1}{\xi_M} \right| \right\rangle = 5\log_{10}\left(\frac{M'}{M}\right). \tag{G9}$$

For the case of changing the reflectivity from R to R', the difference on the $\gamma_1^{(k)}$ term is given by

$$\left\langle \left| \frac{\gamma_1(R')}{\xi_M} \right| \right\rangle - \left\langle \left| \frac{\gamma_1(R)}{\xi_M} \right| \right\rangle = 20\log_{10}\left(\frac{1-R'}{1-R}\right). \tag{G10}$$

The term multiplying $(d-1)$ (slope) also changes to $10\log_{10} R'$.

3. Derivation of Eq. (8)

In this subsection, we now consider a dependency with the size of the system, N. We aim to get an equation that allows us to improve the delay resolution quadratically without an exponential scaling of M. We argued that only the SNR of $\gamma_1^{(k)}$ has a dependency on N through the term $\mathcal{C}_{\mathrm{SNR}}$ in Eq. (G2). In the size intervals we have considered for this paper (up to $N = 12$), $\langle|\gamma_1 \xi_M^{-1}|\rangle$ remained practically constant with N, as it is shown in Fig. 7. We have thus taken the approximation in which we consider no dependency on N.

The strategy to get a quadratic increase in resolution as we increase N, which ensures that the normalized IPC at least remains constant, is conceived as follows. We know that for higher values of R it is less costly to resolve further

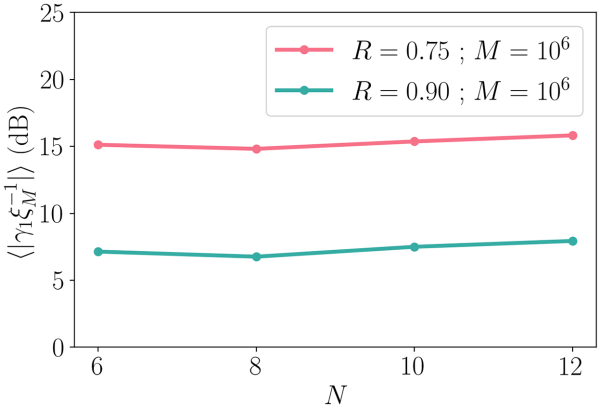

FIG. 7. SNR of $\gamma_1^{(k)}$ as a function of N. The different lines correspond to a different value of R, while $M = 10^6$ for both.

terms into the past. We are thus interested in increasing the value of R with N. We consider the following function:

$$R(N) = 1 - \frac{\mathcal{C}}{N^2}, \tag{G11}$$

in which $\mathcal{C}$ is an arbitrary constant. We recall the SNR equality from Eq. (G3) that we used to obtain Eq. (G5), taking into account that now neither R nor M remain constant. We take the delay so that, for a fixed SNR, $d(N) = \alpha N^2$, where α is an arbitrary constant. This will ensure that the resolution scales quadratically with N. We can rewrite the equality in Eq. (G3) as

$$\sqrt{M}(1-R)^2 R^{d-1} = \sqrt{M'}(1-R')^2 (R')^{d'-1}, \tag{G12}$$

where M, R, and d are functions of N and M', R', and d' are functions of N'. The equality above can be reordered as

$$\sqrt{\frac{M'}{M}} = \left(\frac{1-R}{1-R'}\right)^2 \frac{R^{d-1}}{(R')^{d'-1}}. \tag{G13}$$

We now turn our attention to the last term regarding the R^{d-1} and $(R')^{d'-1}$. If we substitute the chosen functions $R(N)$ and $d(N)$ in the term R^{d-1}, we can perform a series expansion of N^{-1}, yielding

$$\begin{aligned} R(N)^{d(N)-1} &= \left(1 - \frac{\mathcal{C}}{N^2}\right)^{\alpha N^2 - 1} \\ &= e^{-\mathcal{C}\alpha} + \mathcal{O}\left(N^{-2}\right). \end{aligned} \tag{G14}$$

Thus, the fraction of $R^{d-1}/(R')^{d'-1} \simeq 1$. With this, Eq. (G13) can be rewritten as

$$\frac{M'}{M} \simeq \left(\frac{N'}{N}\right)^8. \tag{G15}$$

We end up with $M(N) \propto N^8$, which is the polynomial scaling condition from Eq. (8).

[1] S. Slussarenko and G. J. Pryde, Photonic quantum information processing: A concise review, Appl. Phys. Rev. **6**, 041303 (2019).

[2] S. Takeda and A. Furusawa, Toward large-scale fault-tolerant universal photonic quantum computing, APL Photon. **4**, 060902 (2019).

[3] H. Wang, J. Qin, X. Ding, M. C. Chen, S. Chen, X. You, Y. M. He, X. Jiang, L. You, Z. Wang, C. Schneider, J. J. Renema, S. Höfling, C. Y. Lu, and J. W. Pan, Boson Sampling with 20 Input Photons and a 60-Mode Interferometer in a 1014-Dimensional Hilbert Space, Phys. Rev. Lett. **123**, 250503 (2019).

[4] H. S. Zhong, *et al.*, Quantum computational advantage using photons, Science **370**, 1460 (2020).

[5] J.-i. Yoshikawa, S. Yokoyama, T. Kaji, C. Sornphiphatphong, Y. Shiozawa, K. Makino, and A. Furusawa, Invited article: Generation of one-million-mode continuous-variable cluster state by unlimited time-domain multiplexing, APL Photon. **1**, 060801 (2016).

[6] Y. Cai, J. Roslund, G. Ferrini, F. Arzani, X. Xu, C. Fabre, and N. Treps, Multimode entanglement in reconfigurable graph states using optical frequency combs, Nat. Commun. **8**, 15645 (2017).

[7] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, Nat. Commun. **5**, 4213 (2014).

[8] P. L. McMahon, A. Marandi, Y. Haribara, R. Hamerly, C. Langrock, S. Tamate, T. Inagaki, H. Takesue, S. Utsunomiya, K. Aihara, R. L. Byer, M. M. Fejer, H. Mabuchi, and Y. Yamamoto, A fully programmable 100-spin coherent Ising machine with all-to-all connections, Science **354**, 614 (2016).

[9] T. Honjo, T. Sonobe, K. Inaba, T. Inagaki, T. Ikuta, Y. Yamada, T. Kazama, K. Enbutsu, T. Umeki, R. Kasahara, K. ichi Kawarabayashi, and H. Takesue, 100,000-spin coherent Ising machine, Sci. Adv. **7**, eabh0952 (2021).

[10] D. Pierangeli, G. Marcucci, and C. Conti, Large-Scale Photonic Ising Machine by Spatial Light Modulation, Phys. Rev. Lett. **122**, 213902 (2019).

[11] F. Böhm, G. Verschaffelt, and G. Van der Sande, A poor man's coherent Ising machine based on opto-electronic feedback systems for solving optimization problems, Nat. Commun. **10**, 3538 (2019).

[12] F. Wyffels and B. Schrauwen, A comparative study of reservoir computing strategies for monthly time series prediction, Neurocomputing **73**, 1958 (2010). subspace Learning / Selected papers from the European Symposium on Time Series Prediction,

[13] X. Lin, Z. Yang, and Y. Song, Short-term stock price prediction based on echo state networks, Expert Syst. Appl. **36**, 7313 (2009).

[14] I. Ilies, H. Jaeger, O. Kosuchinas, and M. Rincon, *Stepping forward through echoes of the past: Forecasting with echo state networks*, Tech. Rep. (2007).

[15] K. Nakajima and I. Fischer, *Reservoir Computing: Theory, Physical Implementations, and Applications* (Springer, Singapore, 2021).

[16] P. Coulibaly, Reservoir computing approach to great lakes water level forecasting, J. Hydrol. (Amst) **381**, 76 (2010).

[17] F. Triefenbach, A. Jalalvand, B. Schrauwen, and J.-p. Martens, in *Advances in Neural Information Processing Systems*, Vol. 23, edited by J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta (Curran Associates, Inc., 2010).

[18] L. Wang, Z. Wang, and S. Liu, An effective multivariate time series classification approach using echo state network and adaptive differential evolution algorithm, Expert Syst. Appl. **43**, 237 (2016).

[19] D. Verstraeten, B. Schrauwen, M. D'Haene, and D. Stroobandt, An experimental unification of reservoir computing methods, Neural Netw. **20**, 391 (2007). ,

[20] G. Tanaka, T. Yamane, J. B. Héroux, R. Nakane, N. Kanazawa, S. Takeda, H. Numata, D. Nakano, and A. Hirose, Recent advances in physical reservoir computing: A review, Neural Netw. **115**, 100 (2019).

[21] D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, Parallel photonic information processing at gigabyte per second data rates using transient states, Nat. Commun. **4**, 1364 (2013).

[22] K. Vandoorne, P. Mechet, T. Van Vaerenbergh, M. Fiers, G. Morthier, D. Verstraeten, B. Schrauwen, J. Dambre, and P. Bienstman, Experimental demonstration of reservoir computing on a silicon photonics chip, Nat. Commun. **5**, 3541 (2014).

[23] L. Larger, A. Baylón-Fuentes, R. Martinenghi, V. S. Udaltsov, Y. K. Chembo, and M. Jacquot, High-Speed Photonic Reservoir Computing Using a Time-Delay-Based Architecture: Million Words per Second Classification, Phys. Rev. X **7**, 011015 (2017).

[24] G. Van Der Sande, D. Brunner, and M. C. Soriano, Advances in photonic reservoir computing, Nanophotonics **6**, 561 (2017).

[25] P. Mujal, R. Martínez-Peña, J. Nokkala, J. García-Beni, G. L. Giorgi, M. C. Soriano, and R. Zambrini, Opportunities in quantum reservoir computing and extreme learning machines, Adv. Quantum Technol. **4**, 2100027 (2021).

[26] S. Ghosh, K. Nakajima, T. Krisnanda, K. Fujii, and T. C. H. Liew, Quantum neuromorphic computing with reservoir computing networks, Adv. Quantum Technol. **4**, 2100053 (2021).

[27] D. Marković, A. Mizrahi, D. Querlioz, and J. Grollier, Physics for neuromorphic computing, Nat. Rev. Phys. **2**, 499 (2020).

[28] P. Mujal, R. Martínez-Peña, G. L. Giorgi, M. C. Soriano, and R. Zambrini, Time-series quantum reservoir computing with weak and projective measurements, npj Quantum Inf. **9**, 16 (2023).

[29] J. Chen, H. I. Nurdin, and N. Yamamoto, Temporal Information Processing on Noisy Quantum Computers, Phys. Rev. Appl. **14**, 024065 (2020).

[30] J. Nokkala, R. Martínez-Peña, G. L. Giorgi, V. Parigi, M. C. Soriano, and R. Zambrini, Gaussian states of continuous-variable quantum systems provide universal and versatile reservoir computing, Commun. Phys. **4**, 53 (2021).

[31] R. Medeiros de Araújo, J. Roslund, Y. Cai, G. Ferrini, C. Fabre, and N. Treps, Full characterization of a highly multimode entangled state embedded in an optical frequency comb using pulse shaping, Phys. Rev. A **89**, 053828 (2014).

[32] J. Roslund, R. M. de Araújo, S. Jiang, C. Fabre, and N. Treps, Wavelength-multiplexed quantum networks with ultrafast frequency combs, Nat. Photon. **8**, 109 (2014).

[33] J. Nokkala, F. Arzani, F. Galve, R. Zambrini, S. Maniscalco, J. Piilo, N. Treps, and V. Parigi, Reconfigurable optical implementation of quantum complex networks, New J. Phys. **20**, 053024 (2018).

[34] K. Fujii and K. Nakajima, Harnessing Disordered-Ensemble Quantum Dynamics for Machine Learning, Phys. Rev. Appl. **8**, 024030 (2017).

[35] R. Martínez-Peña, G. L. Giorgi, J. Nokkala, M. C. Soriano, and R. Zambrini, Dynamical Phase Transitions in Quantum Reservoir Computing, Phys. Rev. Lett. **127**, 100502 (2021).

[36] R. A. Bravo, K. Najafi, X. Gao, and S. F. Yelin, Quantum Reservoir Computing using Arrays of Rydberg Atoms, PRX Quantum **3**, 030325 (2022).

[37] J. Nokkala, Online quantum time series processing with random oscillator networks (2021), ArXiv:2108.00698.

[38] Y. Shaked, Y. Michael, R. Z. Vered, L. Bello, M. Rosenbluh, and A. Pe'er, Lifting the bandwidth limit of optical homodyne measurement with broadband parametric amplification, Nat. Commun. **9**, 609 (2018).

[39] N. Takanashi, A. Inoue, T. Kashiwazaki, T. Kazama, K. Enbutsu, R. Kasahara, T. Umeki, and A. Furusawa, All-optical phase-sensitive detection for ultra-fast quantum computation, Opt. Express **28**, 34916 (2020).

[40] H. M. Wiseman and G. J. Milburn, Quantum theory of field-quadrature measurements, Phys. Rev. A **47**, 642 (1993).

[41] H. M. Wiseman and G. J. Milburn, Quantum Theory of Optical Feedback via Homodyne Detection, Phys. Rev. Lett. **70**, 548 (1993).

[42] T. Kouadou, F. Sansavini, M. Ansquer, J. Henaff, N. Treps, and V. Parigi, Spectrally shaped and pulse-by-pulse multiplexed multimode squeezed states of light (2022), ArXiv:2209.10678.

[43] L. S. Madsen, F. Laudenbach, M. F. Askarani, F. Rortais, T. Vincent, J. F. F. Bulmer, F. M. Miatto, L. Neuhaus, L. G. Helt, M. J. Collins, A. E. Lita, T. Gerrits, S. W. Nam, V. D. Vaidya, M. Menotti, I. Dhand, Z. Vernon, N. Quesada, and J. Lavoie, Quantum computational advantage with a programmable photonic processor, Nature **606**, 75 (2022).

[44] H. Tomoda, T. Yoshida, T. Kashiwazaki, T. Umeki, Y. Enomoto, and S. Takeda, Programmable time-multiplexed squeezed light source, Opt. Express **31**, 2161 (2023).

[45] M. Chen, N. C. Menicucci, and O. Pfister, Experimental Realization of Multipartite Entanglement of 60 Modes of a Quantum Optical Frequency Comb, Phys. Rev. Lett. **112**, 120505 (2014).

[46] W. N. Plick, F. Arzani, N. Treps, E. Diamanti, and D. Markham, Violating bell inequalities with entangled optical frequency combs and multipixel homodyne detection, Phys. Rev. A **98**, 062101 (2018).

[47] Y. Cai, J. Roslund, V. Thiel, C. Fabre, and N. Treps, Quantum enhanced measurement of an optical frequency comb, npj Quantum Inf. **7**, 82 (2021).

[48] M. V. Larsen, X. Guo, C. R. Breum, J. S. Neergaard-Nielsen, and U. L. Andersen, Fiber-coupled EPR-state generation using a single temporally multiplexed squeezed light source, npj Quantum Inf. **5**, 46 (2019).

[49] J. Dambre, D. Verstraeten, B. Schrauwen, and S. Massar, Information processing capacity of dynamical systems, Sci. Rep. **2**, 514 (2012).

[50] R. Martínez-Peña, J. Nokkala, G. L. Giorgi, R. Zambrini, and M. C. Soriano, Information processing capacity of spin-based quantum reservoir computing systems, Cognitive Computation (2020),.

[51] H. Vahlbruch, M. Mehmet, K. Danzmann, and R. Schnabel, Detection of 15 dB Squeezed States of Light and Their Application for the Absolute Calibration of Photoelectric Quantum Efficiency, Phys. Rev. Lett. **117**, 110801 (2016).

[52] J. Nokkala, R. Martínez-Peña, R. Zambrini, and M. C. Soriano, High-performance reservoir computing with fluctuations in linear networks, IEEE Trans. Neural Netw. Learn. Syst. **33**, 2664 (2022).

[53] L. C. G. Govia, G. J. Ribeill, G. E. Rowlands, H. K. Krovi, and T. A. Ohki, Quantum reservoir computing with a single nonlinear oscillator, Phys. Rev. Res. **3**, 013077 (2021).

[54] W. D. Kalfus, G. J. Ribeill, G. E. Rowlands, H. K. Krovi, T. A. Ohki, and L. C. G. Govia, Hilbert space as a computational resource in reservoir computing, Phys. Rev. Res. **4**, 033007 (2022).

[55] U. Hübner, N. B. Abraham, and C. O. Weiss, Dimensions and entropies of chaotic intensity pulsations in a single-mode far-infrared nh_3 laser, Phys. Rev. A **40**, 6354 (1989).

[56] A. Weigend and N. Gershenfeld, in *IEEE International Conference on Neural Networks* (1993), p. 1786, vol.3.

[57] M. Inubushi and K. Yoshimura, Reservoir computing beyond memory-nonlinearity trade-off, Sci. Rep. **7**, 10199 (2017).

[58] K. Harkhoe and G. Van der Sande, Delay-based reservoir computing using multimode semiconductor lasers: Exploiting the rich carrier dynamics, IEEE J. Sel. Top. Quantum Electron. **25**, 1502909 (2019).

[59] P. Kumar, M. Jin, T. Bu, S. Kumar, and Y.-P. Huang, Efficient reservoir computing using field programmable gate array and electro-optic modulation, OSA Continuum **4**, 1086 (2021).

[60] S. Marzen, Difference between memory and prediction in linear recurrent networks, Phys. Rev. E **96**, 032308 (2017).

[61] D. Brunner, M. C. Soriano, and G. Van der Sande, *Photonic Reservoir Computing* (De Gruyter, Berlin, 2019).

[62] M. Spagnolo, J. Morris, S. Piacentini, M. Antesberger, F. Massa, A. Crespi, F. Ceccarelli, R. Osellame, and P. Walther, Experimental photonic quantum memristor, Nat. Photon. **16**, 318 (2022).

[63] S. A. Khan, F. Hu, G. Angelatos, and H. E. Türeci, Physical reservoir computing using finitely-sampled quantum systems (2021), ArXiv:2110.13849.

[64] M. Lukoševičius and H. Jaeger, Reservoir computing approaches to recurrent neural network training, Comput. Sci. Rev. **3**, 127 (2009).

[65] Z. Konkoli, in *Advances in Unconventional Computing: Volume 1: Theory*, edited by A. Adamatzky (Springer International Publishing, Cham, 2017), p. 573.

[66] G. Adesso, S. Ragy, and A. R. Lee, Continuous variable quantum information: Gaussian states and beyond, Open Syst. Inf. Dyn. **21**, 1440001 (2014).

[67] A. Serafini, *Quantum Continuous Variables: A Primer of Theoretical Methods* (CRC Press, Boca Raton, 2017).

[68] M. G. Genoni, L. Lami, and A. Serafini, Conditional and unconditional Gaussian quantum dynamics, Contemp. Phys. **57**, 331 (2016).

[69] J. Eisert, S. Scheel, and M. B. Plenio, Distilling Gaussian States with Gaussian Operations is Impossible, Phys. Rev. Lett. **89**, 137903 (2002).

[70] S. L. Braunstein, Squeezing as an irreducible resource, Phys. Rev. A **71**, 055801 (2005).

[71] G. Cariolaro and G. Pierobon, Reexamination of Bloch-Messiah reduction, Phys. Rev. A **93**, 062115 (2016).

[72] P. Mujal, J. Nokkala, R. Martínez-Peña, G. L. Giorgi, M. C. Soriano, and R. Zambrini, Analytical evidence of nonlinearity in qubits and continuous-variable quantum reservoir computing, J. Phys.: Complexity **2**, 045008 (2021).

[73] L. C. G. Govia, G. J. Ribeill, G. E. Rowlands, and T. A. Ohki, Nonlinear input transformations are ubiquitous in quantum reservoir computing, Neuromorphic Comput. Eng. **2**, 014008 (2022).