Mask-Less Asynchronous Time-Delay Reservoir Computing Using a Passive Photonic Integrated Circuit

Mohab Abdalla, Raphael Cardoso, Paul Jimenez, Mauricio Gomes de Queiroz, Andreas Boes, Guanghui Ren, Arnan Mitchell, Ian O'Connor, Fabio Pavanello

To cite this version:

Mohab Abdalla, Raphael Cardoso, Paul Jimenez, Mauricio Gomes de Queiroz, Andreas Boes, et al.. Mask-Less Asynchronous Time-Delay Reservoir Computing Using a Passive Photonic Integrated Circuit. Journal of Lightwave Technology, 2024, 10.1109/JLT.2024.3434538 . hal-04739811

HAL Id: hal-04739811
https://hal.science/hal-04739811v1

Submitted on 16 Oct 2024

This article has been accepted for publication in IEEE/OSA Journal of Lightwave Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JLT.2024.3434538

Mask-less asynchronous time-delay reservoir computing using a passive photonic integrated circuit

Mohab Abdalla, Raphael Cardoso, Paul Jimenez, Mauricio Gomes de Queiroz, Andreas Boes, Guanghui Ren, Arnan Mitchell, Ian O'Connor, Fabio Pavanello

***Abstract*—Photonic time-delay reservoir computing schemes usually employ an input mask as a means of performance enhancement. However, input masking usually necessitates a domain conversion, requiring a signal to be treated before sending it to the reservoir. More recent implementations have explored further ways of performance enhancement, whether through operating in the asynchronous regime, or by using post-filtering approaches. In this numerical study, we analyze the task-independent performance of a passive integrated photonic reservoir, and show that it can achieve good results on some benchmark tasks in the absence of an input mask. We also consider the effects of post-filtering and operating in the asynchronous regime through a parameter space exploration. The proposed scheme enables ultra-fast processing speeds while simultaneously reducing the associated power and complexity costs of the associated electronics. We compare the obtained results with the case of using a mask, and also with other schemes from the literature, showing comparable performance on the investigated tasks.**

***Index Terms*—Reservoir computing, integrated photonics.**

I. INTRODUCTION

RESERVOIR computing (RC) is a machine learning approach which exploits nonlinear dynamics to project data onto higher dimensional spaces. This makes it possible to separate multiple classes with simple linear classifiers, even for tasks that would normally require larger and deeper neural networks. Historically, the motivation was to circumvent the problems of training large-scale artificial neural networks, especially recurrent ones, implemented in software. Currently, the interest has shifted more towards physical RC implementations which leverage the dynamics of physical systems to do the computing, and which have been explored in various domains such as electronics and photonics, among others [1]. Time delay reservoir computing (TDRC) extends RC to time-multiplexing approaches through the use of delayed dynamical systems [2] as time-multiplexed cyclic reservoir networks [3]. This reduces the physical nodes requirement down to a single node in feedback configuration, which simplifies hardware implementation, but comes at the expense of computational speed. The output of this node is sampled in time to yield independent states which express different features of the reservoir's response to a given input, where the feedback length is usually designed to match approximately the input rate. To achieve high speed RC, it is important to choose dynamical systems with a short enough timescale, but which can still be captured by today's electronic readouts. This is the case with passive all-optical reservoirs [4], [5]. Photonics is naturally well-suited for RC, owing to opportunities in speed, efficiency, and parallelism [6]. This has led to increasing interest in photonic RC implementations, whether in bulk [7], [8], [9], [10], on-chip [4], [11], [12], [13], or hybrid schemes [14], [15].

Recently, some interest has been directed towards the exploration of the asynchronous regime [16], [17], [18], where the delay time of the feedback loop is unconventionally chosen to be neither equivalent nor close to that of the input rate. This synchronicity convention stemmed from the initial approach of viewing TDRC systems as the network equivalent of spatially multiplexed reservoirs, with the nodes connected in a cyclic fashion [3]. Another conventional preference has been the use of an input masking protocol to prolong a reservoir's activity and thus improve performance, but which introduces complications in practical cases as well as associated power costs, in addition to a bottleneck for ultra-high speed signal processing [19]. In this study, we tackle this problem by dispensing with the input mask altogether, which allows the bypassing of a possible optical/electronic/optical conversion at the input, in case the signal to be fed to the reservoir is already in the optical domain (for example in telecom applications [20]). We use task-independent metrics alongside standard benchmark tasks to judge the performance of an all-optical integrated reservoir over its design space, with a specific focus on the feedback time and applied phase shift. The considered photonic reservoir has an ultrashort timescale governed by the waveguide lengths. Thus the scheme is only limited by the readout electronics' RF bandwidth. [5]. We also take into account the associated effects of the electronic readout in our study, mainly the photodiode noise and filtering with subsequent 12-bit quantization, to understand its implications on the performance. Lastly, we show the correlation between the performance on benchmark tasks and the different task-independent metrics within the explored design space.

M. Abdalla, R. Cardoso, P. Jimenez, M.G. de Queiroz, Ian O'Connor are with Univ. Lyon, Ecole Centrale de Lyon, INSA Lyon, UCB Lyon, CPE Lyon, CNRS, Lyon Institute of Nanotechnology, UMR5270, Ecully, 69130, France.

M. Abdalla, G. Ren, A. Mitchell are with Integrated Photonics and Applications Centre (InPAC), School of Engineering, RMIT University, Melbourne, VIC 3000, Australia.

A. Boes is with the Institute for Photonics and Advanced Sensing (IPAS), University of Adelaide, Adelaide, SA 5005, Australia.

F. Pavanello is with Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, Grenoble INP, IMEP-LAHC, 38000, Grenoble, France.

E-mail: mohab.abdalla@ec-lyon.fr, fabio.pavanello@cnrs.fr

II. METHODS

The performance of an RC scheme can be quantified in a number of ways; the most obvious is by using standard benchmark tasks. However, this constrains the evaluation to these specific tasks and does not provide information about the reservoir's overall memory/nonlinearity capabilities. To give a more comprehensive account of the RC performance, task-independent metrics must be employed as well. The information processing capacity (IPC) [21] is a task-independent metric which has been used to quantify the performance of a number of reservoirs in a number of dedicated studies [22], [23], [24], [25]. However, when considering a large design space (as in our case), IPC calculation would require large computational time and resources. Instead, we opted for a combination of other metrics which can reflect the memory/nonlinearity abilities of a reservoir, while requiring less computing resources.

A. Task-independent tests

The first metric considered is the linear memory capacity [26], which showcases the ability to reconstruct past inputs using a weighted linear combination of the reservoir's presently probed states. An input sequence u of length 2000, drawn from a uniform distribution $\in [0, 0.5)$, is used to train the model with linear regression to reconstruct k inputs into the past. The performance is then evaluated with the square of the Pearson correlation coefficient r_{c}^2:

$$\mathrm{MC}_k = \frac{\mathrm{cov}(u[n-k], y_k[n])^2}{\mathrm{var}(u[n]) var(y_k[n]))} = r_{\mathrm{c}}^2(u[n-k], y_k[n]) \quad (1)$$

where $\mathrm{cov}(.)$ is the covariance, $\mathrm{var}(.)$ is the variance, $u[n]$ is the input at the current step, $u[n-k]$ is the input delayed by k discrete steps, and $y_k[n]$ is the predicted stream (after training). For this study, we calculate the total linear memory capacity $\mathrm{MC}_{\mathrm{tot}}$ by summing up the first $m = 100$ terms. Theoretically speaking, the sum includes the infinite past, i.e. all the $m \in [1, \infty)$ are to be considered. A similar truncation strategy to [27] is used, where we do not consider the terms where $\mathrm{MC}_k < 0.01$.

$$\mathrm{MC}_{\mathrm{tot}} = \sum_{k=1}^{m} \mathrm{MC}_k \quad (2)$$

The second metric is the computational ability (CA), which is a measure of RC nonlinearity strength or computation power. It is calculated by performing two tests on the reservoir: the kernel quality and the generalization tests [28], [29]. The kernel quality test evaluates how the reservoir can map different inputs to sufficiently different states. For this test, we construct 100 input streams, with each being a sequence of length 100, drawn from an independent and identical distribution with values $\in [0, 0.5)$. All 20 nodes corresponding to the 100th input are then collected to construct a 100×20 matrix (#input streams$\times N$). The generalization test reflects how well the reservoir can map the same input samples under different initial conditions. Using the previously generated 100 different input streams, we concatenate to each an identical sequence of 10 inputs, and we collect the node responses corresponding to the 110th output sample to construct a 100×20 matrix. Using singular value decomposition with the appropriate threshold [30], the ranks of their corresponding matrices are then found. The normalized CA is then calculated:

$$\mathrm{CA} = (\mathrm{KQR} - \mathrm{GR})/N \quad (3)$$

where KQR is the kernel quality matrix rank (from the kernel quality test), GR is the generalization matrix rank (from the generalization test), and N is the number of nodes in the reservoir. Thus, high KQR and low GR are desirable.

B. Benchmark tasks

Each realization over the entire design space was trained on 3 benchmark tasks: NARMA-k, XOR-k, and the Santa Fe dataset [31]. The NARMA-k task is given by:

$$y[n] = 0.3y[n-1] + 0.05y[n-1]\sum_{i=1}^{k} y[n-i] + 1.5u[n]u[n-k] + 0.1 \quad (4)$$

where the input sequence u is drawn from a uniform distribution $\in [0, 0.5)$. The task is to predict $y[n]$ given $u[n]$. The performance on these tasks is evaluated by the normalized mean square error (NMSE), which is given by:

$$\mathrm{NMSE} = \frac{1}{n\sigma_{y_{\mathrm{t}}}^2}\sum_{i=1}^{n}(y_{\mathrm{t}} - y_{\mathrm{p}})^2 \quad (5)$$

where y_{t} is the target (true) value and y_{p} is the predicted value, n is the number of data points, and $\sigma_{y_{\mathrm{t}}}^2$ is the variance of the sequence of true values. For the XOR task, the commonly used bit error rate (BER) was also evaluated, which is given by:

$$\mathrm{BER} = \frac{\text{\#incorrectly classified bits}}{\text{length of test set}} \quad (6)$$

Training for all the benchmark tasks was done using ridge regression, with the regularization term $\alpha = 10^{-4}$ to prevent overfitting. For the linear memory capacity, linear regression was employed since this task does not require a test set. For the NARMA-k task, an input sequence of length 3500 was used for training and 500 for testing. For the XOR=k task, 1000 bits were used for training and 3000 for testing. For the Santa Fe task, the training sequence was 3000 samples, while 1000 were used for testing.

C. Simulation setup

The considered photonic circuit [5], shown in Fig. 1, consists of passive waveguides and 3-dB directional couplers on a low-loss, integrated photonic platform, such as silicon nitride ($\mathrm{Si_3N_4}$) [32] or lithium niobate ($\mathrm{LiNbO_3}$) on insulator [33]. The circuit is modeled using the transfer matrix approach and numerically solved with the following delayed coupled equations, which describe the temporal evolution of the electric field everywhere in the system (Fig. 1).

$$\begin{bmatrix} E_1(t) \\ E_2(t) \end{bmatrix} = \sqrt{\gamma_c}\begin{bmatrix} -i\kappa & r \\ r & -i\kappa \end{bmatrix}\begin{bmatrix} E_{\mathrm{in}}(t) \\ E_{\mathrm{fb}}(t) \end{bmatrix} \quad (7)$$

This article has been accepted for publication in IEEE/OSA Journal of Lightwave Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JLT.2024.3434538

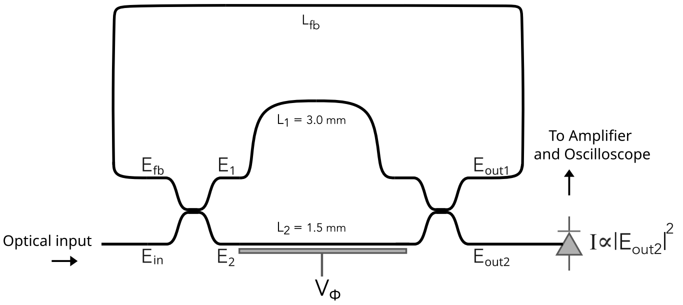


Fig. 1. Passive photonic reservoir. Depending on the desired L_{fb}, the feedback waveguide is either straight, with bends, or in a spiral configuration.

$$\begin{bmatrix} E_{\mathrm{out1}}(t) \\ E_{\mathrm{out2}}(t) \end{bmatrix} = \sqrt{\gamma_c} \begin{bmatrix} \sqrt{\gamma_1} r e^{-i\beta L_1} & -\sqrt{\gamma_2} i\kappa e^{-i(\beta L_2+\Phi)} \\ -\sqrt{\gamma_1} i\kappa e^{-i\beta L_1} & \sqrt{\gamma_2} r e^{-i(\beta L_2+\Phi)} \end{bmatrix} \times \begin{bmatrix} E_1(t-\tau_1) \\ E_2(t-\tau_2) \end{bmatrix} \tag{8}$$

$$E_{\mathrm{fb}}(t) = \sqrt{\gamma_{\mathrm{fb}}} E_{\mathrm{out1}}(t-\tau_{\mathrm{fb}}) \exp\left(-i\beta L_{\mathrm{fb}}\right) \tag{9}$$

Eqs. 7-9 show the circuit model, where γ_c is the percentage of optical power exiting from the coupler, $\gamma_{1,2,\mathrm{fb}} = 10^{-AL/10}$ are the fractions of power after waveguide propagation for a loss factor A [dB/m] and waveguide lengths $L_{1,2,\mathrm{fb}}$ [m], corresponding to the lengths of the upper MZI arm, bottom MZI arm, and the feedback loop, respectively. Since we are considering 50/50 couplers, the cross and through field coupling coefficients are equivalent ($\kappa = r$). For waveguide parameters, $\beta = 2\pi n_{\mathrm{eff}}/\lambda_0$ [m^{-1}] is the propagation constant of the guided mode with effective refractive index n_{eff}, $\tau_{1,2,fb}$ [s] are the delay times of the upper MZI arm, bottom MZI arm, and feedback waveguide, respectively, and Φ [rad] is the applied phase shift on the bottom arm. The application of phase shift through e.g. the voltage on a heater allows changing the reservoir's dynamics through controlling both the feedback strength and phase by adjusting the interference at the directional couplers. For the electronic readout, the photodetector is modeled by adding the associated noise components of shot noise and thermal noise, and then low pass-filtering the signal with the cutoff as the chosen photodetector bandwidth. The signal is then converted to a voltage through a transimpedence amplifier. Finally, the signal is binned to yield the final output as if it were passed through a 12-bit analog-to-digital converter (ADC), as would be found on most high-end digital oscilloscopes. The photodetector noise is modeled by:

$$\sigma_{\mathrm{th}}^2 = 4k_{\mathrm{B}}Tf_{\mathrm{c}}/R_{\mathrm{l}} \tag{10}$$

$$\sigma_{\mathrm{sh}}^2 = 2q(I_{\mathrm{p}} + I_{\mathrm{d}})f_{\mathrm{c}} \tag{11}$$

$$\sigma_{\mathrm{tot}}^2 = \sigma_{\mathrm{th}}^2 + \sigma_{\mathrm{sh}}^2 \tag{12}$$

$$I_{\mathrm{p}} = r_{\mathrm{p}}P_{\mathrm{opt}} \tag{13}$$

$$V_{\mathrm{p}} = R_{\mathrm{l}}I_{\mathrm{p}} \tag{14}$$

where σ_{th}^2 is the thermal noise variance, σ_{sh}^2 is the shot noise variance, σ_{tot}^2 is the total noise variance, k_{B} is the Boltzmann constant, T is the temperature, f_{c} is the photodetector 3-dB bandwidth (cutoff-frequency), R_{l} is the load resistance, q is the electronic charge, I_{p} is the current proportional to the input optical power P_{opt} through responsivity r_{p}, I_{d} is the dark current, and V_{p} is the voltage converted through load resistance R_{l}. The voltage after passing through the ADC is then quantized by

$$V_{\mathrm{Q}} = \mathrm{floor}\left(\frac{V_{\mathrm{p}}}{\Delta_{\mathrm{q}}}\right) \times \Delta_{\mathrm{q}} \tag{15}$$

where Δ_{q} is the ADC quantization step, given by $\Delta_{\mathrm{q}} = \left(V_{\mathrm{RefHi}} - V_{\mathrm{RefLo}}\right)/(2^Q - 1)$, which depends on the considered voltage range and number of bits of resolution Q of the ADC.

III. RESULTS AND DISCUSSION

We consider the passive all-optical integrated reservoir scheme shown in Fig. 1 operating with 10 mW laser power ($\lambda = 1550$ nm) at a fixed datarate of 10 GBit/s, which corresponds to a synchronized delay time $\tau = 100$ ps. The reservoir output is then sampled at 200 GSa/s to yield 20 time-multiplexed nodes per input clock cycle. We sweep the feedback time (50 values), in the form of the feedback waveguide's length, from 0.2τ (corresponding to approximately 3.0 mm on Si$_3$N$_4$ and 2.65 mm on LNOI) to 5τ (approximately 7.51 cm on Si$_3$N$_4$ and 6.62 cm on LNOI), and the applied phase shift (49 values) on the bottom MZI arm from 0 to 2π. The performance is then recorded with the relevant metric for each test/task. Further information about the simulation parameters can be found in Table II in Appendix A.

A. Dispensing with the input mask

An input masking protocol is generally used in TDRC schemes to keep the reservoir continuously perturbed to delay it from falling to a steady state, and thereby aiding in improving the reservoir's memory [34], [35]. However, there are limitations imposed by the use of a mask in practice (especially in photonic implementations) which call for alternative approaches [19]. Masks normally have to be prepared electronically, which may necessitate a domain conversion at the input, instead of allowing an incoming optical signal to be directly processed by the reservoir. Moreover, for high speed implementations in the GHz range, the input mask preparation not only introduces an associated energy cost, but can also put additional constraints on modulation speeds, especially on the arbitrary waveform generator's output samplerate, which would be used to realize the mask. For example, a 10 GSa/s input signal, when masked for a reservoir of 20 nodes, would require a 200 GSa/s signal, which is not feasible.

One solution to avoid the input mask is by shortening the feedback delay time τ_{fb} to fractions of the input clock cycle τ. If the reservoir's timescale is short enough - as in the case of passive optical reservoirs - this would allow the same input sample to interact with past version(s) of itself, such that the system is kept continuously perturbed during the interval of one input sample.

The effect produced can be explained in two complementary ways:

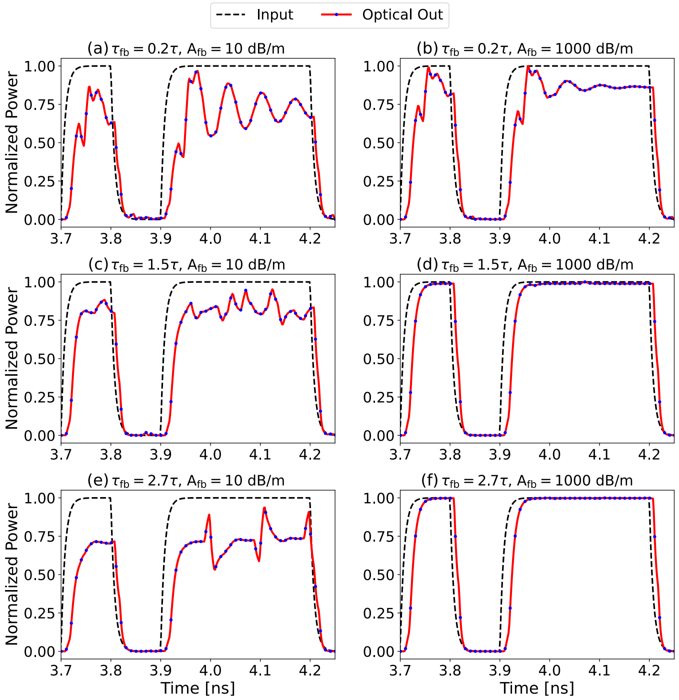


Fig. 2. Effect of τ_{fb} and A_{fb} on the dynamics and effective dimensionality of the reservoir. Blue dots represent the sampled nodes. τ is the delay time synchronized to the input clock cycle (100 ps).

1) The input is interacting with itself and thus increases the complexity, or ’richness’ of the output signal;
2) The shorter feedback lengths allow the reservoir to more quickly ‘forget’ inputs further in the past, resulting in a lower GR, which in turn increases CA.

It is especially important that the feedback losses A_{fb} are kept to a minimum, as this prolongs the reservoir’s activity and delays the time taken to reach steady state (Fig. 2(a,c)). As shown in Fig. 2(b,d,f), higher A_{fb} results in many nodes sharing the same state, which means that the nodes are not linearly independent and thus reduces the effective dimensionality of the reservoir. Furthermore, the delay time also affects the effective reservoir dimensionality, as longer delays (with respect to input clock cycle) would increase the amount of nodes that share the same state (Fig. 2(e)). One must also ensure that the node states are sufficiently distinguishable above the system’s noise floor, which is primarily resulting from the electronic readout and photodetector sensitivity. Furthermore, signal filtering also affects the effective dimensionality of the reservoir, especially if the timescale of the filtering is close to the input clock cycle or slower. For this reason, it is of interest in this study to dispense with the masking procedure and observe how well the reservoir can solve some tasks when fed the input signal directly, in the presence of the electronic readout. As will be seen, a signal interacting with itself in the sub-τ regime can provide the necessary computational power to carry out some tasks with moderate memory requirements.

B. Task-independent performance

In Fig. 3, we show the performance of the reservoir with different feedback lengths and under applied phase shift from 0 to 2π on $\mathrm{MC_{tot}}$, CA, and its constituent metrics GR and KQR. Here, we consider all the outputs passing through a photodetector with 40 GHz bandwidth and a 12-bit ADC. For feedback times less than τ, it is observed that $\mathrm{MC_{tot}}$ is low in general, while CA is in general high. A closer look at the constituent metrics of CA (Fig. 3(c,d)) shows that GR lowers with decreasing feedback time, especially in the sub-τ region, while also showing some dependence on the applied phase shift. However, for KQR, the variation is smaller over the entire design space, with an especially lower rank in the sub-τ region. Since GR becomes considerably lower in this region, the overall CA is highest. Less variation is observed in KQR when compared to GR. These variations are phase-dependent as the nonlinearity is effectively performed by the strength of the interference between the fields at the couplers, which is controlled by the phase shifter. Overall, these results show that there is opportunity in scaling down the feedback length while having enough memory and high CA to do some tasks in the absence of input masking, as will be explored in the following sections.

C. Benchmark task: NARMA-k

In this section, we consider the performance on the nonlinear autoregressive moving average (NARMA-k) benchmark task. This task requires both memory and nonlinearity. By varying the maximum number of memory steps k, we challenge the reservoirs in the design space with different memory/nonlinearity requirements. Furthermore, we also consider the temporal XOR-k task ($u[n] \oplus u[n-k]$) and the Santa Fe chaotic laser prediction, which are not displayed here for the sake of brevity, but which are taken into account in sections III-D and III-F. For all NARMA-k, a thresholding trend can be observed for a feedback time longer than the most distant memory. This is due to the absence of interaction between the first input and earlier inputs into the past, which are essentially skipped. Since the NARMA-k task considers the aggregate of previous inputs (Eq. 4), it depends on the total memory until the chosen k. This is observed in Fig. 3 where, beyond a certain feedback time, some earlier components of the memory are missing and thus lead to significantly worse performance. This sensitivity to earlier k reduces as k increases (as more terms are being aggregated). Additionally, for k’s which are further back in time, their attenuated power in the circuit becomes more sensitive to readout noise which is the reason this task does not achieve a low enough NMSE for k higher than 5. In the sub-τ region, it can be seen that a low NMSE can be achieved for all reported NARMA-k, however with differing trends due to the different memory/nonlinearity requirements from the RC system. Furthermore, upon closer examination, there seems to be a striking resemblance in pattern between both NARMA-4 and NARMA-5, and CA and GR results in Fig. 3(b,c). This is explored further in the section on correlation (section III-F). The effect of reservoir size on the performance of this benchmark task is also investigated in the supplementary material.

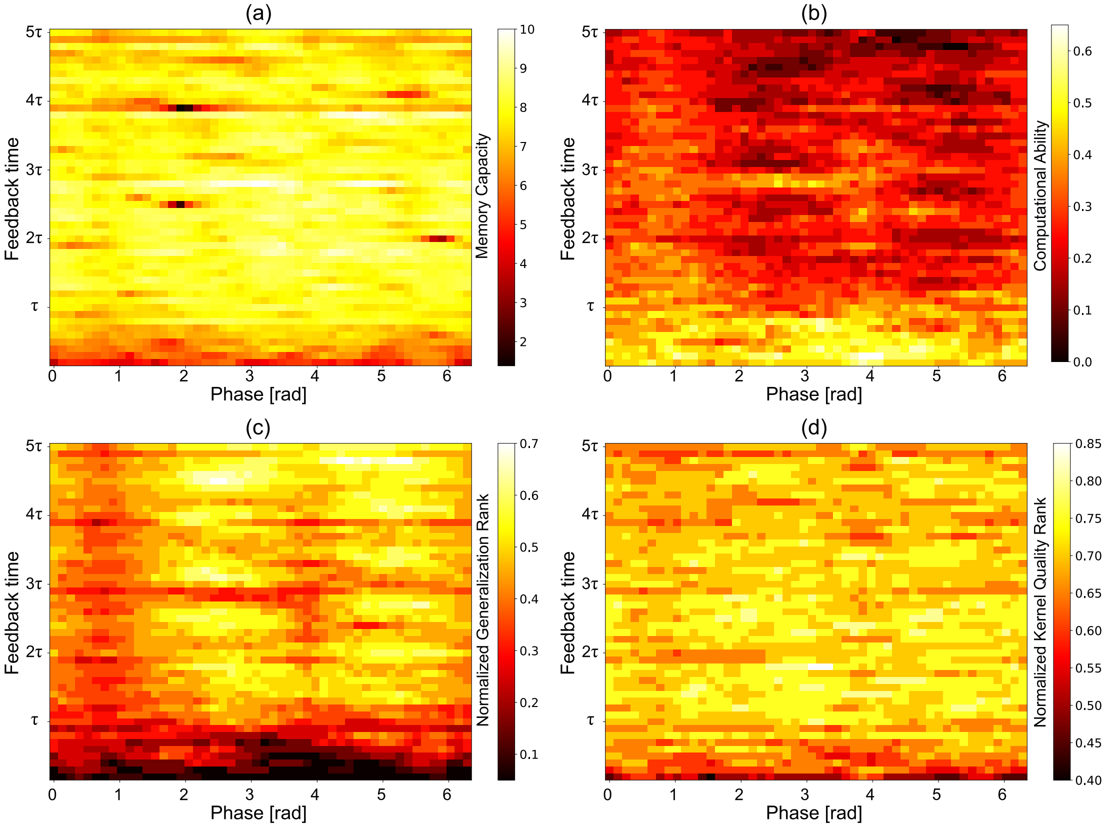


Fig. 3. Results on task-independent tests for different feedback delay τ_{fb}: (a) MC_{tot}, (b) CA, (c) GR, and (d) KQR.

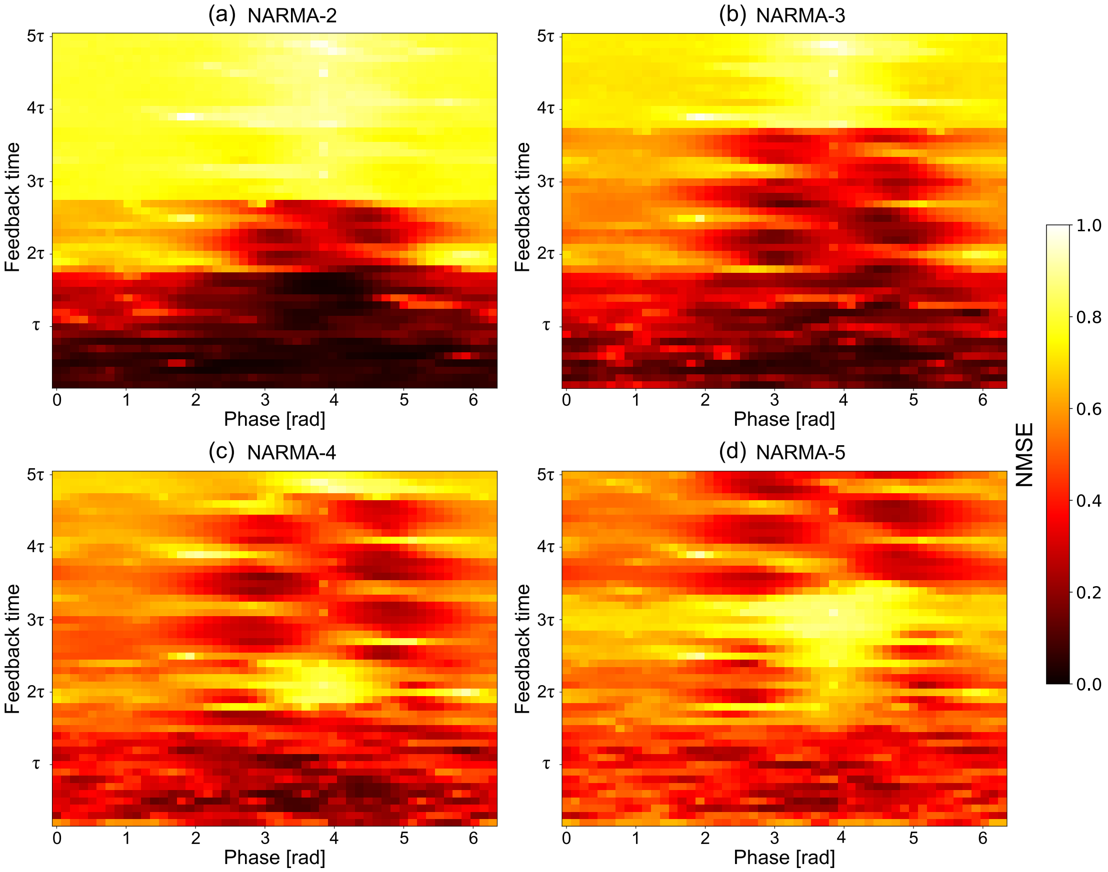


Fig. 4. Performance over the design space on the NARMA-k benchmark task. Small τ_{fb} achieve excellent performance, especially for lower k.

D. Performance comparison

In this section, we compare the obtained results on the mask-less protocol with results simulating the same architecture considering a pseudo-randomly generated mask from an i.i.d. $\in (0,1]$, which is synchronized to the 20 nodes of the reservoir. Here, the same mask is considered for all the reported tasks. Even though in practice it would be difficult to generate such a mask due to limitations on the RF signal generation equipment (as discussed in section III-A), it is still interesting from the point of view of a numerical study to assess how the performance compares to a masked scheme. Additionally, we also compare to results reported in the literature for some RC schemes, including integrated TDRC [36], [11], [15], [12], integrated spatial-multiplexed RC [37], and bulk [38]. Similar to section III-C, we scanned the parameter space in the length-phase space and considered a fixed high-speed 40 GHz photodetector with a subsequent 12-bit ADC. The best performance on the considered tasks is reported in Table I. For the NARMA$-k$ task, it can be seen that the mask-less performance is close to the masked performance, and is also better than the simulated performance of the RC scheme in [38] except for NARMA-5. Since the test set of the XOR task is 3000 bits, the BER resolution is 3.33×10^{-4}. Therefore, we have also included the NMSE between brackets to observe the performance variations between XOR-1 and XOR-2 which both yield $\mathrm{BER} = 0.0$. It can be seen that the unmasked performance is good up to XOR-3, beyond which the masked scheme outperforms it. However, the mask-less BER is still close to the state of the art performance [37], [12]. For the Santa Fe task, we can observe no significant difference between the optimal NMSE for the masked and mask-less scheme. Compared to state of the art, the mask-less scheme achieves better performance than [11] and performs worse than the simulated scheme in [15].

Furthermore, it can be seen in Table I that our scheme can provide a large enough $\mathrm{MC_{tot}}$ compared to state of the art integrated photonic TDRC schemes, and the optimal $\mathrm{MC_{tot}}$ improves by 1 when using the masked scheme. For the majority of tasks considered here, the performance obtained for the unmasked reservoir is comparable to that of the masked reservoir, and even sometimes slightly outperforms it. However, for tasks which require a larger memory, the mask indeed provides a considerable improvement, e.g. in NARMA-5 and XOR-4,5. When comparing to examples from the literature, it is important to consider the different factors involved. For example, the number of nodes, training and test dataset lengths, regularization parameter(s), and post-processing are almost different in every case, which makes it difficult to have a completely fair comparison. Nevertheless, it is shown that the mask-less scheme achieves results that are within the vicinity of state of the art performance, and comparable to those obtained from applying the mask on the proposed scheme.

TABLE I
PERFORMANCE COMPARISON OF PROPOSED RC VS. EXISTING SCHEMES

Task	**Masked**	**Mask-less**	**State of the Art**
NARMA-2	9.37×10^{-3}	6.69×10^{-3}	5×10^{-2} [38]*
NARMA-3	$1.73{\times}10^{-2}$	3.75×10^{-3}	6×10^{-2} [38]*
NARMA-4	$6.42{\times}10^{-2}$	7.95×10^{-2}	0.12 [38]*
NARMA-5	$9.93{\times}10^{-2}$	0.149	0.14 [38]*
XOR-1	$0.0\ (7.49 \times 10^{-5})$	$0.0\ (1.72 \times 10^{-5})$	10^{-3}[37] $10^{-0.7}$[12]
XOR-2	$0.0\ (7.51 \times 10^{-3})$	$0.0\ (1.26 \times 10^{-2})$	2×10^{-3}[37] $10^{-0.9}$[12]
XOR-3	$0.0\ (4.68 \times 10^{-3})$	6.67×10^{-4} (0.10)	2×10^{-2}[37] $10^{-2.6}$[12]†
XOR-4	$0.0\ (5.39 \times 10^{-3})$	3.20×10^{-2} (0.22)	7×10^{-2}[37] — [12]
XOR-5	$0.0\ (5.19 \times 10^{-3})$	2.30×10^{-2} (0.19)	— [37] — [12]
Santa Fe	0.102	0.108	0.135 [11], 4×10^{-2} [15]
$\mathrm{MC_{tot}}$	10.06	9.06	~1.0 [15]‡, 1.5 [36], 6.0 [11]
Nodes	20	20	124 [36], 16 [37], 35 [38] 23 [15], 23 [11], 5 [12]

For the XOR task BER is reported, NMSE is between brackets.
Simulation results [15], [37], [38], experimental results [11], [12], [36].
— means unreported.
* NMSE calculated from the reported Pearson correlation (r_c) scores as NMSE$= 1 - r_c^2$.
† Post-processing involved.
‡ Ref.[15] reports $\mathrm{MC_{tot}} = 2$ without external feedback, but the authors consider also the term $\mathrm{MC_0}$.

E. Effect of post-filtering

In this section, we consider the effect of post-filtering at the readout. Filtering has been recently investigated [39] as a means of performance improvement for optoelectronic reservoirs. This is of particular interest for passive photonic reservoirs as the readout imprints features on the final output signal which need to be considered. Here, we sweep the photodetector's bandwidth (20 points), ranging from 2 GHz to 40 GHz, while considering the task-independent metrics and the NARMA-k task. The results are not independent of the applied phase shift. However, for visualization purposes we choose one value of phase shift, and plot the results with respect to feedback time. The results on $\mathrm{MC_{tot}}$, as shown in Fig. 5 (a), show an overall increasing trend with increasing bandwidth. In Fig. 5 (b), there is a consistently high CA in the sub-τ region from 5 GHz and above. Thus, for tasks that require less $\mathrm{MC_{tot}}$ and more CA, one could in principle scale down the reservoir while being less constrained by readout electronics. Furthermore, for some longer feedback lengths, it seems that lower bandwidth post-filtering improves CA largely by aiding in lowering (improving) GR, while for some cases increasing (improving) KQR. For the most part, however, KQR increases quasi-monotonically with increasing bandwidth, similar to Fig. 5 (a). Also, it can be seen that smaller τ_{fb} are more tolerant to lower bandwidths in terms of CA. We also consider the NARMA-k benchmark task, and report the behavior under different photodetector bandwidths, as shown in Fig. 6. For NARMA-2 and NARMA-3, excellent performance is obtainable in the sub-τ region, where there is a large degree of tolerance for lower bandwidths. On the other hand, NARMA-4 and NARMA-5 seem to benefit more from the slower dynamics of longer delay lines. We conclude that the output readout affects the dimensionality of the reservoir, mostly reducing the effective number of nodes through a combination of noise and linear filtering, which could essentially drown some of the independent features which become insufficiently distinguishable after passing through the output electronics. Thus, shorter feedback lengths, which provide stronger and more frequent dynamics, seem to be particularly robust to the readout's considered effects.

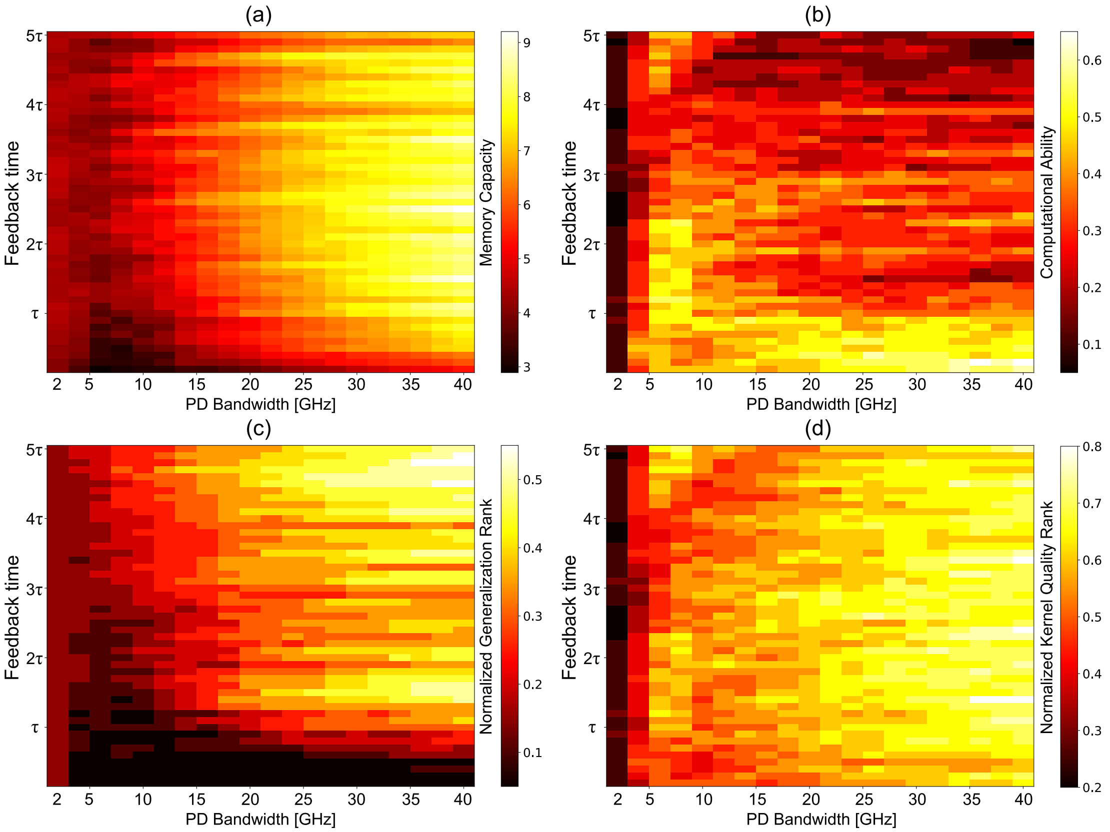


Fig. 5. Performance on the task-independent tests, with respect to photodetector bandwidth and feedback time τ_{fb}.

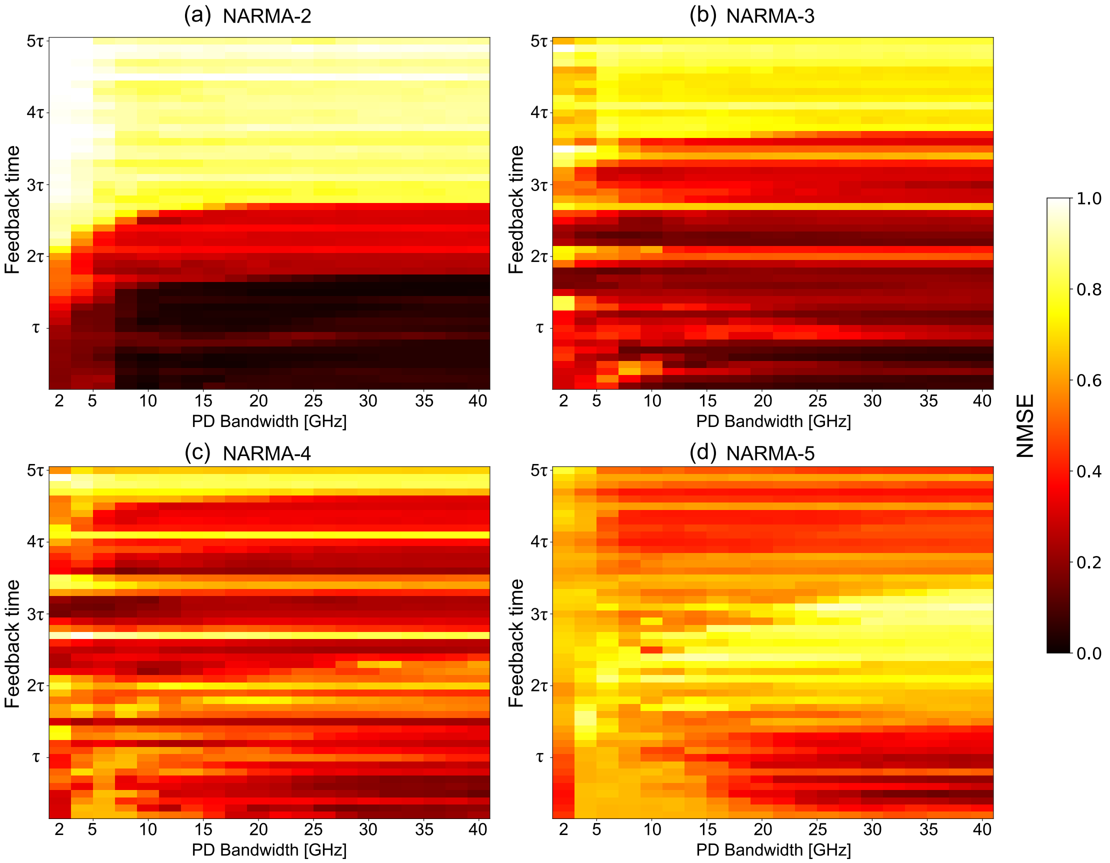
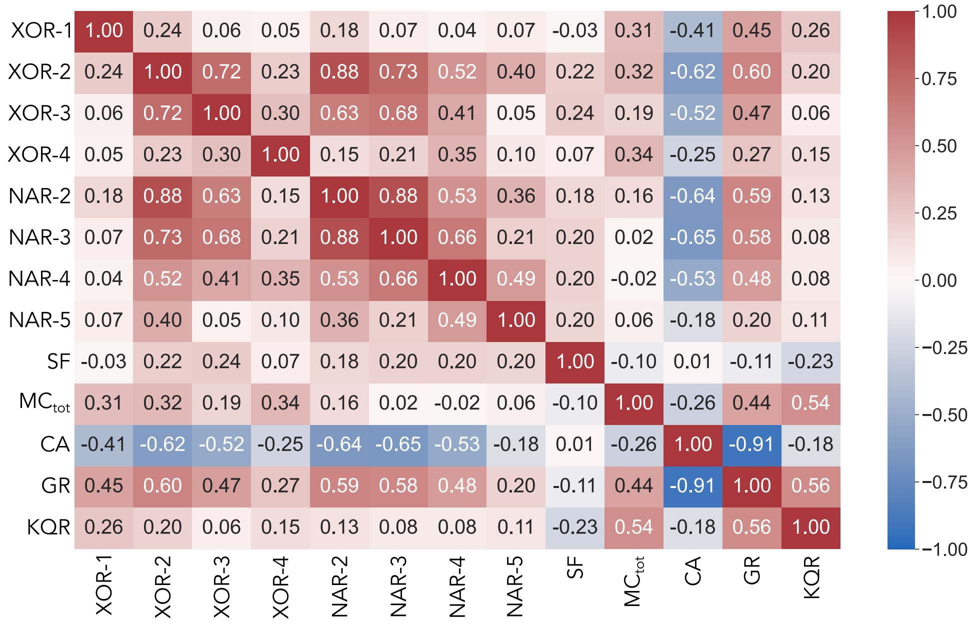


Fig. 6. Performance on NARMA-k benchmark task. For some delay times τ_{fb}, post-filtering assists in lowering the NMSE.



	XOR-1	XOR-2	XOR-3	XOR-4	NAR-2	NAR-3	NAR-4	NAR-5	SF	$\mathrm{MC_{tot}}$	CA	GR	KQR
XOR-1	1.00	0.24	0.06	0.05	0.18	0.07	0.04	0.07	-0.03	0.31	-0.41	0.45	0.26
XOR-2	0.24	1.00	0.72	0.23	0.88	0.73	0.52	0.40	0.22	0.32	-0.62	0.60	0.20
XOR-3	0.06	0.72	1.00	0.30	0.63	0.68	0.41	0.05	0.24	0.19	-0.52	0.47	0.06
XOR-4	0.05	0.23	0.30	1.00	0.15	0.21	0.35	0.10	0.07	0.34	-0.25	0.27	0.15
NAR-2	0.18	0.88	0.63	0.15	1.00	0.88	0.53	0.36	0.18	0.16	-0.64	0.59	0.13
NAR-3	0.07	0.73	0.68	0.21	0.88	1.00	0.66	0.21	0.20	0.02	-0.65	0.58	0.08
NAR-4	0.04	0.52	0.41	0.35	0.53	0.66	1.00	0.49	0.20	-0.02	-0.53	0.48	0.08
NAR-5	0.07	0.40	0.05	0.10	0.36	0.21	0.49	1.00	0.20	0.06	-0.18	0.20	0.11
SF	-0.03	0.22	0.24	0.07	0.18	0.20	0.20	0.20	1.00	-0.10	0.01	-0.11	-0.23
$\mathrm{MC_{tot}}$	0.31	0.32	0.19	0.34	0.16	0.02	-0.02	0.06	-0.10	1.00	-0.26	0.44	0.54
CA	-0.41	-0.62	-0.52	-0.25	-0.64	-0.65	-0.53	-0.18	0.01	-0.26	1.00	-0.91	-0.18
GR	0.45	0.60	0.47	0.27	0.59	0.58	0.48	0.20	-0.11	0.44	-0.91	1.00	0.56
KQR	0.26	0.20	0.06	0.15	0.13	0.08	0.08	0.11	-0.23	0.54	-0.18	0.56	1.00

Fig. 7. Correlation matrix between different tasks, XOR-k, NARMA-k (NAR-k), Santa Fe (SF), and the task independent metrics.

F. Correlation between different tasks

The relation between task-specific performance and task-independent metrics has also been a topic of recent interest [40]. In this section, we shed light on the relation between different tasks through observing the correlation between the reservoir's performance on different tasks over the chosen design space, where we consider the varying feedback length and applied phase shift, and a fixed photodetector bandwidth of 40 GHz. We employ the Pearson correlation coefficient (r_c) to identify linear relations between the performance on various tasks/metrics over the design space. We take into consideration all the presented tasks thus-far, namely the temporal bitwise XOR-k task (with k from 1 to 4 bits into the past), the NARMA-k task (with k from 2 to 5 inputs in the past), the Santa Fe chaotic laser prediction, as well as $\mathrm{MC_{tot}}$, CA, and its constituent metrics (GR, KQR). In this case, for the XOR task, we also employ the NMSE as BER does not yield results falling within a normal distribution. When $0 < r_\mathrm{c} \leq 1$, a positive correlation exists, and when $-1 < r_\mathrm{c} < 0$ it is a negative correlation. A strong positive correlation indicates that two tasks perform similarly well (or similarly poorly) over the design space, while a strong negative correlation would mean that the tasks require opposite configurations. If a strong positive correlation exists between two tasks, we can infer that it is highly likely that the system can be configured in such a way to solve those two tasks at the same time, i.e. the tasks can be solved with the same hyperparameter configuration of the reservoir.

A strong positive correlation is observed between XOR tasks and NARMA tasks for some k. For example, $r_\mathrm{c}(\mathrm{xor2}, \mathrm{narma2}) = 0.88$, which suggests that they can benefit from sharing the same hyperparameter configuration within this design space, even though they are different families of tasks. The same applies to other cases: $r(\mathrm{xor2}, \mathrm{narma3}) = 0.73$, $r_\mathrm{c}(\mathrm{narma2}, \mathrm{narma3}) = 0.88$ and $r_\mathrm{c}(\mathrm{xor2}, \mathrm{xor3}) = 0.72$. These same tasks also exhibit a strong positive correlation with CA (the two quantities CA and NMSE improve in opposite directions). Looking at CA's constituent metrics, it is observed that this is mainly due to the stronger correlation with GR, as the correlation with KQR is generally weak with all benchmark tasks. This is due to the small variations in KQR over the explored design space, which was explained in section III-B, and is the reason why CA itself does not show a strong linear relation with KQR. Similarly, for the Santa Fe task, weak relations are established due to minimal variation of NMSE over the design space. Beyond the sub-τ region, $\mathrm{MC_{tot}}$ does not change considerably, and hence the average or weak correlations with other tasks/metrics. Therefore, for this scheme, CA is a more telling metric of the reservoir's performance on the different tasks than $\mathrm{MC_{tot}}$, especially considering the fact that GR is in itself a test of memory. In fact, the individual memory components MC_k can give a better picture on how much the reservoir can remember for tasks that target specific inputs into the past, such as XOR-k. For other RC schemes, which can have more variations of the nonlinearity, a stronger correlation with KQR over the design space would be expected. In any case, the correlation matrix not only shows inherent relations between different tasks over that space, but is also a useful tool to judge a reservoir's ability to solve multiple desired tasks on the same signal.

IV. CONCLUSION

In this work we have explored the potential for mask-less photonic TDRC using an all-optical, passive, integrated architecture. For applications with an already-incoming optical signal, dispensing with the mask enables bypassing the usual domain conversion(s) at the input layer. In addition to reductions in associated energy and complexity costs, this enables the processing of higher signal bitrates B_m as the mask normally requires $N \times B_\mathrm{m}$ of signal generation/modulation bandwidth, thus restricting either the size of the reservoir or the overall speed of the system. We have explored this within the chosen design space of the minimum complexity architecture, focusing on the feedback length, and the applied phase shift which can be freely tuned post-fabrication. The best obtained results show small differences between masked and mask-less performance for tasks with moderate memory requirements. Additionally, we considered the effects of post-filtering and showed that in the sub-τ region, some tasks can be performed even with considerably strong linear filtering, which degrades the linear independence of the sampled nodes. Furthermore, we have shown the correlation between the reservoir performance on the considered tasks as as a way to understand the relation between the performance on different benchmark tasks and the task-independent metrics within the chosen design space. This would allow further optimization of the reservoir design to handle multiple tasks using the same configuration. This could be useful for other reservoir implementations and serve as a practical tool to aid in design of multitasking RC. Future studies may consider nonlinear effects of the considered platform, along with relevant parameters such as varying the signal-to-noise ratio (SNR), waveguide losses, input wavelength, MZI arm lengths, and coupling ratio of the splitters, which would influence the dynamics of the proposed reservoir and the effective (usable) number of nodes. Moreover, the number of nodes required for certain tasks may be less than what is considered here, which means that they can operate at faster input clock cycles (e.g. tasks showing tolerance to lower bandwidth post-filtering). Finally, other metrics such as IPC may be applied to quantify the reservoir performance and provide insight into the computational power of the intensity-conversion nonlinearity used here and in similar passive photonic RC schemes.

APPENDIX A
SIMULATION PARAMETERS

TABLE II
PARAMETERS OF THE OPTICAL AND ELECTRONIC SIMULATIONS

Parameter	Value	Description
τ_{fb}	scanned over [20 ps;500 ps]	Feedback time
τ	100 ps	Synchronized delay time
L_{fb}	$\tau_{\mathrm{fb}} \times c/n_{\mathrm{g}}$	Feedback length
Φ	scanned over [0;2π]	Applied phase shift
B_{m}	10 GBit/s	Input rate
f_{s}	200 GSa/s	Sampling rate (electronic)
N	$f_{\mathrm{s}}/B_{\mathrm{m}} = 20$	Number of nodes
Δt	1.6 ps	Simulation timestep (optical)
λ_0	1550 nm	Laser wavelength
n_g	1.996	Group index of Si_3N_4
γ_{c}	0.966	$P_{\mathrm{out}}/P_{\mathrm{in}}$ of directional coupler
κ	$\sqrt{0.5}$	Cross coefficient of coupler
r	$\sqrt{0.5}$	Through coefficient of coupler
$A_{(\mathrm{fb})}$	20 dB/m	(Feedback) waveguide loss
P_{in}	10 mW	Input laser power (amplitude)
r_{p}	0.8	Photodetector responsivity
I_{d}	2 nA	Photodetector dark current
R_{l}	100 Ω	Load resistance
f_{c}	scanned over [2 GHz;40 GHz]	Photodetector bandwidth
Q	12 bits	ADC bits of resolution
V_{RefHi}	1.2 V	ADC upper bound
V_{RefLo}	0.0 V	ADC lower bound

APPENDIX B
EFFECT OF ENOB

ADCs in oscilloscopes typically have a lower effective number of bits (ENOB) than their designed resolution, especially when operating at high frequencies. Therefore, it is interesting to see how the proposed RC scheme would perform with such low ENOB. We consider how the scheme performs in the sub-τ region, showing the optimum values obtained for some of the tasks presented earlier in the work under the harsh conditions of a single oscilloscope acquisition, which is equivalent of greatly reducing the considered SNR in this work. Thus, we consider the cases of 5-8 bits of ADC resolution to mimic the effect of a lower ENOB. For XOR-k tasks we report the lowest achieved BER, while for NARMA-k and Santa Fe tasks we report the lowest achieved NMSE.

TABLE III
PERFORMANCE ON LOWER ENOB

Task	5 bits	6 bits	7 bits	8 bits
XOR-1	0.0	0.0	0.0	0.0
XOR-2	8.3×10^{-3}	6.6×10^{-4}	0.0	0.0
NARMA-2	0.13	0.08	0.05	0.04
NARMA-3	0.29	0.21	0.16	0.12
SF	0.46	0.27	0.22	0.17

While this heavily degrades the signal quality, the proposed scheme can still solve the above tasks effectively, albeit with less accuracy. This degradation can be easily mitigated by averaging multiple signal acquisitions, which is usually done in practice and can improve the ENOB by multiple effective bits of precision, in accordance with the number of signals acquired and averaged. An important consideration here would be the signal stability during the total acquisition time of the multiple signals to be averaged. This is done in practice by using a temperature controller to keep the photonic circuit thermally stable during the total acquisition time. However, when considering real-time processing, one limitation is that averaging over multiple acquisitions would decrease the overall throughput of the system. Higher input optical powers would also further improve the SNR and minimize the needed amplification after detection, reducing their associated noise. Under these considerations, the simulation results obtained in this work should approximate the experimental performance.

ACKNOWLEDGMENTS

This work was supported by the ECLAUSion project which has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 801512. We also acknowledge the support of the International Associated Laboratory in Photonics between France and Australia (LIA ALPhFA). This work was also carried out within the framework of the PHASEPUF project supported by the French "Agence Nationale de la Recherche" under project number ANR-20-CE39-0004.

REFERENCES

[1] 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 Networks*, vol. 115, pp. 100–123, 7 2019.

[2] L. Appeltant, M. C. Soriano, G. V. D. Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer, "Information processing using a single dynamical node as complex system," *Nature Communications*, vol. 2, 2011.

[3] A. Rodan and P. Tiňo, "Minimum complexity echo state network," *IEEE Transactions on Neural Networks*, vol. 22, pp. 131–144, 1 2011.

[4] K. Vandoorne, P. Mechet, T. V. Vaerenbergh, M. Fiers, G. Morthier, D. Verstraeten, B. Schrauwen, J. Dambre, and P. Bienstman, "Experimental demonstration of reservoir computing on a silicon photonics chip," *Nature Communications*, vol. 5, 3 2014.

[5] M. Abdalla, C. Zrounba, R. Cardoso, P. Jimenez, G. Ren, A. Boes, A. Mitchell, A. Bosio, O. Ian, and F. Pavanello, "Minimum complexity integrated photonic architecture for delay-based reservoir computing," *Optics Express*, 3 2023.

[6] G. V. D. Sande, D. Brunner, and M. C. Soriano, "Advances in photonic reservoir computing," pp. 561–576, 2017.

[7] Y. Paquot, F. Duport, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman, and S. Massar, "Optoelectronic reservoir computing," *Scientific Reports*, vol. 2, 2012.

[8] Q. Vinckier, F. Duport, A. Smerieri, K. Vandoorne, P. Bienstman, M. Haelterman, and S. Massar, "High-performance photonic reservoir computer based on a coherently driven passive cavity," *Optica*, vol. 2, p. 438, 5 2015.

[9] Y. Yang, P. Zhou, P. Mu, and N. Li, "Time-delayed reservoir computing based on an optically pumped spin vcsel for high-speed processing," *Nonlinear Dynamics*, vol. 107, pp. 2619–2632, 2 2022.

[10] R.-Q. Li, Y.-W. Shen, B.-D. Lin, J. Yu, X. He, and C. Wang, "Scalable wavelength-multiplexing photonic reservoir computing," *APL Machine Learning*, vol. 1, 9 2023. [Online]. Available: https://pubs.aip.org/aml/article/1/3/036105/2902962/Scalable-wavelength-multiplexing-photonic

[11] K. Harkhoe, G. Verschaffelt, A. Katumba, P. Bienstman, and G. V. der Sande, "Demonstrating delay-based reservoir computing using a compact photonic integrated chip," *Optics Express*, vol. 28, p. 3086, 2 2020.

[12] D. Bazzanella, S. Biasi, M. Mancinelli, and L. Pavesi, "A microring as a reservoir computing node: Memory/nonlinear tasks and effect of input non-ideality," *Journal of Lightwave Technology*, vol. 40, pp. 5917–5926, 9 2022.

[13] C. Ma, J. V. Kerrebrouck, H. Deng, S. Sackesyn, E. Gooskens, B. Bai, J. Dambre, and P. Bienstman, "Integrated photonic reservoir computing with an all-optical readout," *Optics Express*, vol. 31, p. 34843, 10 2023.

[14] D. Brunner and I. Fischer, "Reconfigurable semiconductor laser networks based on diffractive coupling," *Opt. Lett.*, vol. 40, no. 16, pp. 3854–3857, Aug 2015. [Online]. Available: https://opg.optica.org/ol/abstract.cfm?URI=ol-40-16-3854

[15] G. Donati, C. R. Mirasso, M. Mancinelli, L. Pavesi, and A. Argyris, "Microring resonators with external optical feedback for time delay reservoir computing," *Optics Express*, vol. 30, p. 522, 1 2022.

[16] F. Stelzer, A. Röhm, K. Lüdge, and S. Yanchuk, "Performance boost of time-delay reservoir computing by non-resonant clock cycle," *Neural Networks*, vol. 124, pp. 158–169, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0893608020300125

[17] T. Hülser, F. Köster, L. Jaurigue, and K. Lüdge, "Role of delay-times in delay-based photonic reservoir computing [invited]," *Optical Materials Express*, vol. 12, p. 1214, 3 2022.

[18] J.-Y. Tang, B.-D. Lin, Y.-W. Shen, R.-Q. Li, J. Yu, X. He, and C. Wang, "Asynchronous photonic time-delay reservoir computing," *Optics Express*, vol. 31, p. 2456, 1 2023.

[19] A. Argyris, J. Schwind, and I. Fischer, "Fast physical repetitive patterns generation for masking in time-delay reservoir computing," *Scientific Reports*, vol. 11, 12 2021.

[20] S. Sackesyn, C. Ma, J. Dambre, and P. Bienstman, "Experimental realization of integrated photonic reservoir computing for nonlinear fiber distortion compensation," *Optics Express*, vol. 29, p. 30991, 9 2021.

[21] J. Dambre, D. Verstraeten, B. Schrauwen, and S. Massar, "Information processing capacity of dynamical systems," *Scientific Reports*, vol. 2, 2012.

[22] B. Vettelschoss, A. Röhm, and M. C. Soriano, "Information processing capacity of a single-node reservoir computer: An experimental evaluation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 6, pp. 2714–2725, 2022.

[23] K. Harkhoe and G. V. der Sande, "Task-independent computational abilities of semiconductor lasers with delayed optical feedback for reservoir computing," *Photonics*, vol. 6, 2019.

[24] S. Tsunegi, T. Kubota, A. Kamimaki, J. Grollier, V. Cros, K. Yakushiji, A. Fukushima, S. Yuasa, H. Kubota, K. Nakajima, and T. Taniguchi, "Information processing capacity of spintronic oscillator," *Advanced Intelligent Systems*, vol. 5, no. 9, p. 2300175, 2023. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/aisy.202300175

[25] 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*, vol. 15, no. 5, pp. 1440–1451, Sep 2023. [Online]. Available: https://doi.org/10.1007/s12559-020-09772-y

[26] H. Jaeger, "Tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the" echo state network" approach," 2002.

[27] A. Röhm, L. Jaurigue, and K. Lüdge, "Reservoir computing using laser networks," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 1, pp. 1–8, 2020.

[28] R. A. Legenstein and W. Maass, "2007 special issue: Edge of chaos and prediction of computational performance for neural circuit models," *Neural Networks*, vol. 20, pp. 323–334, 2007. [Online]. Available: https://api.semanticscholar.org/CorpusID:7249094

[29] Appeltant, *PhD Thesis: Reservoir computing based on delay-dynamical systems*, 2012. [Online]. Available: www.vubpress.be

[30] M. Dale, J. F. Miller, S. Stepney, and M. A. Trefzer, "A substrate-independent framework to characterize reservoir computers," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 475, 6 2019.

[31] A. Weigend and N. Gershenfeld, "Results of the time series prediction competition at the santa fe institute," in *IEEE International Conference on Neural Networks*, 1993, pp. 1786–1793 vol.3.

[32] K. A. Buzaverov, A. S. Baburin, E. V. Sergeev, S. S. Avdeev, E. S. Lotkov, M. Andronik, V. E. Stukalova, D. A. Baklykov, I. V. Dyakonov, N. N. Skryabin, M. Y. Saygin, S. P. Kulik, I. A. Ryzhikov, and I. A. Rodionov, "Low-loss silicon nitride photonic ics for near-infrared wavelength bandwidth," *Opt. Express*, vol. 31, no. 10, pp. 16 227–16 242, May 2023. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-31-10-16227

[33] A. Boes, B. Corcoran, L. Chang, J. Bowers, and A. Mitchell, "Status and potential of lithium niobate on insulator (lnoi) for photonic integrated circuits," *Laser and Photonics Reviews*, vol. 12, 4 2018.

[34] L. Appeltant, G. V. D. Sande, J. Danckaert, and I. Fischer, "Constructing optimized binary masks for reservoir computing with delay systems," *Scientific Reports*, vol. 4, 1 2014.

[35] Y. Kuriki, J. Nakayama, K. Takano, and A. Uchida, "Impact of input mask signals on delay-based photonic reservoir computing with semiconductor lasers," *Optics Express*, vol. 26, p. 5777, 3 2018.

[36] K. Takano, C. Sugano, M. Inubushi, K. Yoshimura, S. Sunada, K. Kanno, and A. Uchida, "Compact reservoir computing with a photonic integrated circuit," *Opt. Express*, vol. 26, no. 22, pp. 29 424–29 439, Oct 2018. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-26-22-29424

[37] A. Katumba, M. Freiberger, P. Bienstman, and J. Dambre, "A multiple-input strategy to efficient integrated photonic reservoir computing," *Cognitive Computation*, vol. 9, no. 3, pp. 307–314, Jun 2017. [Online]. Available: https://doi.org/10.1007/s12559-017-9465-5

[38] J. Henaff, M. Ansquer, M. C. Soriano, R. Zambrini, N. Treps, and V. Parigi, "Optical phase encoding in a pulsed approach to reservoir computing," *Opt. Lett.*, vol. 49, no. 8, pp. 2097–2100, Apr 2024. [Online]. Available: https://opg.optica.org/ol/abstract.cfm?URI=ol-49-8-2097

[39] G. O. Danilenko, A. V. Kovalev, E. A. Viktorov, A. Locquet, D. S. Citrin, and D. Rontani, "Impact of filtering on photonic time-delay reservoir computing," *Chaos*, vol. 33, 1 2023.

[40] T. Hülser, F. Köster, K. Lüdge, and L. Jaurigue, "Deriving task specific performance from the information processing capacity of a reservoir computer," *Nanophotonics*, 3 2022.