Photonic information processing beyond Turing: an optoelectronic implementation of reservoir computing

L. Larger,[1] **M. C. Soriano,**[2] **D. Brunner,**[2,*] **L. Appeltant,**[3] **J. M. Gutierrez,**[4] **L. Pesquera,**[4] **C. R. Mirasso,**[2] **and I. Fischer**[2]

[1]*UMR CNRS FEMTO-ST 6174/Optics Department, University of Franche-Comté, 16 Route de Gray, 25030 Besançon cedex, France*
[2]*Instituto de Física Interdisciplinar y Sistemas Complejos, IFISC (UIB-CSIC), Campus Universitat de les Illes Balear, E-07122 Palma de Mallorca, Spain*
[3] *Applied Physics Research Group (APHY), Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussel, Belgium*
[4]*Instituto de Física de Cantabria, IFCA (UNICAN-CSIC), Av. Los Castros s/n, E-39005 Santander, Spain*

[*]*dbrunner@ifisc.uib-csic.es*

Abstract: Many information processing challenges are difficult to solve with traditional Turing or von Neumann approaches. Implementing unconventional computational methods is therefore essential and optics provides promising opportunities. Here we experimentally demonstrate optical information processing using a nonlinear optoelectronic oscillator subject to delayed feedback. We implement a neuro-inspired concept, called Reservoir Computing, proven to possess universal computational capabilities. We particularly exploit the transient response of a complex dynamical system to an input data stream. We employ spoken digit recognition and time series prediction tasks as benchmarks, achieving competitive processing figures of merit.

OCIS codes: (190.3100) Instabilities and chaos; (200.3050) Information processing; (250.4745) Optical processing devices.

References and links

1. D. A. B. Miller, M. H. Mozolowski, A. Miller, and S. D. Smith, "Nonlinear optical effects in insb with a cw co laser," Opt. Commun. **27**, 133–136 (1978).
2. E. Abraham and S. D. Smith, "Optical bistability and related devices," Rep. Prog. Phys. **45**, 815–885 (1982).
3. J. L. O'Brien, "Optical quantum computing," Science **7**, 1567–1570 (2007).
4. H. J. Caulfield and S. Dolev, "Why future supercomputing requires optics," Nat. Photonics **4**, 261 (2010).
5. R. S. Tucker, "The role of optics in computing," Nat. Photonics **4**, 405 (2010).
6. D. A. B. Miller, "Correspondence to the editor," Nat. Photonics **4**, 406 (2010).
7. H. Jaeger and H. Haas, "Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication," Science **304**, 78–80 (2004).
8. D. V. Buonomano and W. Maass, "State-dependent computations: Spatiotemporal processing in cortical networks," Nat. Rev. Neurosci. **10**, 113–125 (2009).
9. J. P. Crutchfield, L. D. William, and S. Sudeshna, "Introduction to focus issue: Intrinsic and designed computation: Information processing in dynamical systems beyond the digital hegemony," Chaos **20**, 037101 (2010).
10. L. Appeltant, M. C. Soriano, G. Van der 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," Nat. Commun. **2**, 468 (2011).

11. M. Rabinovich, R. Huerta, and G. Laurent, "Transient dynamics of neural processing," Science **321**, 48–50 (2008).
12. K. Vandoorne, W. Dierckx, B. Schrauwen, D. Verstraeten, R. Baets, P. Bienstman, and J. Campenhout, "Toward optical signal processing using photonic reservoir computing," Opt. Express **16**, 11182–11192 (2008).
13. K. Vandoorne, J. Dambre, D. Verstraeten, B. Schrauwen, and P. Bienstman, "Parallel reservoir computing using optical amplifiers," IEEE Trans. Neural Netw. **22**, 1469–1481 (2011).
14. A. Neyer, E. Voges, "Dynamics of electrooptic bistable devices with delayed feedback," IEEE J. Quantum Electron. **18**, 2009–2015 (1982).
15. L. Larger, J.-P. Gordegbuer, and V. S. Udalsov, "Ikeda–based nonlinear delayed dynamics for application to secure optical transmission systems using chaos," C. R. Phys. **5**, 669–681 (2004).
16. K. E. Callan, L. Illing, Z. Gao, D. J. Gauthier, and E. Schöll, "Broadband chaos generated by an optoelectronic oscillator," Phys. Rev. Lett. **104**, 113901 (2010).
17. L. Larger and J. M. Dudley, "Optoelectronic chaos," Nature **465**, 41–42 (2010).
18. K. Ikeda, "Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system," Opt. Commun. **30**, 257–261 (1979).
19. T. Erneux, L. Larger, M. W. Lee, and J. Goedgebuer, "Ikeda hopf bifurcation revisited," Physica D **194**, 49–64 (2004).
20. F. T. Arecchi, G. Giacomelli, A. Lapucci, and R. Meucci, "Two–dimensional representation of a delayed dynamical system," Phys. Rev. A **45**, R4225–R4228 (1993).
21. D. Verstraeten, B. Schrauwen, D. Stroobandt, and J. Van Campenhout, "Isolated word recognition with the liquid state machine: a case study," Inf. Process. Lett. **30**, 521–528 (2005).
22. R. F. Lyon, "A computational model of filtering, detection, and compression in the cochlea," Proc. of the IEEE Int. Conf. Acoust., Speech, Signal Processing (1982).
23. A. E. Hoerl and R. W. Kennard, "Ridge Regression: Applications to Nonorthogonal Problems" Technometrics **12**, 69-82 (1970).
24. A. S. Weigend and N. A. Gershenfeld, "Time series prediction: Forecasting the future and understanding the past," ftp://ftp.santafe.edu/pub/Time-Series/Competition (1993).
25. A. Rodan and P. Tino, "Minimum complexity echo state network," IEEE Trans. Neural Netw. **22**, 131–144 (2011).
26. L. J. Cao, "Support vector machines experts for time series forecasting," Neurocomputing **51**, 321-339 (2003).
27. D. Psaltis, D. Brady, X. G. Gu, and S. Lin, "Holography in artificial neural networks," Nature **343**, 325–330 (1990).
28. Y. Paquot, F. Duport, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman, S. Massar, "Optoelectronic Reservoir Computing," http://arxiv.org/abs/1111.7219

1. Introduction

Optical information processing is a vision originating from the 1970s [1, 2], but due to power consumption, volume and scaling issues, interest decayed in the 1980s. Notwithstanding, optical information processing has been receiving reawakened interest with the evolution of photonic technologies and quantum computing [3]. The potential role of optics in supercomputing is again under consideration [4–6].

Inspired by the way the brain processes information, neuroscience, neural network, and dynamical systems communities have been proposing novel computational concepts [7–9]. These concepts are fundamentally different from the standard Turing or von Neumann Machine methods, which are widely implemented in most computational systems. One of these concepts is known as Echo State Network [7], Liquid State Machine [8] or more generally as Reservoir Computing (RC). RC is based on the computational power of complex recurrent networks operating in a dynamical and transient-like fashion. In standard neural networks recurrent networks have been employed, however resulting in difficulties to train network connection weights. RC benefits from the advantages of recurrent neural networks, while at the same time avoiding the problems in the training procedure. A schematic illustration of the network structure typically considered in RC, is shown in Fig. 1(a). These complex networks (or reservoirs) usually consist of a large number (10^2 to 10^3) of randomly connected nonlinear dynamical nodes receiving the information to be processed via input signals. These input signals are injected from l input channels into m reservoir nodes, with random weights w^i_{lm}. The reservoir response, i.e. the response of the network to the input signal, is evaluated at the read-out nodes j via a linear
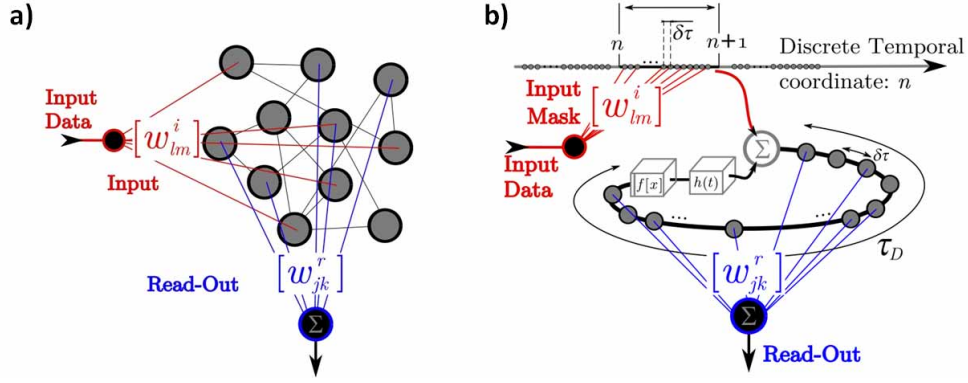
Fig. 1. Schematic representation of RC based on **(a)** a complex network of nonlinear nodes or **(b)** a single nonlinear element subject to delayed feedback via time multiplexing, where $f(x)$ stands for the the system's nonlinear transformation and $h(t)$ denotes the system's impulse response, respectively.

weighted sum of k node states, with coefficients w^r_{jk}. Due to the characteristics of the reservoir and its large number of dynamical elements (degrees of freedom), complex classification tasks and any nonlinear approximation can, in principle, be realized [7, 8, 10].

Without input, the reservoir is typically set to operate in an asymptotically stable, fixed point, state. When excited by an external stimulus (i.e. the information to be processed), the reservoir might, however, exhibit complex transient dynamics. The transient dynamical states, essential for information processing purposes in this scheme, must comply with certain characteristics. If two input signals are similar enough within a certain range, a sufficiently similar transient response must be generated by the reservoir (approximation property). If two input signals belong to different classes, their transient states must sufficiently differ (separation property). These two properties, together with a short-term (fading) memory of the system, are crucial for the computational performance of RC [7, 8]. Similar mechanisms have been reported in real physiological systems [11]. In addition, RC requires the system to be trained with known signals. During this training phase the read-out weights are optimized, enabling subsequent processing of untrained signals belonging to the same class as those used in the training procedure [10].

The experimental implementation of traditional RC brings a key challenge with it. The reservoir is usually composed of a relatively large number of nonlinear nodes interconnected in a network. For instance, a photonic LSM based on a network of coupled Semiconductor Optical Amplifiers (SOA) has recently been proposed and simulated [12, 13]. However, considering the physical complexity of the reservoir, the approach of many nodes is technologically highly demanding and often unrealistic. These constrains can be overcome by replacing the complex network of many elements with an approach based on a single nonlinear element subject to long delayed feedback via time multiplexing [10]. Delay systems are well known to be high dimensional and they have been shown to exhibit a sufficiently large number of different transient states. Despite its simplicity (scalar nonlinear dynamical system, but with a long delay) this system can perform certain tasks as well as traditional reservoirs [10]. A schematic representation of this approach is shown in Fig. 1(b). Here, the complex network is replaced by a reservoir consisting of a single nonlinear element with delayed feedback. The network nodes are distributed along the delay line and the data injection is realized via time multiplexing. From a practical point of view, a big advantage of our scheme is the possible simplification of
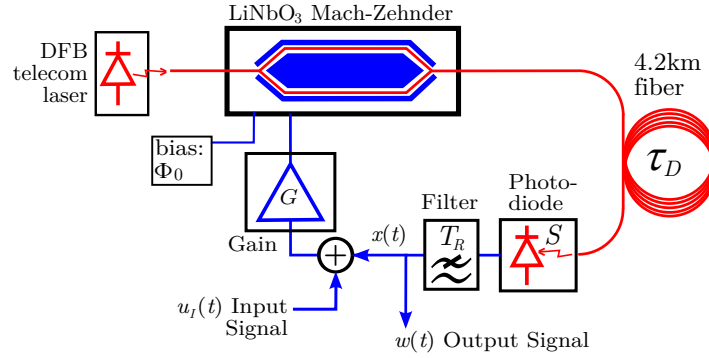
Fig. 2. Optoelectronic implementation of RC with a single nonlinear element subject to delayed feedback.

a hardware implementation.

In the following, we demonstrate the first experimental realization of optical-based RC using a single nonlinear optoelectronic device subject to delay feedback. Our experiments prove that the RC concept can be transfered from the electronic [10] to the optical domain, using optoelectronic hardware. Moreover, by using a different nonlinearity we show that the particular type of the nonlinearity seems not to be crucial. An advantage of the particular choice of nonlinearity in this manuscript is that it allows us to study the dependence of the RC performance on the shape of the nonlinearity in detail. This is achieved by tuning a single experimental parameter. Finally, our experiment demonstrates the potential for a high bandwidth realization of RC.

2. Experimental setup

The scheme we propose is based on a simple and efficient delay-coupled photonic system, depicted in Fig. 2. This setup was originally proposed as a modern integrated optics version allowing for the exploration of optical chaos [14–16], as exhibited by an Ikeda ring cavity [18]; it was also later successfully modified and used in the framework of broadband optical chaos communications [15], and highlighted as a system for studying fundamental characteristics and applications of complex dynamics including RC [17]. Our implementation consists of several key components. We employ a standard telecommunication wavelength DFB diode laser (20 mW) emitting at 1550 nm. An integrated telecom Mach-Zehnder modulator (MZM, LiNbO$_3$) provides an electro-optic nonlinear modulation transfer function ($\sin^2$−function). A long optical fiber implements the delayed feedback loop and a photodiode is employed for optical detection. An electronic feedback circuit closes the nonlinear delay loop, connecting its output to the MZM input electrode. This circuit serves several purposes. It acts as a low pass filter, with a characteristic response time T_R. It allows to add the input information $u_I(t)$ to the delayed signal $x(t)$, and amplifies this signal before it is applied to the MZM to allow for sufficient nonlinear operation. In addition, it provides the data output $w(t)$.

Our experimental system provides direct access to key parameters, e.g. the nonlinearity gain β and the offset phase of the MZM Φ_0, enabling easy tunability of nonlinearity and dynamical behaviors. Parameter β is controlled via the laser diode power, while Φ_0 is controlled by the DC bias input of the MZM. In the absence of input signal, the system is set to operate in a steady (fixed point) state by keeping β at a sufficiently low value. By setting the system in the steady state, a consistent response of the device to the same input signal is guaranteed.

The signal in the feedback loop can be described by the following scalar equation:

$$\varepsilon \dot{x}(s) + x(s) = \beta \, \sin^2[\mu \, x(s-1) + \rho \, u_I(s-1) + \Phi_0], \qquad (1)$$

where ρ is the relative weight of the input information compared to the feedback signal x and μ corresponds to the feedback scaling. Parameter $\varepsilon = T_R/\tau_D$ is the oscillator response time normalized to the delay and $s = t/\tau_D$ is the normalized time. Setting $\rho = 0$, the system performs the well known Ikeda dynamics [18], whose bifurcation diagram has already been intensively explored in the literature [19]. In the RC approach, the dynamics typically remain in a fixed point when it is not excited by an input information ($\beta < 1$). Dynamical complexity occurs during the transient response of the nonlinear delay system when it is excited by the input information.

In delay systems, the dynamical degrees of freedom are distributed along the delay line [20]. Therefore, we define virtual nodes by dividing the total delay interval of length τ_D, realized by 4.2 km optical fiber, into subintervals of length θ [10]. At the end of each subinterval we extract the respective virtual node states. By this, we aim at mimicking the nodes of traditional reservoirs. Unlike traditional RC, connectivity between virtual nodes is limited to local couplings including few nearest neighbors. The extent of the coupling is determined by the characteristic response time (T_R) of the nonlinear delayed feedback loop through its impulse response. The longer (shorter) T_R is relative to the separation θ, the more (less) consecutive virtual nodes are connected. Temporal separations θ slightly smaller than T_R were found to yield the best RC performance [10]. Additional to this short time (local) coupling, a long time coupling originates from the delayed feedback, as explicitly written in Eq.(1).

In order to evaluate the performance of the system, the transient response of the reservoir needs to be processed for a given task. This dedicated processing is carried out by one or several read-out nodes. Each read-out node is defined by a linear weighted sum of the virtual node states. As it is also the case in traditional RC processing, the read-out weights are obtained via a training procedure. This training optimizes the linear separation of the virtual node states, excited by the input information to be processed. A parallel read-out of the virtual nodes can be obtained by simply tapping the delay line at the node positions. Each virtual node is scaled with a weight that needs to be determined from the training stage. In our scheme, a sequential read-out is also possible via time multiplexing, making it more practical and ideally suited for an experimental realization. We have sequentially read out the full transient response of the nonlinear delay dynamics and performed an off-line training procedure using a dedicated toolbox [13].

In our experiments we have chosen a number of $N_N = 400$ virtual nodes [10], a delay time of $\tau_D = 20.87 \, \mu$s, i.e. $\theta = \tau_D/N_N = 52.18$ ns. With the internal system timescale of $T_R = 240$ ns, we calculate a ratio of $T_R/\theta \simeq 4.6$ between the system response time and node width. It is worth mentioning that other values of N_N and τ_D yield similar results, as long as the indicated relative scaling is fulfilled. This is of particular relevance when the proposed setup has to be extended to an ultra-fast version involving standard high speed telecom components.

To evaluate the performance of our system we perform two challenging tasks typically used as benchmark in machine learning and neural network computing: spoken digit recognition and time series prediction. We would like to emphasize at this point that data injection and the classification are in this work computed off line. For RC, the input data is multiplied with a discrete mask, and some additional pre-processing depending on the task at hand. The post processing of the reservoir readout only consists of a linearly weighted sum. As such, both steps could in the future be implemented into the experimental realization with high bandwidth components. The training procedure, which is also carried out offline, once performed, does not affect the bandwidth of the online operation. Accordingly, the achievable bandwidth of an
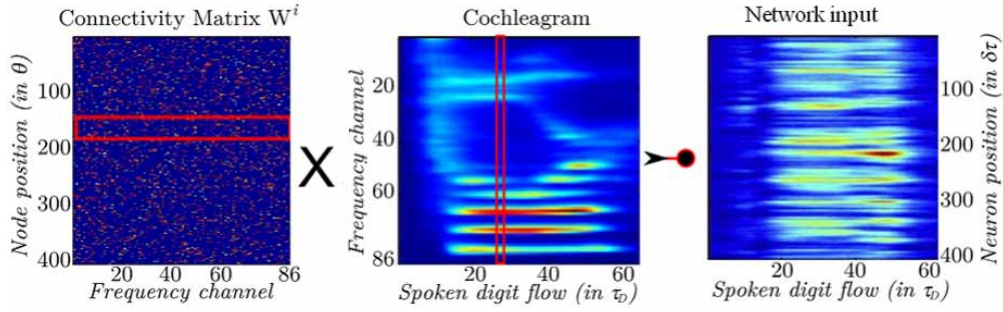
Fig. 3. Injection of a spoken digit into the reservoir showing the input connectivity matrix (left), a Cochleagram of a spoken digit (middle) and the resulting input data of the network (right). In the connect matrix the color code presents the magnitude of the input scaling factors w_{lm}^i, in the Cochleagram and the Network input data the color encodes the amplitudes of the signals, with red (blue) corresponding to large (small) values.

experimental realization consisting of entirely hardware based data injection, reservoir response and classifier readout should be determined by the bandwidth of our reservoir.

3. Benchmark tests for evaluating computational power

Spoken digit recognition is a benchmark test widely used in the field of machine learning and in particular RC [21]. The task of recognizing spoken digits reliably at high speed represents a very demanding computational task. At the same time this test also has a certain appeal due to its practical nature. The standard approach to spoken digit recognition utilizes data pre-processing, which replicates the response of the human Cochlea to sound waves, as depicted in Fig. 3. The Lyon's Cochlear ear model [22] divides the input signal into 86 channels, containing different frequency information, and associating each channel's response to the data input with a firing (excitation) possibility. The input data matrix M_l (dimension N_fxN_s) constructed with the Lyon's Cochlear ear model consists of the corresponding N_f=86 frequency channels and a maximum of N_s=130 samples in time. M_l is multiplied with the input connectivity matrix W_i (dimension N_NxN_f, N_N=400 being the number of virtual nodes in the delay line), creating the data input M_i for the reservoir. Most of the elements w_{lm}^i of the connectivity matrix W_i are set to zero, realizing a sparse and random connectivity between the input layer and the reservoir. The remaining elements are chosen randomly from two discrete mask values, keeping the system in a transient state for the duration of the spoken digit, while also breaking the symmetry between the N_n nodes. The elements of the connectivity matrix remain constant for the duration of the node separation θ. For training the output weights we have randomly chosen 475 spoken digits among a data set of 500, leaving 25 for testing. The read-out weights ω_{jk}^r are calculated from a ridge regression [23] on the system response to the 475 test samples. These weights correspond to the coefficients of a read-out matrix W_r, which is expected to provide the identification of the spoken digit in the form of a so-called target function. The entire training and test procedure is repeated 20 times with different, non-overlapping fragmentations of the 500 speech samples. By following this approach, we minimize the influence of individual speakers and spoken digits on our results, as well as providing statistical information.

 The performance for this task is characterized by the word error rate (WER), as well as a margin. We compute the margin by taking the classifier value of the reservoir's best guess, from which we subtract the classifier value of the second best guess. Figures 4(a) and 4(b) show
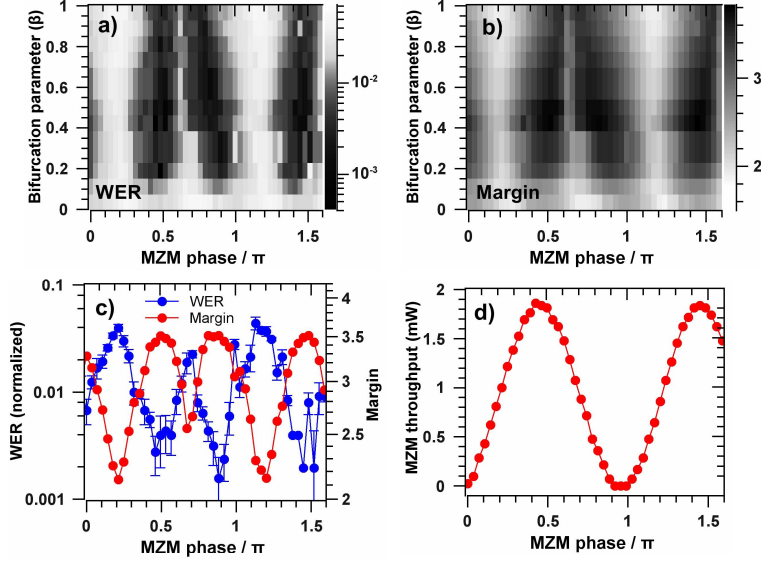
Fig. 4. (a) and (b) show the WER and Margin for spoken digit recognition in the (β, Φ_0)−plane (bifurcation parameter vs. MZM phase). The two figures of merit show a similar dependency on both parameters, with excellent performance at $\beta = 0.3$ and $\Phi_0 = 0.89\pi$. (c) Detailed dependence of the RC performance on the MZM phase at $\beta = 0.3$. (d) MZM transmission function as a function of phase Φ_0.

WER and margin extracted from our experiment, displayed in the (β, Φ_0)−plane. Part (c) of the same Figure provides the Φ_0-dependence for a constant β, while the transmission function of the MZM as a function of Φ_0 is shown in part (d). As demonstrated by the nonlinear transfer function of the MZM, depicted in Fig. 4(d), and by Eq. (1), we can experimentally realize a variety of different nonlinear response properties to data input. These can be directly tuned by scanning the (β, Φ_0)−plane, allowing to control magnitude and sign of the linear, as well as nonlinear response. We can choose to work with settings for different sign and magnitude of slope as well as curvature. Accordingly, our experiment represents not only a powerful electro-optical realization of RC, but at the same time it allows for studying the influence of nonlinearity and dynamical properties on the RC performance. A strong dependence in classification capability of the reservoir is found, with the WER ranging from (7.24 ± 0.79) % down to only (0.04 ± 0.017) %. The systematic dependence of the WER on Φ_0 shows the importance of the nonlinearity for the classification performance. We find the lowest WER always to be at points close, but not equal, to the local extrema of the nonlinear response. Around these points the nonlinearity can be approximated by a quadratic function. The optimal operational point has a tendency to be shifted from the local extrema towards the side with a negative slope in the response function. Corresponding points, sharing the same nonlinearity, differ in stability properties of the fixed point for a change in sign of the slope [19]. Besides operating around the local extrema of the response function, we can tune the operating point to the vicinity of the inflection point, making its response almost linear. Here the performance strongly decreases, highlighting the importance of the nonlinearity for classification tasks. When changing β, we find the optimal operational conditions for intermediate values. As soon as β is sufficiently large ($\beta > 0.1$) the performance does not critically depend on β, as long as Φ_0 is kept optimized. An increase in β, however, results in a growing sensitivity on Φ_0. In the absence of

feedback (μ=0), the system's performance significantly degrades, with the best classification yielding a WER of 1.84 %. Removing the delayed feedback strips the system of its memory, which is thus proven to be beneficial for successful spoken digit classification using our setup. Figure 4(c) shows the WER and margin as a function of Φ_0 for $\beta = 0.3$ and $\rho \simeq \pi$ in more detail. Error bars are extracted from three independent measurements, repeated under identical experimental conditions. It can be seen that good performance is not limited to a single point, with a WER remaining below 0.5% for the range $0.75\pi \leq \Phi_0 \leq 0.95\pi$.

We further evaluated the performance of our system by addressing the one-time-step prediction task of a time series recorded from a far-infrared laser operating in a chaotic state [24]. The one-time-step prediction is performed by feeding the reservoir only one explicit data point at a time. Information about points further in the past are present in the system only implicitly due to its internal, fading memory. To evaluate the performance of our RC approach we computed the normalized mean square error (NMSE) between a sequence of predicted points and their corresponding targets. The results for the one-time-step prediction are depicted in Fig. 5. For $\beta = 0.2$ (blue points), we again find a strong dependence of the NMSE on the MZM phase Φ_0 and therefore on the characteristics of the nonlinearity. For $\Phi_0 = 0.1\pi$ we obtain the lowest prediction error with a NMSE$= 0.124 \pm 4 \times 10^{-4}$. For the task of time series prediction the system's performance is optimized for Φ_0 being shifted further away from the local extrema in the response function, closer towards the inflection point. In addition, the system's performance significantly degrades for these values of Φ_0 corresponding to the local extrema. This is different to the behavior obtained in the spoken digit recognition task, where at these values of Φ_0 the performance was not optimal, still the loss in performance was far less significant. We interpret this as a manifestation of the importance of the memory for the one-time-step prediction task, however, a small amount of nonlinearity is still required for obtaining good performance. To provide evidence that the performance indeed stems from the interplay of high-dimensional mapping and nonlinearity and not from the nonlinearity alone, we in addition plot the data obtained when disconnecting the feedback line (red points, $\mu = 0$). The lower performance without feedback loop (i.e. memory) is clearly visible. Data presented for $\beta = 0.2$ shows consistently better optimal performance for $\Phi_0 < 0.5\pi$, where the slope of Eq. (1) is positive. For the case of zero feedback the performance is almost symmetric around $\Phi_0 = 0.5\pi$, again indicating that this effect might be connected to properties of the system's memory. Timeseries prediction based on numerical methods achieved even lower prediction errors (below 1 % using echo state networks [25] or support vector machines [26]), however neglecting noise and finite experimental precision, and even more, externally feeding the reservoir several data points at a time.

4. Conclusion

Our results prove that a simple nonlinear optoelectronic system subject to delayed feedback can efficiently perform RC, a non-Turing type of computation. The presented experiments encourage a new approach to optical information processing, representing a flexible and efficient, potentially low power-consuming device with excellent computational performance. Using RC, parallel and high speed optical processing becomes feasible without the difficulty of training the entire connection topology of the network [27], which is an advantage over classical optical neural networks. Laser diodes and other nonlinear optical elements with dynamical bandwidths easily reaching 10 GHz should allow for an all-optical implementation of the reservoir. An evaluation of speed limitations due to all-optical data input and data classification requires, however, more detailed studies. Our approach serves multipurpose information processing, as demonstrated by the two different computational tasks carried out in the experiments. We note that a related experiment is reported in [28]. Our demonstrated results should not be limited to an optoelectronic oscillator and might be transferred to all-optical
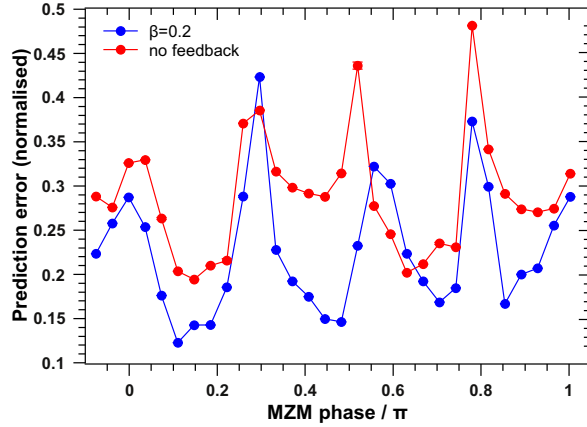
Fig. 5. MZM phase dependence of the RC performance in a time series prediction task, using the Santa-Fe data set. Best performance for $\beta = 0.2$ is found around $\Phi_0 = 0.1\pi, \Phi_0 = 0.5\pi, \Phi_0 = 0.7\pi$ and $\Phi_0 = 0.85\pi$ phase values in the vicinity of local extrema of the transfer function of the MZM (see Figs. 4(d), 1(a), and 1(b)).

implementations. This would allow for direct interconnection between optical communication and information processing.

Major work needs to be done in the future in order to explore the full potential of our approach, including scaling possibilities. In addition, implementation of more advanced features, e.g. enhancing the connectivity of the virtual network, real-time post-processing and plasticity rules to optimize the reservoir for the corresponding task during the training phase, are foreseen.

Acknowledgments