Classification of IQ-Modulated Signals Based on Reservoir Computing With Narrowband Optoelectronic Oscillators

Haoying Dai and Yanne K. Chembo, *Senior Member, IEEE*

***Abstract*— We numerically perform the classification of IQ-modulated radiofrequency signals using reservoir computing based on narrowband optoelectronic oscillators (OEOs) driven by a continuous-wave semiconductor laser. In general, the OEOs used for reservoir computing are wideband and are processing analog signals in the baseband. However, their hardware architecture is inherently inadequate to directly process radiotelecom or radar signals, which are modulated carriers. On the other hand, the high-Q OEOs that have been developed for ultra-low phase noise microwave generation have the adequate hardware architecture to process such multi-GHz modulated signals, but they have never been investigated as possible reservoir computing platforms. In this article, we show that these high-Q OEOs are indeed suitable for reservoir computing with modulated carriers. Our dataset (DeepSig RadioML) is composed with 11 analog and digital formats of IQ-modulated radio signals (BPSK, QAM64, WBFM, etc.), and the task of the high-Q OEO reservoir computer is to recognize and classify them. Our numerical simulations show that with a simpler architecture, a smaller training set, fewer nodes and fewer layers than their neural network counterparts, high-Q OEO-based reservoir computers perform this classification task with an accuracy better than the state-of-the-art, for a wide range of parameters. We also investigate in detail the effects of reducing the size of the training sets on the classification performance.**

***Index Terms*— Reservoir computing, optoelectronic oscillators, nonlinear oscillators, IQ modulation formats, radio modulation recognition.**

I. Introduction

MACHINE learning (ML) has already enacted a deep impact on modern science and technology. Unlike Turing-von Neumann machines that require input data and known rules to provide outputs, ML platforms use input data and known outputs to infer rules or underlying structure.

The impact of machine learning is expected to be transformational in wireless communications networks, which are submitted to a (genuinely) exponential increase of traffic, en route to reach a zettabyte per year within 3 years. Therefore, the radiofrequency (RF) spectrum is becoming increasingly congested with signals sharing the same frequency band. Processing this amount of information is an unprecedented and difficult challenge, which could be in part addressed by machine learning [1].

Effective adoption of ML for wireless networks is nevertheless facing several challenges. The first one is that it is quite common for ML approaches to be solely based on software, that is, the input signal needs to be converted to a computer-compatible format before ML-software processing. Along the same line, in most instances, the machine learning platform requires transduction or conversion at both the input and output stages. A second challenge is that in the spectral domain, wireless signals are generally narrowly centered around a carrier frequency – up to 70 GHz for the next generation of wireless networks. An optimal ML platform should be able to process the modulated signal, without demodulation to the baseband. A third challenge is that high throughput wireless networks need to be interconnected with optical fiber networks both at the backhaul and fronthaul level, since optical fiber is the optimal technological solution to ensure high-bandwidth connection between the network nodes. Consequently, from the hardware perspective, ML solutions should ideally have the capability to handle both microwave and lightwave signals, in order to interact seamlessly with both types of signals in these heterogeneous wireless-fibered networks.

Reservoir computing (RC) based on optoelectronic oscillators is a machine learning approach that can address the three aforementioned challenges. OEOs have a hybrid architecture that concatenates a laser-driven optical fiber line and an electronic branch, and they have been thoroughly investigated from both the fundamental and applied points for view (see review article [2] for a survey of OEO science and technology).

A few years ago, it was shown that wideband OEOs have the capability to perform machine learning based on reservoir computing (or RC) with RF and/or optical input signals [3], [4]. Subsequently, wideband OEOs have been proven to be a robust RC-compatible platform by numerous studies, achieving state-of-the-art performances with several benchmarks such as spoken digit recognition, nonlinear time-series prediction, packet header recognition, signal radar forecasting, and even cancer data classification (see for example [3]–[14]).

Manuscript received January 20, 2021; revised April 6, 2021; accepted April 12, 2021. Date of publication April 20, 2021; date of current version April 28, 2021. This work was supported in part by the University of Maryland through the Minta Martin Fellowship and in part by the Northrop Grumman–University of Maryland Seed Grant Program. *(Corresponding author: Yanne K. Chembo.)*

The authors are with the Department of Electrical and Computer Engineering, University of Maryland at College Park, College Park, MD 20742 USA, and also with the Institute for Research in Electronics and Applied Physics (IREAP), University of Maryland at College Park, College Park, MD 20742 USA (e-mail: dhy@terpmail.umd.edu; ykchembo@umd.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JQE.2021.3074132.

Digital Object Identifier 10.1109/JQE.2021.3074132

This analog computing has also investigated for all-optical systems as well (see for example [15]–[30]). This computational power is partially rooted in the fact that RC only features 3 layers (input, reservoir and output, as explained in detail in [31]–[33]), and requires to train only the output layer while the reservoir (or network) layer is fixed: this distinctive property results in a drastic reduction of processing latency, ideal for large throughput.

An important property of OEOs for wireless networks is their versatility, in the sense that they can process input information either in the microwave domain (RF signals) or lightwave domain (optical fiber data, etc.). Noteworthy is the fact that OEOs can operate virtually at any frequency up to 100 GHz, and have the intrinsic capability handle high-bandwidth signals and perform ultrafast processing, such as in the experiment reported in [9] where OEO-based RC permitted to achieve a record of one million words per second.

In this article, we will extend the concept of OEO-based RC for the classification of RF modulated signals. Wideband OEOs are not suitable for this task, as they essentially operate in the baseband. However, narrowband OEOs, initially developed for the purpose of ultralow phase noise microwave generation, are capable of processing RF modulated signals despite the fact that they have never been considered for the purpose of RC or ML in general. Our main objective here is to therefore to show, for the first time to the best of our knowledge, that narrowband OEOs can be used for RC in wireless communications. Another objective is to show that even with a reduced training set size, the reservoir computer can still provide a high level of performance – sometimes better than the state-of-the-art – with a faster computing speed in comparison to other ML alternatives. We consider a benchmark where an RF signal is IQ-modulated using 11 different formats, and with different levels of signal-to-noise ratio. The task of the RC computer will be to classify these incoming signals with the highest accuracy but smaller training set sizes.

The article is organized as follows. Section II is devoted to the description of the narrowband OEO that will be used for RC, while the dynamical equations ruling its interaction with the incoming IQ-modulated RF signal are derived in Sec. III. The dataset of IQ-modulated signals is described in Sec. IV, and Sec. V describes in detail the RC procedure implemented by the narrowband OEO for this classification task. Our numerical simulations are presented in Sec. VI, and their relevance for the field of ML in wireless communications is discussed. The last section recapitulates our main results and concludes the article.

II. The System

The system under study is a single-loop OEO with an input port in the RF branch for the signal to be processed. A schematic representation of this oscillator is displayed in Fig. 1, and it consists of the following elements in a closed feedback loop: (i) A wideband integrated Mach-Zehnder modulator, seeded by a continuous-wave (CW) semiconductor laser of optical power P; the modulator is characterized by a DC

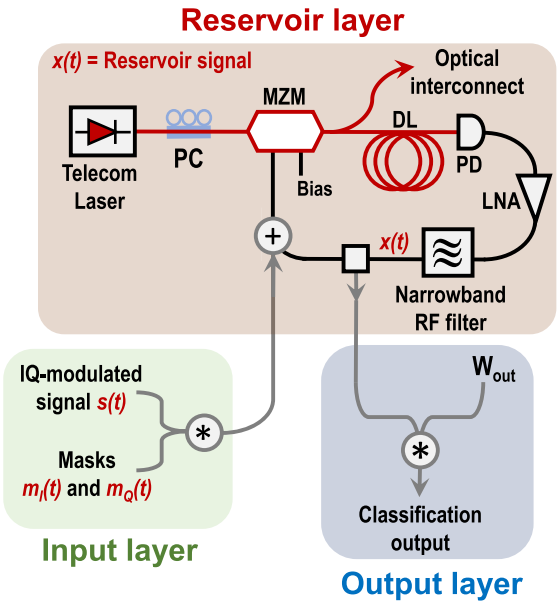

Fig. 1. Schematical representation of the reservoir computer based on a narrowband OEO. DL: Delay line; LNA: Low-noise amplifier. MZM: Mach-Zehnder modulator; PD: Photodiode.

half-wave voltage $V_{\pi_{DC}}$ and an RF half-wave voltage $V_{\pi_{RF}}$, which are both of the order of few volts. (ii) The modulated laser beam at the output of the modulator is launched into an optical fiber delay line of length $L = 4$ km and group velocity index n_g, that creates a time delay of $T = n_g L/c = 20$ μs, where c is the velocity of light in vacuum; This delay line also induces microwave ring-cavity modes with a free-spectral range (FSR) equal to $\Omega_T/2\pi = 1/T = 50$ kHz. Note that for our purpose, a physical optical fiber delay line is not an absolute necessity; the time delay can be implemented using a digital signal processing (DSP) of field programmable gate array (FPGA) board. (iii) The time-delayed optical signal is converted into an electrical signal by a photodiode with a conversion factor S, in units of V/W. (iv) The electrical signal is then sent to narrowband microwave RF filter, that has the role to select the carrier frequency $\Omega_0/2\pi$ and the allowed bandwidth $\Delta\Omega/2\pi$ around it. This central frequency can be anywere in the 1-100 GHz range, while the bandwidth is typically few tens of MHz wide. In all cases, the RF quality factor $Q_{\rm RF} = \Omega_0/\Delta\Omega$ of the narrowband RF filter is generally in excess of 100. (v) The filtered electrical is subsequently amplified with a gain factor G, before being connected back as a voltage $V(t)$ to the RF electrode of the Mach-Zehnder modulator. All optical and electrical losses are accounted for with a single attenuation factor κ. (vi) The IQ-modulated RF signal to be processed is eventually inserted in the feedback loop via a microwave coupler, placed just before the RF electrode of the modulator.

It is important to emphasize here the essential differences between narrowband and wideband OEOs. As discussed in the review article [2], both systems are Ikeda-like oscillators characterized by the very same four elements (linear gain, filter, delay and nonlinearity) and feature similar bandwidth-delay products ($\Delta\Omega \times T \sim 100$), but the difference in parameters $\Delta\Omega$ and T leads to substantially different dynamical properties.

Indeed, wideband OEOs feature short time delay and a large bandwidth filter, while conversely, narrowband OEO feature a long time delay and a narrow bandwidth filter. The wideband OEOs generally output spectrally flat and broad signals, including hyperchaotic ones. As a consequence, it is intuitively logical to use them to process baseband electrical signals through RC. However, narrowband OEOs were intended to generate a single tone RF signals, so that it is less intuitive to foresee how they can be useful for RC. Nevertheless, as it will be shown in this article, RF modulated signals have a spectral spread that falls within these narrowband OEO bandwidths, and can therefore be processed.

III. The Model

A. The Narrowband OEO Model

The narrowband OEO obeys an integro-differential equation that can be written under the simplified and dimensionless form

$$x + \tau \frac{dx}{dt} + \frac{1}{\theta} \int_{t_0}^{t} x(\xi) d\xi = \beta \cos^2[x_T + \phi], \tag{1}$$

where $x(t) = \pi V(t)/2V_{\pi_{RF}}$ is the dimensionless voltage at the input of the Mach-Zehnder modulator, $x_T \equiv x(t - T)$ is its time-delayed counterpart, $\beta = \pi \kappa GSP/2V_{\pi_{RF}}$ is the normalized feedback gain, $\phi = \pi V_B/2V_{\pi_{DC}}$ is the Mach-Zehnder off-set phase. In wideband OEOs, the parameters $\tau = 1/2\pi f_{\mathrm{H}}$ and $\theta = 1/2\pi f_{\mathrm{L}}$ are the typical timescales corresponding to the low- and high-frequency cutoffs f_{L} and f_{H} of two cascaded first-order filters. However, in narrowband OEOs, they are written instead as $\tau = 1/\Delta\Omega$ and $\theta = \Delta\Omega/\Omega_0^2$ and jointly describe the characteristic timescales associated with the resonant bandpass filter.

However, Eq. (1) is not suitable for the study of the narrowband OEO because of the very large splitting between the various timescales [35]. This problem is circumvented by acknowledging that the narrowband filter forces the oscillations to be quasi-sinusoidal with frequency Ω_0, so that their complex envelope carries all the relevant dynamical information. Therefore, we can write the solution of Eq. (1) as signal with a carrier of frequency Ω_0, which is slowly modulated by a complex-valued amplitude $\mathcal{A}(t)$ following

$$x(t) = \frac{1}{2}\mathcal{A}(t)e^{i\Omega_0 t} + \frac{1}{2}\mathcal{A}^*(t)e^{-i\Omega_0 t} \tag{2}$$

where the envelope $\mathcal{A}(t) = A(t)e^{i\psi(t)}$ varies slowly relatively to the carrier of frequency Ω_0 – a condition that mathematically translates to $|\dot{\mathcal{A}}(t)| \ll \Omega_0|\mathcal{A}(t)|$.

It was shown that the envelope obeys the delay-differential equation [35]–[37]

$$\dot{\mathcal{A}} = -\mu\mathcal{A} - 2\mu\gamma\, e^{-i\sigma}\, \mathrm{Jc}_1[2|\mathcal{A}_T|]\mathcal{A}_T, \tag{3}$$

where $\mu = \Delta\Omega/2$ is the half-bandwidth of the radio-frequency filter, $\gamma = \beta \sin\, 2\phi$ is the effective gain of the feedback loop, $\sigma = \Omega_0\, T$ is the round-trip phase shift of the microwave, Jc_1 is the Bessel-cardinal function defined as $\mathrm{Jc}_1(x) = \mathrm{J}_1(x)/x$, and $\mathcal{A}_T \equiv \mathcal{A}(t - T)$. The phase condition of the oscillation is $\gamma\, e^{-i\sigma} = -1$, and it is convenient to set $\sigma = \pi$ (modulo 2π), so that $\gamma \geq 0$.

Equation (3) typically has four dynamical regimes as a function of γ [36]. For $\gamma < 1$, the system does not oscillate and the trivial fixed point is stable ($\mathcal{A} = 0$). At $\gamma = 1$, the system undergoes a primary Hopf bifurcation and the system oscillates with a constant complex-valued envelope, i.e. $x(t)$ is a sinusoidal signal corresponding to a stable limit-cycle (this regime is the one that is useful for ultrapure microwave generation). As the gain is increased and reaches the critical value $\gamma_{\mathrm{cr}} = \{2\mathrm{Jc}_1[\mathrm{J}_0^{-1}(0)]\}^{-1} \simeq 2.31$, the oscillator undergoes a secondary Hopf bifurcation, also known as a Neimark-Sacker bifurcation, where the amplitude of the sinusoidal signal becomes periodically modulated $x(t)$–in other words, envelope undergoes a primary Hopf bifucation and the limit-cycle bifurcates to a torus. This (multi-)periodic regime bifurcates to chaos when the gain is increased beyond a value around 2.8.

IV. The OEO Model With IQ-Modulated Signal Injection

A. IQ-Modulated Signal Dataset

The data set we have chosen to evaluate the performance of narrowband OEOs is made of radiofrequency signals with IQ modulation ("I" for *in-phase*, and "Q" for *quadrature*). The waveforms $I(t)$ and $Q(t)$ are baseband signals, and the IQ-modulation scheme encodes them onto a radio carrier whose frequency can be as high as few hundred GHz. This modulation scheme is popular in modern architectures of wireless networks because it allows to transmit high-capacity modulation formats (constellations with more bits per symbol), and therefore permits to optimize the bandwidth of the transmission link.

As the airspace becomes crowded by many RF signals sharing the same frequency band allocation, it becomes critical to have the capacity to identify, and the large throughput of data favors hardware analog recognition in order to avoid latency and network congestion. Typical tests involve for example RF fingerprinting hardware imperfections or emitted modulation formats. In this article, we are considering the latter case, which is particularly important in many instances such as for dynamic spectrum access protocols, which require the optimization of spectrum allocation. The efficiency of this optimization procedure greatly relies on the efficient radio recognition, and while prior works are mostly based on feature engineering from temporal signals, analog computing techniques are promising for directly processing these signals in the physical layer.

From the perspective of machine learning, the radio modulation recognition task could be characterized as a classification problem that requires correct prediction of the modulation scheme directly from the radio signals. In order to train and test our proposed narrowband OEO-based RC, we will exploit the IQ radiofrequency modulation dataset proposed by O'Shea, Corgan and Clancy (Deepsig dataset RADIOML 2016.04C–see [34], [38]). This dataset consists of 11 modulation signals generated by GNU radio, as displayed in Fig. 3. Eight of these modulations are digital modulations, namely BPSK (binary phase-shift keying), BFSK (binary frequency-shift keying), QPSK (quadrature phase-shift keying), CPFSK

(continuous phase-frequency-shift Keying), PAM4 (4-level pulse-amplitude modulation), 8PSK (8 phase-shift keying), 16QAM (16 quadrature amplitude modulation), and 64QAM (64 quadrature amplitude modulation), while the remaining three are analog modulations schemes, namely WB-FM (wide-band frequency modulation), AM-SSB (single-sideband amplitude modulation), and AM-DSB (double-sideband amplitude modulation).

Each example in the dataset contains 128 sampled points for both in-phase and quadrature components. For each type of modulation, signals are collected with different signal-to-noise ratio (SNR) ranging from -20 dB to 18 dB with a step of 2 dB. Hence, the in-phase and quadrature components of one example could be expressed by two vectors of length 128 written as $\{X_I^i(1), X_I^i(2), \ldots, X_I^i(128)\}$ and $\{X_Q^i(1), X_Q^i(2), \ldots, X_Q^i(128)\}$, where subscripts I and Q indicates the signal is in-phase or quadrature component and the superscript indicates that it corresponds to the i^{th} example in the dataset.

B. Dynamical Model for the Driven OEO

We need to derive an envelope equation for the narrowband OEO when it is driven by an IQ-modulated signal. The modulated IQ signal could be expressed as:

$$s(t) = I(t)\cos\Omega_0 t + Q(t)\sin\Omega_0 t \tag{4}$$

and according to the OEO architecture presented in Fig. 1, this IQ-modulated signal influences the dynamics of the OEO following

$$x + \tau\frac{dx}{dt} + \frac{1}{\theta}\int_{t_0}^{t} x(\xi)d\xi = \beta\cos^2[x_T + \rho s(t) + \phi], \tag{5}$$

where ρ is a dimensionless normalization factor weighting the amplitude of the signal $s(t)$ in the feedback loop. In order to determine the envelope equation corresponding to Eq. (5), we first rewrite the modulated terms inside the nonlinear function $\cos^2$ as

$$x_T + \rho s(t) = q(t)\cos\Omega_0 t + r(t)\sin\Omega_0 t \tag{6}$$

where

$$q(t) = \rho I(t) + A_T\cos(\psi_T - \sigma) \tag{7}$$

$$r(t) = \rho Q(t) - A_T\sin(\psi_T - \sigma) \tag{8}$$

with

$$A_T \equiv |\mathcal{A}(t-T)| \text{ and } \psi_T \equiv \arg\mathcal{A}(t-T). \tag{9}$$

We can now rewrite the signal injected into RF electrode of the Mach-Zehnder modulator as

$$x_T + \rho s(t) = \frac{1}{2}\mathcal{Z}(t)e^{i\Omega_0 t} + \frac{1}{2}\mathcal{Z}^*(t)e^{-i\Omega_0 t} \tag{10}$$

with $\mathcal{Z}(t)$ being its complex-valued envelope explicitly defined through

$$|\mathcal{Z}(t)| = \sqrt{q^2(t) + r^2(t)} \tag{11}$$

$$\arg\mathcal{Z}(t) = \arctan\frac{q(t)}{r(t)} - \frac{\pi}{2}. \tag{12}$$

Performing the same algebra as the one that led to Eq. (3), we are led to the envelope equation

$$\dot{\mathcal{A}} = -\mu\mathcal{A} - 2\mu\gamma\,\mathrm{Jc}_1[2|\mathcal{Z}|]\mathcal{Z}, \tag{13}$$

which governs the dynamics of the narrowband OEO envelope when driven by an IQ modulated signal.

V. Reservoir Computing Procedure

A. Conventional Reservoir Computer

Reservoir computing is a particular subset of supervised machine learning algorithm that arose as a one of the leading paradigms in the area of neuromorphic computing. This approach, which is conceptually close to the ones of echo state network (ESN) or a liquid-state machine (LSM), is a brain-inspired framework that can perform efficient and real-time processing for temporal information [31]–[33], [39]). Unlike the usual neural networks where the full connectivity matrix is updated via a backpropagation algorithm (for example), reservoir computing is based on the optimization of "read-out" coefficients only. The reservoir computer therefore encodes the information signal to be processed into the highly dimensional dynamical system.

There are two main advantages justifying our choice of reservoir computing for our project: *(i)* Only the output layer needs to be trained, while the reservoir remains static. This feature greatly simplifies the determination of the optimal weights. Once the optimal read-out coefficients are obtained, they are used to project the nonlinear transient dynamics of the reservoir onto an identifiable target state. *(ii)* The second advantage is that this simpler architecture permits faster training and testing than in conventional neural networks. Since, high-throughput inherently imposes a stringent constraint on computation time, the possibility provided by reservoir computing to expedite the training appears as a critical advantage for our purpose.

The signals at the input layer are fed into the RC, such that the reservoir and output states $\mathbf{x}$ and $\mathbf{y}$ are sequentially updated at discrete time-steps $n \in \mathbb{N}$ following

$$\mathbf{x}(n+1) = \mathbf{f}_{\text{NL}}[\mathbf{W}_{\text{in}}\cdot\mathbf{u}(n) + \mathbf{W}\cdot\mathbf{x}(n)] \tag{14}$$

$$\mathbf{y}(n+1) = \mathbf{W}_{\text{out}}\cdot\mathbf{x}(n+1), \tag{15}$$

where $\mathbf{u}(n)$ is the input signal while $\mathbf{f}_{\text{NL}}$ is a nonlinear vector function. The optimal readout connectivity matrix $\mathbf{W}_{\text{out}}$ is obtained via ridge regression through

$$\mathbf{W}_{\text{out}}^{\text{opt}} = \mathbf{M}_{\mathbf{y}}\cdot\mathbf{M}_{\mathbf{x}}^{\mathsf{T}}\cdot[\mathbf{M}_{\mathbf{x}}\cdot\mathbf{M}_{\mathbf{x}}^{\mathsf{T}} + \lambda\cdot\mathbf{I}]^{-1}, \tag{16}$$

where $\mathbf{M}_{\mathbf{x}}$ is a matrix that suitably concatenates the internal state $\mathbf{x}$ generated with some specific training input vectors $\mathbf{u}$, while $\mathbf{M}_{\mathbf{y}}$ is the target (or *teaching*) matrix yielding the desired classification outcome, $\mathbf{I}$ is the identity matrix and $\lambda \ll 1$ is a small regularization factor for to avoid the overfitting of the inversion problem. The prediction task becomes a regression problem if the teaching matrix is high-dimensional (quasi-continuous approximation), and a classification problem when it is low-dimensional. Despite its simplicity, RC provides a powerful analog computing algorithm with performances

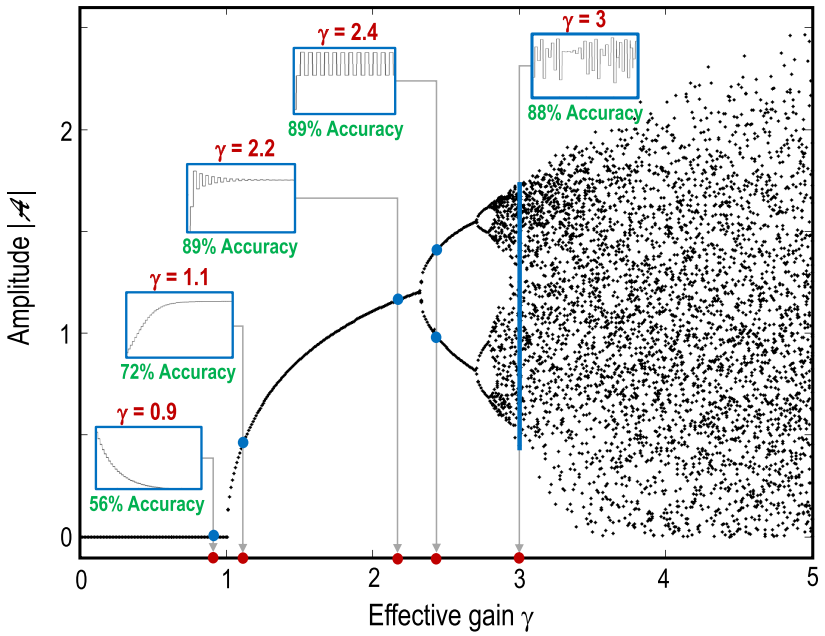

Fig. 2. Bifurcation diagram of the narrowband OEO as the effective gain γ is increased. The system undergoes a Hopf bifurcation at $\gamma = 1$, and a Neimark-Sacker bifurcation at $\gamma = 2.31$. The timetraces in blue inserts display the transient dynamics of the non-driven OEO to the final state for 5 gains values of interest. The best classification accuracy achieved with a training set of $230 \times 11 = 2530$ samples and SNR = 18 dB is indicated as well for these 5 gain values (ρ is scanned from 1.5 to 2.5). The highest classification accuracy in our RC simulations (88.94% for $\gamma = 2.2$ and $\rho = 1.63$) is obtained with only 1 trained layer: It outperforms the state-of-the-art accuracy of 87% which is achieved using neural networks with 5 trained layers and ~ 4800 (i.e twice as many) training samples [34]. Interestingly, our work shows that optimal performance is achieved around the Neimark-Sacker bifurcation, and *not* around the primary Hopf bifurcation like in RC using broadband OEOs.

comparable to or even better than RNNs on several benchmarks [40]. Interested readers could refer to [41] for a more general and detailed review.

B. Narrowband OEO-Based RC Implementation

In principle, a time-delay system has an infinite dimensionality that can be exploited to map input data into arbitrary high dimensional space. In OEO-based RC, this property is used to excite a large number of virtual neurons in the temporal domain as traditional RC does in spatial domain. The first implementation of RC with time-delayed systems was implemented using an electronic Mackey-Glass oscillator [42]. We refer the reader to that reference, as well as to [3], [4], [11], [14] for an in-depth presentation of RC using time-delayed systems.

In order to implement RC with our narrowband OEO, we need first to determine the functions $r(t)$ and $q(t)$ introduced in Sec. IV-B, so that we can compute the complex-valued envelope $\mathcal{Z}(t)$ needed to simulate Eq. (13). Before being fed into reservoir layer, the in-phase and quadrature components are processed into two token sequence as shown in Section IV. From these two discrete token sequence, another two continuous time series are generated by holding every sampled data point of in-phase and quadrature component $X_I^i(k)$ and $X_Q^i(k)$ for system delay time T. In this way, the discrete sample points of one example are processed into two continuous time series which satisfies:

$$I(t) \equiv X_I^i(k), \quad (k-1)T \le t \le kT \tag{17}$$

$$Q(t) \equiv X_Q^i(k), \quad (k-1)T \le t \le kT. \tag{18}$$

While conventional RC couples all data points of one example into RC layer in parallel, the narrow band optoelectronic RC exploits the concept of time-division multiplexing by feeding only one data point into RC layer during one delay time T. In each delay time T, a data point is coupled to the virtual nodes after multiplication with the input mask. Thus, the input mask should be a signal produced temporally with period equals to the delay time T so that these weights are fixed for all input points. In our simulations, we use six-valued masks, following the approach proposed in [43]. To stimulate N virtual nodes in the RC layer, the weight value of mask must be separated by time interval $\Delta T = T/N$, with $N = 400$ in our case. As there are two input streams corresponding to in-phase and quadrature components, we introduce two masks $m_I(t)$ and $m_Q(t)$ to couple them into RC layer following:

$$J_I(t) = I(t) \cdot m_I(t) \text{ and } J_Q(t) = Q(t) \cdot m_Q(t). \tag{19}$$

The values of $\mathcal{A}(t)$ at the end of each time interval of ΔT is sampled to give the reservoir states of each virtual nodes. Thus, the output of RC layer value corresponding to j^{th} virtual node of i^{th} example of input data is:

$$\mathsf{A}_{ij} = \mathcal{A}[i \cdot T + (N-j) \cdot \Delta T]. \tag{20}$$

The RC states of all data points in the whole training set are collected and vectorized to train the weights of read-out layer, which lead to an optimization problem to search for optimal read-out layer weights $\mathbf{W}^{\text{RO}}$:

$$\mathbf{W}_{\text{opt}}^{\text{RO}} = \underset{\mathbf{W}^{\text{RO}}}{\text{argmin}} \sum_{i=1}^{N_{\text{Train}}} \|\mathbf{W}^{\text{RO}}\mathbf{A}_{i\cdot} - \mathbf{Y}_i\|^2 + \lambda\|\mathbf{W}^{\text{RO}}\|^2 \tag{21}$$

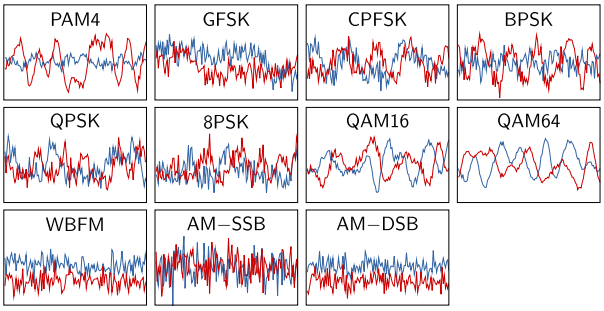

Fig. 3. Examples of normalized timetraces of the I (blue) and Q (red) signals for our 11 modulation formats, when the SNR = 0 dB. Each timetrace has 128 sample points.

where $\mathbf{A}_{i\cdot}$ is the RC states of i^{th} training example, and $\mathbf{Y}_i$ is the corresponding label. The optimal read-out layer weights $\mathbf{W}_{\text{opt}}^{\text{RO}}$ could be obtained via ridge regression algorithm [40]:

$$\mathbf{W}_{\text{opt}}^{\text{RO}} = \mathbf{A}(\mathbf{A}^T\mathbf{A} + \lambda\mathbf{I})^{-1}\mathbf{Y} \tag{22}$$

where $\mathbf{A}$ is a suitably designed matrix that concatenates all internal RC states A_{ij}, $\mathbf{Y}$ stands for the labels of training set and λ is a small regulation term to prevent overfitting.

VI. Simulation Results and Discussion

We consider the following parameters for our numerical simulations: The delay time is $T = 20$ μs, the narrowband filter bandwidth is $\Delta\Omega/2\pi = 20$ MHz, the roundtrip phase shift is $\sigma = \Omega_0\ T = \pi$, the insertion gain for the input data is $\rho = 5$, and the regularization term for ridge regression algorithm is $\lambda = 0.001$.

Here we use the dataset mentioned in Section IV-A as benchmarks to test our narrow band OEO-based RC. Unless otherwise stated, we extract 60 training samples and 60 test samples for every type of modulation and for each SNR, which leads to a training set and a test set with equal size of 660. Before multiplying with input mask, we normalized these signals first so that each training example's power sums to one, following:

$$X_{I,Q}^{i\prime}[j] = \sum_{k=1}^{128} \frac{X_{I,Q}^{i}[j]}{\sqrt{X_I^i[k]^2 + X_Q^i[k]^2}} \tag{23}$$

As a classification task, the performance metrics is the classification accuracy, which is defined as the percentage of correctly predicted modulation samples with regard to all predicted samples. We also introduce the confusion matrix as another pertinent indicator that helps understand where are the sources of error in our RC classifier.

As explained in Fig. 2, our best accuracy performance is 88.94% (with a training set size of 2530 samples and $\rho = 1.62$), which is better than the state-of-the-art at 87% (with a training set size of 4800 samples). The OEO-based RC is already demonstrating its potential with this performance, but in this work, one of our key objectives is to investigate how the RC performs with even *smaller* training set sizes.

Our results are presented in Table I, which displays the accuracy of our classifier as a function of the gain γ and the signal-to-noise ratio, when the training set size is reduced to 660 samples. As shown in Fig. 2, the parameter γ is directly related to the dynamical regime of narrowband OEO. We can see from the table, the RC performs poorest when $\gamma = 0.9$, which corresponds to the oscillator being in the trivial equilibrium when not driven by the IQ-modulated signal [$x(t) \equiv 0$ and $\mathcal{A}(t) \equiv 0$]. This result seems to be in contradiction with the conventional wisdom gained from RC with broadband OEOs, where normalized gain values below unity provide optimal performance. In the case of narrowband OEOs, the dynamical regime where RC is the most efficient is located around the secondary bifurcation, i.e., the one where the envelope (and not the carrier) undergoes a Hopf bifurcation. This situation actually makes sense when one realizes that the I and Q signals only contribute to the envelope dynamics, and not to the one of the carrier: as a consequence, only the Hopf instability related to the envelope might contribute to the RC performance. This is exactly what is observed as the value of the gain is increased: The RC efficiency is ameliorated when γ is in the area surrounding the Neimark-Sacker bifurcation at $\gamma_{\text{cr}} = 2.31$. It is indeed unconventional to operate a reservoir computer beyond a fixed point but previous work had already shown that this approach is valid under certain conditions for RNNs [44]. To the best of our knowledge, RC has never been operated in a chaotic regime, but in our case, the performance at $\gamma = 3$ remains comparable to the maximal efficiency of our classifier. This is explained by the fact that chaos in for this gain parameter remains low-dimensional, in many points similar to the laminar chaos recently discussed in the literature [45]. However, our simulations show that if the gain is increased any further, the system becomes more chaotic and the performance steadily decreases: For an SNR = 0 dB, the accuracy drops to $\simeq 50\%$ for $\gamma = 3.5$, and to $\sim 40\%$ for $\gamma = 4.2$.

TABLE I

Classification Accuracy as a Function of Gain and SNR (with $\rho = 1.61$ and 660 Training Samples)

SNR	−6 dB	0 dB	6 dB	8 dB	10 dB	18 dB
$\gamma = 0.9$	19%	44%	63%	65%	50%	32%
$\gamma = 1.1$	19%	46%	67%	67%	60%	59%
$\gamma = 1.6$	22%	47%	71%	73%	68%	79%
$\gamma = 2.2$	22%	47%	76%	77%	69%	82%
$\gamma = 2.4$	22%	47%	76%	77%	69%	83%
$\gamma = 3.0$	22%	48%	73%	78%	70%	75%

As shown in Fig. 4, we have also investigated the effect of SNR on the RC performance, for the fixed gain value $\gamma = 2.2$, but different training size sets sizes. We can first note that as expected, the RC efficiency is poor when the SNR is low (typically, SNR < -10 dB). In that case, the reservoir computer basically operates as a random classifier, allocating the 11 labels following a uniform probability distribution with accuracy $1/11 \simeq 9.09\%$. As the noise is reduced, the reservoir computer progressively begins to outperform the random classifier, and it is noteworthy that even when the noise is 4 times more powerful than the signal (SNR $= -6$ dB), RC can still distinguish features and perform much better than a random guess. Higher SNR values lead to a significant

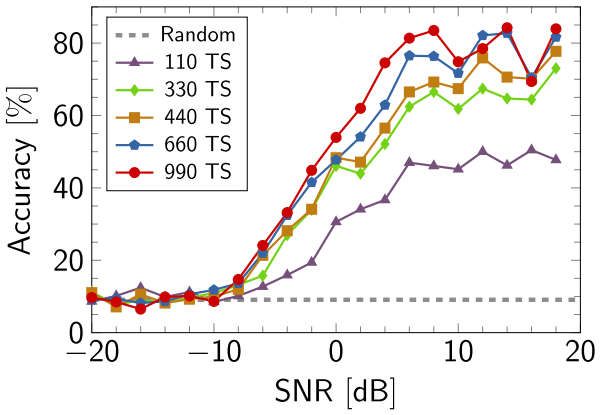

Fig. 4. Variation of the classification accuracy of the narrowband reservoir computer with the SNR, for $\gamma = 2.2$. We have considered various set sizes for the training samples (TS), ranging from ~ 100 to ~ 1000 (note that the typical set size for training in neuronal-network-based ML is ~ 5000 – see [34]). The purple line with triangle symbols corresponds to the accuracy when the training set has 110 patterns (10 for each of the 11 modulation formats); The green line with diamond symbols corresponds to the accuracy when the training set involves 220 patterns (20 for each of the 11 modulation formats); etc. The dashed line indicates the performance of a random classifier ($1/11 \simeq 9.09\%$) For very low SNR (i.e., very high levels of noise), the reservoir computer does not performs better than a random classifier. However, as the SNR is increased, the reservoir computer accuracy significantly improves and achieves a performance approaching 90%.

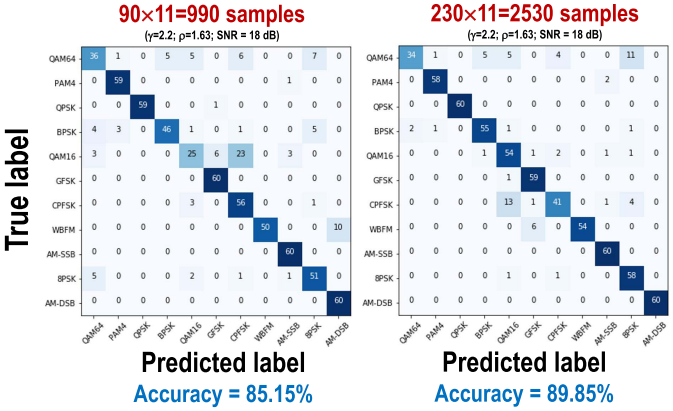

Fig. 5. Confusion matrices for $\gamma = 2.2$, and for two different training sample sizes. The testing set size is contains 660 items, so that the sum for each line must is always 60 (the number of testing samples). One can observe that most errors mainly come from a confusion between WBFM and AM-DSB analog formats, and from a confusion between the digital formats QAM64, QAM16, and the PSK (8, Q, and B).

increase of performance, but quite counter-intuitively, higher SNR does not systematically lead to better accuracy. These anomalies (also present in Table I) are be explained by the variability of the dataset.

The confusion matrices of Fig. 5 give us an insight in relation to the origin of misclassifications. We have reported such matrices for two sets of parameters yielding good performance. It appears that there two confusion clusters generate most of the errors. The first one involves the analog formats WBFM and AM-DSB, and it can be seen that when the reservoir computer classifies one of them with high accuracy, it performs very poorly for the other. This is consistent with the results from [34], where the classifications is performed with neural networks. The second confusion cluster involves the 5 digital formats QAM64, QAM16, 8PSK, QPSK and BPSK, which are sometimes shuffled by the classifier.

VII. Conclusion

In this article, we have investigated the performance of high-Q OEO-based RC for the classification of IQ-modulated signals. We have shown that they have an performance better than the state-of-the-art, while featuring a simpler architecture, which leads to a faster computing capability. We have also explored the effect of training set size reduction, and shown using about ten times less training samples that the state-of-the-art neural networks leads to only a marginal reduction of accuracy for our RC. This robustness property gives to OEO-based RC an edge for real-time computing or ML with nonstationary datasets, and positions them as promising ML systems for the processing of radio signals. Our results also indicate that since OEOs can process both lightwave and microwave signals, they are suitable as real-time analog computing nodes in the physical layer of the next generation of networks, where wireless and fiber communications links are interconnected.

References

[1] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2224–2287, 3rd Quart., 2019.

[2] Y. K. Chembo, D. Brunner, M. Jacquot, and L. Larger, "Optoelectronic oscillators with time-delayed feedback," *Rev. Mod. Phys.*, vol. 91, no. 3, Sep. 2019, Art. no. 035006.

[3] L. Larger *et al.*, "Photonic information processing beyond Turing: An optoelectronic implementation of reservoir computing," *Opt. Exp.*, vol. 20, no. 3, pp. 3241–3249, 2012.

[4] Y. Paquot *et al.*, "Optoelectronic reservoir computing," *Sci. Rep.*, vol. 2, no. 1, p. 287, Feb. 2012.

[5] S. Ortín *et al.*, "A unified framework for reservoir computing and extreme learning machines based on a single time-delayed neuron," *Sci. Rep.*, vol. 5, no. 1, p. 14945, Oct. 2015.

[6] R. Martinenghi, S. Rybalko, M. Jacquot, Y. K. Chembo, and L. Larger, "Photonic nonlinear transient computing with multiple-delay wavelength dynamics," *Phys. Rev. Lett.*, vol. 108, no. 24, Jun. 2012, Art. no. 244101.

[7] M. Hermans, P. Antonik, M. Haelterman, and S. Massar, "Embodiment of learning in electro-optical signal processors," *Phys. Rev. Lett.*, vol. 117, no. 12, Sep. 2016, Art. no. 128301.

[8] G. Van der Sande, D. Brunner, and M. C. Soriano, "Advances in photonic reservoir computing," *Nanophotonics*, vol. 6, no. 3, pp. 561–576, May 2017.

[9] 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*, vol. 7, no. 1, Feb. 2017, Art. no. 011015.

[10] P. Antonik, M. Haelterman, and S. Massar, "Brain-inspired photonic signal processor for generating periodic patterns and emulating chaotic systems," *Phys. Rev. A, Gen. Phys.*, vol. 7, no. 5, May 2017, Art. no. 054014.

[11] D. Brunner, B. Penkovsky, B. A. Marquez, M. Jacquot, I. Fischer, and L. Larger, "Tutorial: Photonic neural networks in delay systems," *J. Appl. Phys.*, vol. 124, no. 15, Oct. 2018, Art. no. 152004.

[12] J. D. Hart, L. Larger, T. E. Murphy, and R. Roy, "Delayed dynamical systems: Networks, chimeras and reservoir computing," *Phil. Trans. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 377, no. 2153, Sep. 2019, Art. no. 20180123.

[13] 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," *Nature Commun.*, vol. 10, no. 1, p. 3538, Dec. 2019.

[14] Y. K. Chembo, "Machine learning based on reservoir computing with time-delayed optoelectronic and photonic systems," *Chaos, Interdiscipl. J. Nonlinear Sci.*, vol. 30, no. 1, Jan. 2020, Art. no. 013111.

[15] K. Hicke, M. A. Escalona-Morán, D. Brunner, M. C. Soriano, I. Fischer, and C. R. Mirasso, "Information processing using transient dynamics of semiconductor lasers subject to delayed feedback," *IEEE J. Sel. Topics Quantum Electron.*, vol. 19, no. 4, Jul. 2013, Art. no. 1501610.
[16] D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, "Parallel photonic information processing at gigabyte per second data rates using transient states," *Nature Commun.*, vol. 4, no. 1, p. 1364, Jan. 2013.
[17] K. Vandoorne *et al.*, "Experimental demonstration of reservoir computing on a silicon photonics chip," *Nature Commun.*, vol. 5, no. 1, p. 3541, Mar. 2014.
[18] C. Mesaritakis, A. Bogris, A. Kapsalis, and D. Syvridis, "High-speed all-optical pattern recognition of dispersive Fourier images through a photonic reservoir computing subsystem," *Opt. Lett.*, vol. 40, no. 14, pp. 3416–3419, Jul. 2015.
[19] Q. Vinckier *et al.*, "High-performance photonic reservoir computer based on a coherently driven passive cavity," *Optica*, vol. 2, pp. 438–446, May 2015.
[20] J. Bueno, D. Brunner, M. C. Soriano, and L. Fischer, "Conditions for reservoir computing performance using semiconductor lasers with delayed optical feedback," *Opt. Exp.*, vol. 25, no. 3, pp. 2401–2412, Feb. 2017.
[21] R. M. Nguimdo, E. Lacot, O. Jacquin, Q. Hugon, G. Van der Sande, and H. G. de Chatellus, "Prediction performance of reservoir computing systems based on a diode-pumped erbium-doped microchip laser subject to optical feedback," *Opt. Lett.*, vol. 42, no. 3, pp. 375–378, Feb. 2017.
[22] F. D.-L. Coarer *et al.*, "All-optical reservoir computing on a photonic chip using silicon-based ring resonators," *IEEE J. Sel. Top. Quantum Electron.*, vol. 24, no. 6, Nov./Dec. 2018, Art. no. 7600108.
[23] Y. Hou *et al.*, "Prediction performance of reservoir computing system based on a semiconductor laser subject to double optical feedback and optical injection," *Opt. Exp.*, vol. 26, no. 8, pp. 10211–10219, Apr. 2018.
[24] A. Katumba *et al.*, "Neuromorphic computing based on silicon photonics and reservoir computing," *IEEE J. Sel. Topics Quantum Electron.*, vol. 24, no. 6, pp. 1–10, Nov. 2018.
[25] J. Vatin, D. Rontani, and M. Sciamanna, "Experimental reservoir computing using VCSEL polarization dynamics," *Opt. Exp.*, vol. 27, no. 13, pp. 18579–18584, Jun. 2019.
[26] D. Yue *et al.*, "Performance optimization research of reservoir computing system based on an optical feedback semiconductor laser under electrical information injection," *Opt. Exp.*, vol. 27, no. 14, pp. 19931–19939, Jul. 2019.
[27] Y. Chen *et al.*, "Reservoir computing system with double optoelectronic feedback loops," *Opt. Exp.*, vol. 27, no. 20, pp. 27431–27440, Sep. 2019.
[28] X. X. Guo, S. Y. Xiang, Y. H. Zhang, L. Lin, A. J. Wen, and Y. Hao, "Polarization multiplexing reservoir computing based on a VCSEL with polarized optical feedback," *IEEE J. Sel. Topics Quantum Electron.*, vol. 26, no. 1, pp. 1–9, Jan. 2020.
[29] A. Argyris *et al.*, "Comparison of photonic reservoir computing systems for fiber transmission equalization," *IEEE J. Sel. Topics Quantum Electron.*, vol. 26, no. 1, pp. 1–9, Jan. 2020.
[30] C. Sugano, K. Kanno, and A. Uchida, "Reservoir computing using multiple lasers with feedback on a photonic integrated circuit," *IEEE J. Sel. Topics Quantum Electron.*, vol. 26, no. 1, pp. 1–9, Jan. 2020.
[31] W. Maass, T. Natschläger, and H. Markram, "Real-time computing without stable states: A new framework for neural computation based on perturbations," *Neural Comput.*, vol. 14, no. 11, pp. 2531–2560, Nov. 2002.
[32] H. Jaeger, "Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication," *Science*, vol. 304, no. 5667, pp. 78–80, Apr. 2004.
[33] D. Verstraeten, B. Schrauwen, M. D'Haene, and D. Stroobandt, "An experimental unification of reservoir computing methods," *Neural Netw.*, vol. 20, no. 3, pp. 391–403, Apr. 2007.
[34] T. J. O'Shea, J. Corgan, and T. C. Clancy, "Convolutional radio modulation recognition networks," in *Proc. Int. Conf. Eng. Appl. Neural Netw.* Springer, 2016, pp. 213–226.
[35] Y. K. Chembo, L. Larger, H. Tavernier, R. Bendoula, E. Rubiola, and P. Colet, "Dynamic instabilities of microwaves generated with optoelectronic oscillators," *Opt. Lett.*, vol. 32, no. 17, pp. 2571–2573, 2007.
[36] Y. K. Chembo, L. Larger, and P. Colet, "Nonlinear dynamics and spectral stability of optoelectronic microwave oscillators," *IEEE J. Quantum Electron.*, vol. 44, no. 9, pp. 858–866, Sep. 2008.
[37] Y. K. Chembo, L. Larger, R. Bendoula, and P. Colet, "Effects of gain and bandwidth on the multimode behavior of optoelectronic microwave oscillators," *Opt. Exp.*, vol. 16, pp. 9067–9072, Jun. 2008.
[38] T. J. O'Shea, J. Corgan, and T. C. Clancy, "Unsupervised representation learning of structured radio communication signals," in *Proc. 1st Int. Workshop Sens., Process. Learn. Intell. Mach. (SPLINE)*, Jul. 2016, pp. 1–5.
[39] G. Tanaka *et al.*, "Recent advances in physical reservoir computing: A review," *Neural Netw.*, vol. 115, pp. 100–123, Jul. 2019.
[40] M. Lukoševičius and H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Comput. Sci. Rev.*, vol. 3, no. 3, pp. 127–149, Aug. 2009.
[41] M. Lukoševičius, H. Jaeger, and B. Schrauwen, "Reservoir computing trends," *KI Künstliche Intelligenz*, vol. 26, no. 4, pp. 365–371, Nov. 2012.
[42] L. Appeltant *et al.*, "Information processing using a single dynamical node as complex system," *Nature Commun.*, vol. 2, no. 1, p. 468, Sep. 2011.
[43] M. C. Soriano *et al.*, "Optoelectronic reservoir computing: Tackling noise-induced performance degradation," *Opt. Exp.*, vol. 21, no. 1, pp. 12–20, Jan. 2013.
[44] B. A. Marquez, L. Larger, M. Jacquot, Y. K. Chembo, and D. Brunner, "Dynamical complexity and computation in recurrent neural networks beyond their fixed point," *Sci. Rep.*, vol. 8, no. 1, p. 3319, Dec. 2018.
[45] J. D. Hart, R. Roy, D. Müller-Bender, A. Otto, and G. Radons, "Laminar chaos in experiments: Nonlinear systems with time-varying delays and noise," *Phys. Rev. Lett.*, vol. 123, no. 15, Oct. 2019, Art. no. 154101.

Haoying Dai received the bachelor's degree in electronics from Southeast University, Nanjing, China, in 2015. He is currently pursuing the Ph.D. degree with the University of Maryland. His research interests include machine learning based on microwave photonic systems for wireless and optical fiber communication networks.

Yanne K. Chembo (Senior Member, IEEE) received the Ph.D. degree in nonlinear dynamics from the University of Yaoundé I, Cameroon, in 2005, and the Ph.D. degree in photonics from the University of the Balearic Islands, Palma de Mallorca, Spain, in 2006. In 2007 and 2008, he was a Post-Doctoral Fellow with the Centre National de la Recherche Scientifique (CNRS), France, and the FEMTO-ST Institute, Besançon, France. In 2009, he was a NASA Postdoctoral Program (NPP) Fellow with the Jet Propulsion Laboratory, Quantum Science and Technology Group, Caltech, Pasadena, USA. From 2010 to 2016, he was the Research Scientist of CNRS, with affiliation to the FEMTO-ST Institute, where he founded and led the Microwave Photonics Group. In 2017, he joined the International GeorgiaTech-CNRS Research Laboratory, Atlanta, USA, where he became the CNRS Research Director in 2018. In 2019, he joined the Institute for Research in Electronics and Applied Physics (IREAP), University of Maryland, where he has been the Head of the Photonic Systems Laboratory. He has coauthored more than 100 articles in refereed international journals and more than 60 international refereed proceedings. His research interests include microwave photonics, machine learning, as well as the applied nonlinear, stochastic, and quantum dynamics of complex photonic systems. He is a fellow of SPIE and OSA. He is an Associate Editor of the *Optics Express* journal (OSA).