

Dynamic Response Prediction of Bi-State Emission of Quantum Dot Lasers Based on Machine Learning

Wild Freitas da Silva Santos (✉ wild.santos@ufba.br)

Federal University of Bahia: Universidade Federal da Bahia

Eduardo Furtado Simas Filho

Federal University of Bahia: Universidade Federal da Bahia

George André Pereira Thé

Federal University of Ceara: Universidade Federal do Ceara

Research Article

Keywords: quantum dot laser, optical filters, extreme learning machine, echo state network

Posted Date: December 2nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1037376/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Dynamic Response Prediction of Bi-State Emission of Quantum Dot Lasers Based on Machine Learning

Wild F. S. Santos^{1*}, Eduardo F. Simas Filho¹ and George A. P. Thé²

^{1*}Digital System Laboratory (PPGEE), Federal University of Bahia (UFBA), R. Prof. Aristίδes Novis, Salvador, 40210-630, Bahia, Brazil.

²Dept. of Teleinformatics Engineering (DETI), Federal University of Ceará (UFC), Campus do Pici s/n, Bloco 725, 60440-900, Fortaleza, Ceará, Brazil.

*Corresponding author(s). E-mail(s): wild.santos@ufba.br;
Contributing authors: eduardo.simas@ufba.br; george.the@ufc.br;

Abstract

Dual-state emission is a common and important phenomenon which takes place in semiconductor Quantum Dot Lasers at different temperature and operating conditions usually investigated from microscopic carrier interaction modeling or even rate-equations based approaches. In this study, we revisit the topic, but the investigation is here performed from a system identification perspective; we built black-box models based on artificial neural networks approach, using the Multilayer Perceptron, the Extreme Learning Machine and a hybrid Echo State Network - Extreme Learning Machine. As a case study, we focused on switch-on transient and its prediction. The study revealed the model was able to separate and to predict, from the solely total power, without using any QDL design parameters, the optical power around the ground state and first excited state lasing lines of InAs/InGaAs quantum dot laser. The error performance was low as a RMSE of 2.81 μ W and MAPE of 0.50% with processing time (training and testing time) of 15.27 s, enabling the alternative model to be used in optical filtering instrumentation as low-resolution and low-cost filters for applications in which it is not economically viable to use a spectrum analyzer, which can be replaced by a simple optical power meter.

Keywords: quantum dot laser, optical filters, extreme learning machine, echo state network

1 Introduction

Semiconductor Quantum Dot Laser (QDL) is an important class of semiconductor emitting devices which presents distinct properties such as sharp optical transitions, low threshold current, high output power, efficiency and large modulation bandwidth and has applications in several fields of medicine, telecommunications and optical instrumentation [1].

Most of its interest comes from the rich carrier dynamics in the active region and in the surrounding layers, which gives rise to the Dual-State Emission (DSE) at certain operating conditions. The DSE consists of the optical power emission in two spectral regions: one around the Ground State (GS) and the other around the first Excited State (ES) lasing lines [2]; if well exploited, it can be useful for applications requiring wide emission spectrum [3] or even for wavelength switching applications.

Due to its characteristics and applicability, the study of the DSE remains relevant, being analyzed with several different focuses, such as analyzing the optical noise that may impact some applications [2], or the evaluation of its properties for applications with external optical feedback [4].

In this work, we revisit the DSE phenomenon in time-domain, but from the perspective of nonlinear regression based on artificial neural networks (ANN), in this case, the Multilayer Perceptron (MLP) [5], the Extreme Learning Machine Networks (ELM) [6] and the hybrid Echo State Network - Extreme Learning Machine Networks (ESN-ELM) [7]. In the approach investigated here, it is proposed a black-box model [8] for the time-domain response of a QDL. It is discussed how it can be used to optically filter the switch-on transient of QDL total output power into the one corresponding to the spectral region around GS and the other one, around the first ES, with no need for prior knowledge of the inherent physics of the device nor its design parameters. Thus, the reader should realize how this can be useful for the purpose of measurement instrumentation

So, we do not use ANN to optimize or help the manufacturing process, like in [9], or to help set parameters for optimization in certain applications as in [10] and [11]. Indeed, as the following pages will reveal along with the results, the concepts introduced by the proposed modeling approach can help to develop the electronics of low-resolution optical filters in those applications in which an optical spectrum analyzer is not a choice for cost issues and only simple optical power meters are available.

2 Theory

2.1 Quantum Dot Semiconductor Lasers

Semiconductor Quantum Dot is a nanometer-sized structure able to confine charge carriers in the three spatial directions, which is known to affect its optical emission properties [2] and, as such, can be used for engineering of emitting or light detection devices. For example, Quantum Dot Lasers rely on the use of quantum dots in the active region and their intrinsic high differential gain to produce coherent light through stimulated photon emission following the recombination of electron-hole pairs confined therein.

In the literature, much of the comprehension of the underlying physics of quantum dot devices can be achieved with the help of a simplified description of carrier scattering events in which they taking place at a regular and periodical time-basis. Figure 1 illustrates that. It shows the QDL carrier dynamics as a very simple three-level system. The first level, the Carrier Reservoir (RS), is where the carriers are injected, followed by a capture process into the ES (second level) (at time τ_{ES}^{RS}) and relax into the GS (third level) (at time τ_{GS}^{ES}). However, carriers also can be excited (thermally), returning from GS to ES (at time τ_{ES}^{GS}), and from ES to RS (at time τ_{RS}^{ES}). Additionally, carriers can recombine spontaneously in all three levels (at time τ_{RS}^{spon} , τ_{ES}^{spon} and τ_{GS}^{spon}) [2].

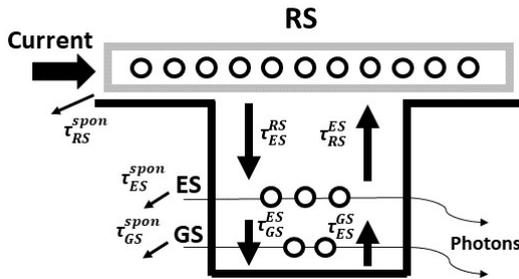


Fig. 1 Schematics of excitonic description for carrier dynamics in a QDL.

Simplified diagrams like the Figure 1 are essential to produce nonlinear coupled rate-equations models. Although these are a popular choice among researchers in the QDL scientific community, they are complex in the number of equations and, for the purpose of proper time-domain response prediction, they usually require sophisticated description of the underlying carrier-carrier scattering phenomena and the photon-matter interaction, which are expressed as:

$$\frac{dN_{RS}}{dt} = \frac{I}{q} + \frac{N_{ES}}{\tau_{RS}^{ES}} - \frac{N_{RS}}{\tau_{ES}^{RS}}(1 - \rho_{ES}) - \frac{N_{RS}}{\tau_{RS}^{spon}} + F_{RS} \quad (1)$$

4 *Dynamic Response Prediction of Bi-State Emission of QDL Based on ML*

$$\begin{aligned} \frac{dN_{ES}}{dt} = & \left(\frac{N_{RS}}{\tau_{RS}} + \frac{N_{GS}}{\tau_{GS}} \right) (1 - \rho_{ES}) - \frac{N_{ES}}{\tau_{ES}} (1 - \rho_{GS}) \\ & - \frac{N_{ES}}{\tau_{RS}} - \Gamma_p v_g g_{ES} S_{ES} - \frac{N_{ES}}{\tau_{ES}^{spont}} + F_{ES} \end{aligned} \quad (2)$$

$$\frac{dN_{GS}}{dt} = \frac{N_{ES}}{\tau_{GS}} (1 - \rho_{GS}) - \frac{N_{GS}}{\tau_{ES}} (1 - \rho_{ES}) - \Gamma_p v_g g_{GS} S_{GS} - \frac{N_{GS}}{\tau_{GS}^{spont}} + F_{GS} \quad (3)$$

$$\frac{dS_{ES}}{dt} = \left(\Gamma_p v_g g_{ES} - \frac{1}{\tau_p} \right) S_{ES} + \beta_{sp} \frac{N_{ES}}{\tau_{ES}^{spont}} + F_{S_{ES}} \quad (4)$$

$$\frac{dS_{GS}}{dt} = \left(\Gamma_p v_g g_{GS} - \frac{1}{\tau_p} \right) S_{GS} + \beta_{sp} \frac{N_{GS}}{\tau_{GS}^{spont}} + F_{S_{GS}} \quad (5)$$

$$\frac{d\phi_{ES}}{dt} = \frac{1}{2} \Gamma_p v_g (g_{GS} k_{ES}^{GS} + g_{ES} \alpha_{ES} + g_{RS} k_{ES}^{RS}) + F_{\phi_{ES}} \quad (6)$$

$$\frac{d\phi_{GS}}{dt} = \frac{1}{2} \Gamma_p v_g (g_{GS} \alpha_{GS} + g_{ES} k_{GS}^{ES} + g_{RS} k_{GS}^{RS}) + F_{\phi_{GS}} \quad (7)$$

where I is the injected current in RS (Fig. 1), q is the elementary charge and ρ_{GS} and ρ_{ES} are the carrier occupation probabilities in ES and GS, respectively. ρ_p is the photon lifetime; γ_p is the optical confinement factor; β_{sp} is the spontaneous emission factor and v_g is the group velocity of the light. α_{GS} and α_{ES} are the carrier population self-contributions in the GS and ES to the linewidth broadening factor (LBF)[2] of GS and ES, respectively. k_{ES}^{RS} and k_{ES}^{GS} are the carrier population cross-contributions in the RS and GS to the LBF of ES, respectively. k_{GS}^{RS} and k_{GS}^{ES} are the carrier population cross-contributions in the RS and ES to the LBF of GS, respectively. g_{GS} and g_{ES} are material gains of GS and ES, respectively, given by [2]:

$$g_{ES} = \frac{\alpha_{ES}}{1 + \zeta_{ES} \frac{S_{ES}}{V_S}} \frac{N_B}{V_B} (2\rho_{ES} - 1) \quad (8)$$

$$g_{GS} = \frac{\alpha_{GS}}{1 + \zeta_{GS} \frac{S_{GS}}{V_S}} \frac{N_B}{V_B} (2\rho_{GS} - 1) \quad (9)$$

where α_{ES} and α_{GS} are the differential gains; ζ_{ES} and ζ_{GS} are the gain compression factors of ES and GS, respectively; N_B is the QDL total number; V_B is the active region volume; and V_S is the volume occupied by the photons in the laser cavity and g_{RS} is the RS gain. F_{RS} , F_{ES} and F_{GS} are the carrier noise sources in RS, ES and GS, respectively. $F_{S_{ES}}$ and $F_{S_{GS}}$ are the spontaneous emission photon noise sources and $F_{\phi_{ES}}$ and $F_{\phi_{GS}}$ are the spontaneous emission phase noise sources in ES and GS, respectively[2].

So, using rate equation modeling for the purpose discussed would be to use a sledgehammer to crack a nut. As an alternative, we consider the QDL and

its intrinsic nonlinear behavior from an input-output pictorial representation, thus encompassing the dynamics from a nonlinear mathematical model able to gather the way this very rich and intricate dynamics appears at the output for different operating conditions. Such a modeling approach is the subject of the upcoming sections.

2.2 Identification and Modeling of Nonlinear Dynamics System

Modeling and identifying nonlinear dynamic systems is a challenging task because nonlinear processes do not share many properties with each other, each system has its unique characteristics, making the ability to describe the different classes of system a universal problem [12].

As a simple example of the identification and modeling systems task, the Fig. 2 illustrates a process, assuming that has a Single Input and a Single Output (SISO), with $\mu(k)$ being the system and model input value at time instant k ; $n(k)$ is the model noise; $y(k)$ is the process output and $y'(k)$ the model output, also known as system output estimate. Estimate and real system output are compared and generate the error $e(k)$, which feeds the model for parameter adjustments and increased reliability.

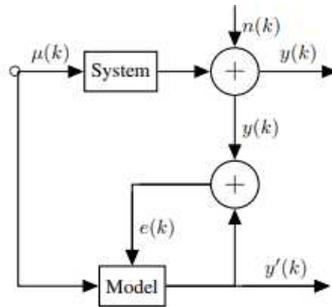


Fig. 2 Systems Identification Model.

There are several ways to classify modeling techniques, one of which is called black-box modeling (or empirical modeling), which uses little or no prior knowledge of the system. So, it is not necessary for the user to have knowledge about the physical laws, the complicated functions and variables, that govern the behavior of the system [8].

The Figure 3 shows the flowchart of the main steps for the black-box modeling of the system. It starts by determining the model input, and then the boot signal of the system, which is the injected current in the case of the laser here investigated. The model architecture is then chosen according to the limitations and objectives of the study. After that, the design takes place with an usual mathematical representation of dynamics systems, in which simplest models are preferred over higher order ones, avoiding the unnecessary increase

in complexity and estimation of unknown parameters; It ends with a validation procedure, in which the model performance is assessed and it is analyzed whether it meets all expectations or if additional refinement is required.

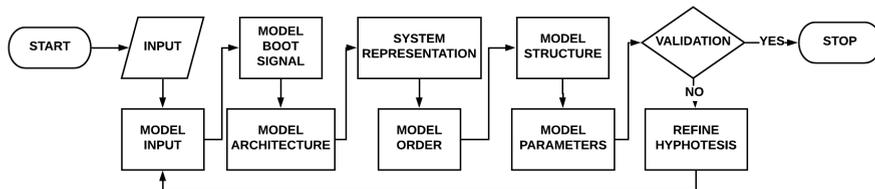


Fig. 3 Dynamic Systems Identification Flowchart.

2.3 Multilayer Perceptron (MLP)

One approach for estimating the nonlinearity of the mathematical modeling of dynamic systems is to use an ANN. ANNs are composed of interconnected layers of processing units called neurons, where the output of a single neuron (y) with n inputs (x) can be defined by:

$$y = f\left(\sum_{j=1}^n w_j x_j + b\right) \quad (10)$$

where b is the bias, w_j is the weight, both are constant, and $f()$ is called an activation function.

In the topology called MLP (Multilayer Perceptron), the output of one neuron feeds the input of all others neurons belonging to the next layer. Therefore, the output of a network with a single node in the output layer and a hidden layer is a non-linear function in the parameters of the type:

$$y = f_s\left[\sum_{i=1}^m w_i f_i\left(\sum_{j=1}^n w_{ij} x_j + b_i\right) + b_s\right] \quad (11)$$

where $f_s()$ is the neuron activation function of the output layer for a single hidden layer feedforward architecture (Single Hidden Layer Feedforward Neural Networks - SLFNs). Such a function need not be equal to $f_i()$, $i = 1, \dots, m$, which in turn, need not be equal to each other. The parameter b_s is the polarization term of the neuron of the output layer, w_i are the weights of the output of each neuron in the hidden layer and w_{ij} are the weights of the input j as regarded by the i -th neuron of the hidden layer.

It is worth mentioning that a MLP must undergo a training phase to define the ideal synaptic weights, thus favoring the modeling itself. A popularly used method for that is the error backpropagation algorithm [5], which basically

has two phases: the first concerns the synaptic weights of the network are randomly initialized, and the input propagates throughout the network, layer by layer, up to the exit. At this stage, changes are confined to the potential activation and outputs of neurons in the network.

In the second step of the training, it occurs a backpropagation of the output, and an error signal is produced and compared with the desired network output. The resulting error signal is propagated through the network, again layer by layer, but this time the propagation is performed backwards, resulting in successive adjustments in the ANN synaptic weights [5].

Other examples of training functions that can be applied in the adjustment of synaptic weights and used in the current work, are Levenberg-Marquardt [13], Bayesian Regularization[14] and Scaled Conjugate Gradient[15].

2.4 Extreme Learning Machine (ELM)

Another ANN that can learn the in-out nonlinear mapping function is the ELM, which is a learning technique for training SLFNs. Compared to other ANN techniques, the main difference is the significant increase in training speed through random generation of input weights and the bias of the hidden layer [6]. Essentially, the nonlinear relation between input and output is written according the following equations:

$$f_L(x) = \sum_{i=1}^L \beta_i h_i(x) = h(x)\beta \quad (12)$$

$$h_i(x) = G(a_i, b_i, x) \quad (13)$$

$$\beta \in \mathbb{R}^{L \times m} \min_{\|H\beta - T\|^2} \quad (14)$$

$$\beta^* = H^+ T. \quad (15)$$

where $\beta = [\beta_1, \dots, \beta_L]^T$ is the weight vector that connects the hidden layer with L nodes and the $m \geq 1$ output nodes; $h(x) = [h_1(x), \dots, h_L(x)]$ is the nonlinear mapping function; $G(a_i, b_i, x)$ (with the parameters of hidden nodes (a, b)) is a nonlinear continuous function; H is the output matrix of the hidden layer (randomly generated); T is the training data matrix; $\|\cdot\|$ is the Frobenius method ; H^+ denotes the Moore – Penrose generalized inverse matrix of H .

In other words, ELM trains SLFN in two main phases: the first related to the random mapping of the characteristics and the second about the linear solution of the parameters. In the first step, ELM randomly initializes the hidden layer to map the input data in a feature space through some nonlinear mapping functions, which can be any nonlinear continuous function (Gaussian, Multi-quadratic, hyperbolic tangent, etc.)[16].

In the second step, the weights are related to the hidden layer and the output layer, indicated by β , and are estimated by decreasing the approximation error, observing the quadratic error, as described in Eq. 14, and the optimal solution found by Eq. 15.

2.5 Echo State Network - Extreme Learning Machine (ESN-ELM)

Despite the attractiveness of the structural simplicity and efficiency of the ELM training process, it is possible to obtain an improvement in its performance by using an Echo State Network (ESN) [17] as the input layer; this leads to a ESN-ELM hybrid model with a layer (the reservoir) to enhance the non-linear characteristic of the ANN and based on the ELM combiner [7].

The Echo State Network (ESN) is an ANN that has three distinct layers: input layer, reservoir and output layer. The ESN recursive equation (without output feedback connections) resembles like the following:

$$x(n) = f(W^{in}u(n) + W_x(n-1)) \quad (16)$$

$$y(n) = f^{out}(W^{out}X(n)) \quad (17)$$

$$W^{out} = ((H^T H)^{-1} H^T S)^T \quad (18)$$

where $u(n)$, $x(n)$ and $y(n)$ are the input, the internal state of the reservoir and the output vector, respectively ; $W^{in} \in \mathbb{R}^{NxK}$, $W \in \mathbb{R}^{NxN}$ and $W^{out} \in \mathbb{R}^{Lx(Nxk)}$ are the weights of the input, reservoir and output layers, respectively; K input units provide external stimulation $u(n) \in \mathbb{R}^K$; N refers to internal units (neurons), having internal states $x(n) \in \mathbb{R}^N$; L output units generating the output signal $y(n) \in \mathbb{R}^L$; $X(n) = [x(n)^T, u(n)^T]^T$ the concatenation of internal states and input vector; $f(\cdot) = [f_1(\cdot), f_2(\cdot), \dots, f_N(\cdot)]^T$ and $f^{out}(\cdot) = [f_1^{out}(\cdot), f_2^{out}(\cdot), \dots, f_L^{out}(\cdot)]^T$ are the internal states and the output activation function, respectively; $H = [X(1), X(2), \dots, X(n, \max)]^T$ and $S = [S(1), S(2), \dots, S(n, \max)]^T$ are the internal state and target output matrix [17].

Additionally, for ESN to work properly, it is vital to choose the proper startup parameters. One of them is the Reservoir Size, which is the number of internal neurons within the ESN reservoir, probably the most determining aspect, because it directly impacts the maximum number of synaptic connections. Very small reservoirs will result in the inability to model accurately, while very large reservoirs can cause problems related to data overfitting.

Another parameter is the Connectivity σ , the fractional number of recurrent connections within the reservoir that affects the reservoir complexity. Too many connections can reduce the diversity of reservoir states, affecting ESN training capability. In practice it is common to use values in the range between [0.01; 0.2] [18].

The third main parameter is the Spectral radius ρ , the reservoir “memory” length. The larger the spectral radius, the more the previous inputs influence the current output, so to ensure the dynamics stability, the size of the spectral radius ρ , which is the highest absolute eigenvalue of the connection weight matrix W , must be correctly controlled.

Another hybrid network alternative using the ESN would be the ESN-MLP. Although it improves the ability for mapping, consequently improving the performance of the ANN, this alternative causes ESN to lose one of its main

advantages, which is the simplicity in the training process, implying longer processing time on training [19].

3 Methods

3.1 Previous Experiment and Characteristics of Experimental Data

The present study used measurements made available by the Semiconductor Optics Group of the Technical University of Darmstadt (Germany). The experiment basically consists of injecting direct current into the QDL, controlling its operating temperature with a specific controller. The QDL optical power passes through an optical insulator (with the aim of allowing the emitted power to travel in only one direction), and then an inference filter (to narrow the spectral band), where it will later be detected by optical power meter and sampled by a fast photo-detector, as showed in [20].

The results of the experiment consists of 147 time-series of switch-on transient (in step response) of InAs/InGaAs QDL separated into GS, ES and total output power as well. Each time-series has 510 samples corresponding to 6 ns long duty-time of a square-wave like electrical driving.

The time-series were acquired for different current amplitudes in which the upper level changed progressively, with steps of 1 mA, from 64.9 mA up to 139.9 mA (@ 20 °C), illustrated in Fig. 4 and 5, and from 114.9 mA up to 185 mA (@ 40 °C), showed in Fig. 6 and 7.

Thus, Figures 4 and 5 illustrate the Output Optical Power (y axis) of the GS and ES, respectively, highlighting the predominance of emission in the GS for the operating temperature of 20 °C. In this case, the ES is only capable of emitting power from an injection current (x axis) higher than 100 mA, but for a short period of time (z axis).

Differently, Figures 6 and 7 show the change in the DSE when the operating temperature is increased. The ES emission is more relevant, in detriment of GS emission. Another relevant aspect is the temporal evaluation of the emission, where it is seen that for emission at 40 °C, the ES state emits first than the GS state, and the two only present parallel emission at the end of the sampled period in higher currents, above of 160 mA.

3.2 Proposed Model

The proposed black-box model consists in an ANN which implements the nonlinear function ϕ in Eq. 19; it represents the mapping from the input space containing the total optical power, $P(n)$; the electrical current, I ; the operating temperature, T ; and the time-vector, $t(n)$; to the output space containing the GS and ES optical power at each discrete time instant, $t(n)$. In turn, Eq. 20 describes the input optical power, $P(n)$, and shows that the model needs a memory of $n - k$ samples of the total optical power to support the time-series prediction.

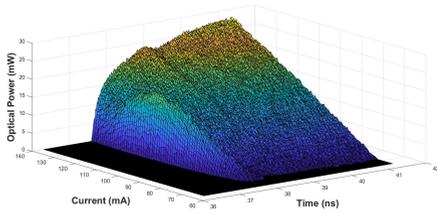


Fig. 4 GS Power @ 20 °C

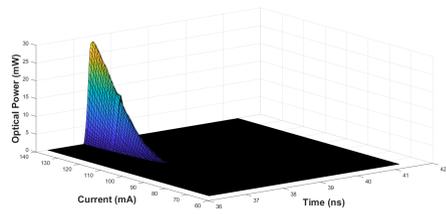


Fig. 5 ES Power @ 20 °C

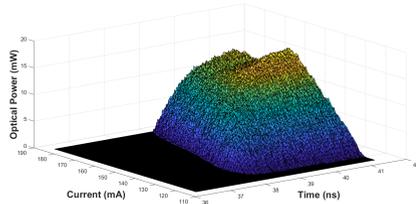


Fig. 6 GS Power @ 40 °C

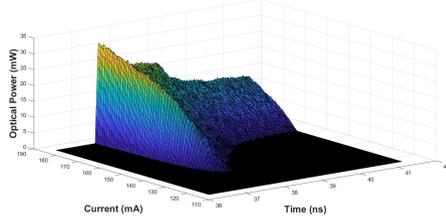


Fig. 7 ES Power @ 40 °C

$$P^{GS}(n), P^{ES}(n) = \phi(P(n), I, T, t(n)) \quad (19)$$

$$P(n) = P^T(n), P^T(n-1), \dots, P^T(n-k) \quad (20)$$

where P^{GS} and P^{ES} are Ground State (GS) and Excited State (ES) optical output power, respectively; ϕ represents the ANN nonlinear function and $P(n)$ is the array of the output power $P^T(n)$ and lagged up to $n - k$.

The Fig. 8 shows the model represented by Eq. 19 and 20, highlighting the outputs (Ground and Excited State Power), resulting from the Total Optical Power (and other inputs), so, the proposed model performs the function of an optical filter.

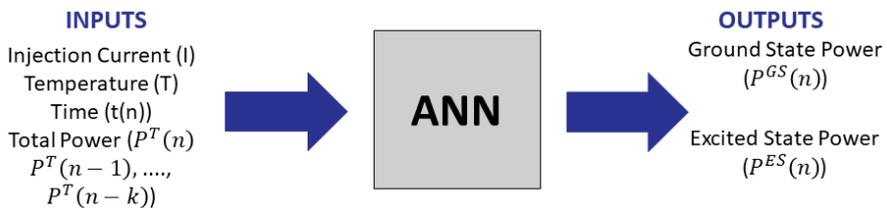


Fig. 8 Proposed Model.

3.3 Design parameters

To determine the quantity of the delayed input samples in Eq. 20, k , we calculated the Auto-correlation Function (ACF) and the Partial Auto-correlation Function (PACF) [21] of the vector P^T . Thus, it is possible to choose the values that have the highest correlation, after sorting them in a descending way,

excluding the values of low relevance, that is, choosing only the values that theoretically help ANN to be able to predict the system dynamics, regardless of its order.

Another design parameter investigated was the size of the hidden layer, which ranged from 1 to 1,000 hidden nodes (HN) with different activation functions. For the ELM model, the functions Sigmoidal (Sine), Radial Basis (Radbis) and Triangular Basis (Tribas) were used, while in the ESN-ELM model, the function the Hyperbolic Tangent (Tanh), as well as the Sine and Tribas function were used.

In the case of the Multilayer Perceptron (MLP), we used the ANN with size ranging from 2 to 50 neurons (Ne) with Levenberg-Marquardt (LM) [13], Bayesian Regularization (BR)[14] and Scaled Conjugate Gradient (SCG)[15] as training function.

The data were grouped into training (70%) and testing (30%) samples to develop the models. Additionally, for the best performance of the algorithm, avoiding distortions due to the influence in the training of certain input variables with a greater dynamic range of values in relation to the others, all samples of learning and test input are normalized to the interval [0; 1].

3.4 Performance Evaluation

This work used one of the main performance indexes in regression model evaluation, the Root Mean Square Error (RMSE), defined by:

$$RMSE = \frac{\sum_{n=1}^m e_n^2}{m} \quad (21)$$

where e_n is the output error at time n , which is obtained by the difference between the real value and the value estimated by the ANN for the set of m output samples evaluated.

The Mean Absolute Percentage Error (MAPE) is also used, defined by:

$$MAPE = \frac{\sum_{n=1}^m EP_n}{m} \quad (22)$$

where EP_n is the percent error of output at time n and m again is the number of output samples evaluated.

4 Results

First, as mentioned in Subsection 3.3, we used the ACF and the PACF of the vector P^T to determinate k in the Eq. 20, so, the Fig. 9 shows the PACF, where the first five lagged values with the highest correlation are those of 1st,

2nd, 3rd, 5th and 6th order (which were used in the model), so the 4th lagged value has no relevant correlation and was discarded as input to the model.

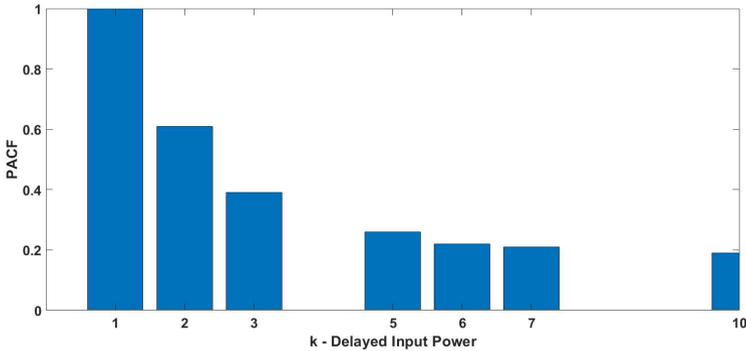


Fig. 9 Partial Auto-correlation Function.

Thus, the Tab. 1 shows the MLP model (top), ELM Model (middle) and ESN-ELM Model (bottom) performance evaluations, highlighting the best results obtained for each training relative to different activation functions. The worst result among those obtained by the MLP and ELM model was with RMSE of $15.47 \mu\text{W}$ and MAPE of 2.59 %; these are low numbers, and as such, we conclude about the application of both model in any verified scenarios. In contrast, the ESN-ELM model presented much higher error values, making its application unfeasible in most of the configurations used.

The MLP best result is using the Levenberg-Marquardt training function without Delayed Input Power (k in Eq.20) and 48 neurons (Ne), resulting in an RMSE of $8.40 \mu\text{W}$ and MAPE of 1.32%. For this training function, increasing of the model order did not sound as a good choice, because it increased the complexity with the insertion of more delayed input powers and the performance got worse.

In contrast, the MLP with the Bayesian Regularization training function has its best result with 1 (one) Delayed Input Power and 46 neurons, with RMSE of $8.69 \mu\text{W}$, and MAPE of 1.46%. However, that result is 3.45% worse than that obtained with the Levenberg-Marquardt function.

In the case of MLP using Scaled Conjugate Gradient function, the model had its worst results, and its best result is using 3 (three) Delayed Input Power elements with only 38 neurons (amount less than the average of the other functions), with RMSE of $9.10 \mu\text{W}$ and MAPE of 2.45%.

On the other hand, the ELM Model showed error lower than MLP for all quantities of delayed input powers and its best overall performance is using 2 (two) Delayed Input Power with 889 Hidden Nodes (HN) and the Radbas as activation function, resulting in a RMSE of $2.81 \mu\text{W}$ and MAPE of 0.50% with processing time (training and testing time) of 15.27 s. It means that the ELM has error 65% lower than that presented by MLP, showing its superiority

for this application, and a processing time much lower than the rate equation processing time.

The best results with the Sine and Tribas activation functions are RMSE of $2.89 \mu\text{W}$ and MAPE of 0.54%, and RMSE of $3.29 \mu\text{W}$ and MAPE of 0.66%, respectively. They are inferior when compared to the results with Radbas, but they are still superior to the best results obtained with the MLP.

Another issue is the relationship between the number of hidden nodes and the amount of delayed optical power, which was, in general, an inversely proportional trend in the ELM Model (except when used with Tribas activation function). This can help to choose the configuration to be implemented in hardware, in future works, provided that the processing capacity and available memory in the device are met.

Despite the expectation of performance improvement with the insertion of the ESN with the ELM, the result obtained was inferior to the ELM model; it led to error levels similar to the other models only using the Hyperbolic Tangent (Tanh) as an activation function. In this configuration, the model had its best result with RMSE of $5.56 \mu\text{W}$ and MAPE of 1.02%, with $k = 1$ and 144 hidden nodes.

The other two activation functions used, Sine and Tribas, had RMSE values much higher than those found in the other models, with $43.69 \mu\text{W}$ and $196.25 \mu\text{W}$, respectively, at its best performances, though the MAPE values presented were not as distorted, with 2.93% and 6.08%, respectively. These results show how important it is to use more than one parameter for model validation, thus preventing inadequate conclusions.

The main reason for the low performance of the ESN-ELM model is the use of non-optimized values for the parameters of the ESN part of the model. One way of mitigating the distortions found, consequently increasing the performance of the model, is the use of optimization algorithms such as Particle Swarm Optimization (PSO) [22] and Artificial Bee Colony (ABC)[23].

In addition to the validation parameters used and for better visualization of the performance achieved so far in this study, the Fig. 10 shows the experimental and theoretically predicted time-series of a switch-on transient experiment with $I = 123 \text{ mA}$ at $20 \text{ }^\circ\text{C}$ (a), and $I = 174 \text{ mA}$ at $40 \text{ }^\circ\text{C}$ (b). The rather good agreement between the solid (experimental) and dashed lines (predicted by the model), which we stress was obtained for a small input memory lag ($k = 2$) suggests that the proposed approach is a good candidate for the embedded electronics of low-resolution spectral analyzers.

5 Conclusion

The alternative of filtering the QDL optical power from ANN black-box model, using as input only the total power, temperature and current, is promising. The GS and ES optical power, that is, the dynamics Bi-State Emission, was

Table 1 Performance Evaluations of The Best Results

MLP MODEL									
<i>LM</i>			<i>BR</i>			<i>SCG</i>			
<i>k</i>	<i>Ne</i>	<i>RMSE</i> (μW)	<i>MAPE</i> (%)	<i>Ne</i>	<i>RMSE</i> (μW)	<i>MAPE</i> (%)	<i>Ne</i>	<i>RMSE</i> (μW)	<i>MAPE</i> (%)
0	48	8.40	1.32	45	8.70	1.51	38	15.47	2.59
1	44	8.60	1.52	46	8.69	1.46	34	9.76	2.76
2	50	9.00	1.67	44	8.84	1.64	44	9.87	2.86
3	45	9.00	1.64	22	8.69	1.78	38	9.10	2.45
4	47	8.97	1.60	39	8.89	1.70	24	9.94	2.41
5	47	9.08	1.71	22	9.15	1.83	50	10.01	2.66
ELM MODEL									
<i>Sine</i>			<i>Radbas</i>			<i>Tribas</i>			
<i>k</i>	<i>HN</i>	<i>RMSE</i> (μW)	<i>MAPE</i> (%)	<i>HN</i>	<i>RMSE</i> (μW)	<i>MAPE</i> (%)	<i>HN</i>	<i>RMSE</i> (μW)	<i>MAPE</i> (%)
0	993	11.32	1.72	891	6.92	1.19	989	7.42	1.26
1	811	2.99	0.55	369	3.11	0.59	949	3.31	0.68
2	737	2.89	0.54	889	2.81	0.50	935	3.36	0.67
3	232	3.91	0.77	572	3.39	0.57	983	3.29	0.66
4	270	4.07	0.77	362	3.76	0.67	978	3.41	0.67
5	141	5.08	0.95	281	4.09	0.81	969	3.52	0.66
ESN-ELM MODEL									
<i>Sine</i>			<i>Tanh</i>			<i>Tribas</i>			
<i>k</i>	<i>HN</i>	<i>RMSE</i> (μW)	<i>MAPE</i> (%)	<i>HN</i>	<i>RMSE</i> (μW)	<i>MAPE</i> (%)	<i>HN</i>	<i>RMSE</i> (μW)	<i>MAPE</i> (%)
0	13	92.32	2.96	378	12.34	1.48	172	243.67	7.12
1	6	64.52	3.33	144	5.56	1.02	135	196.25	6.08
2	14	43.69	2.93	140	5.91	1.15	277	197.49	10.36
3	3	62.59	2.61	178	6.59	1.33	298	240.67	11.07
4	6	49.54	4.19	100	8.58	1.35	73	217.11	9.53
5	1	57.73	3.06	100	8.48	1.73	43	209.15	6.68

predicted with ELM- and MLP-based approaches producing very low prediction errors. Additionally, they presented the processing time in the order of seconds, much lower than the rate equation model.

In the case of the ESN-ELM model, contrary to expectations, it presented, in most of the tested configurations, results much worse than those presented by the other ANNs. Thus, it is necessary to evaluate its performance using optimization algorithms such as PSO and ABC, and again verify the feasibility of its application in the model.

The success of the alternative model presented is relevant in the context of optical filtering instrumentation, with the possibility of exploring the concept in the future implementation. Thus, it is possible to develop low-resolution and low-cost filters for certain applications in which it is not necessary and/or

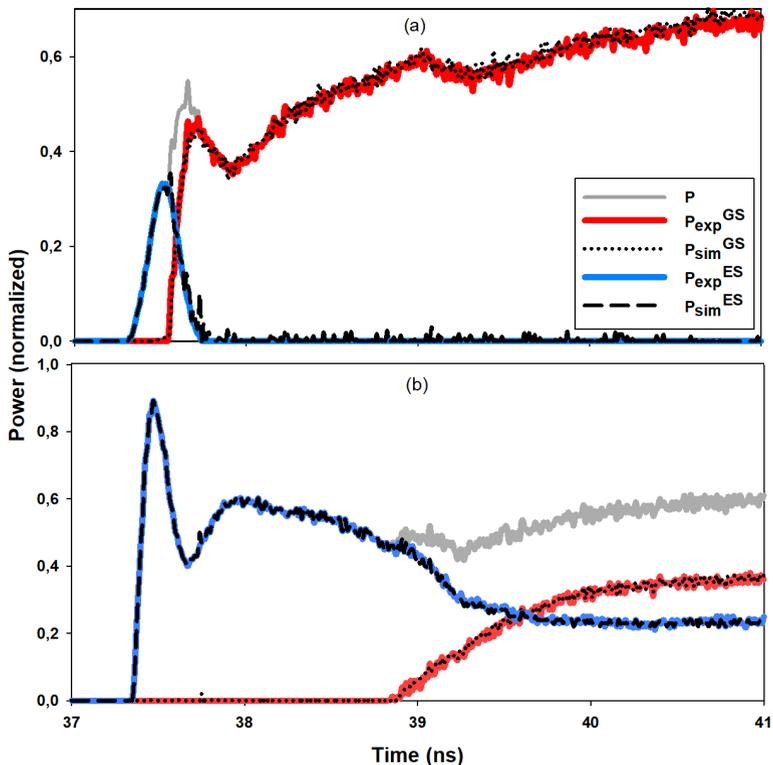


Fig. 10 ELM model time-response with (a) $I = 123$ mA at 20 °C and (b) $I = 174$ mA at 40 °C, where P is the QDL Optical Power; P_{exp}^{GS} and P_{sim}^{GS} are the Ground State Power, experimental and simulated, respectively; P_{exp}^{ES} and P_{sim}^{ES} are the Excited State Power, experimental and simulated, respectively.

economically viable to use of a spectrum analyzer, which can be replaced by a simple optical power meter.

As future work, we will investigate a case of higher spectral resolution at least in continuous-wave operation, checking the ability of the inference machine to predict different lasing lines around the GS and ES. The experimental data necessary for this investigation is available and is currently under preparation.

Acknowledgments. We thank the Semiconductor Optics Group of the Technical University of Darmstadt (Germany). This paper and the research it would not have been possible without their support, providing the experimental data used in the work.

Declarations

Not applicable

References

- [1] Hussein Abdali, R., Abdullattif Ghalib, B.: The effect of delay time on the synchronization of quantum dot semiconductor laser with optical feedback. *Journal of Ovonic Research* **17**, 61–71 (2021)
- [2] Zhou, Y., Duan, J., Grillot, F., Wang, C.: Optical noise of dual-state lasing quantum dot lasers. *IEEE Journal of Quantum Electronics* **56**(6), 1–7 (2020). <https://doi.org/10.1109/JQE.2020.3026090>
- [3] Gioannini, M., Dommermuth, M., Drzewietzki, L., Breuer, S.: Dual-state quantum dot laser self-interferometric sensing. In: 2014 International Semiconductor Laser Conference, pp. 94–95 (2014). <https://doi.org/10.1109/ISLC.2014.180>
- [4] Duan, J., Zhou, Y., Huang, H., Dong, B., Wang, C., Grillot, F.: Dynamic properties of two-state lasing quantum dot laser for external optical feedback resistant applications. In: 2020 International Conference on Numerical Simulation of Optoelectronic Devices (NUSOD), pp. 79–80 (2020). <https://doi.org/10.1109/NUSOD49422.2020.9217720>
- [5] Aggarwal, C.C.: *Neural Networks and Deep Learning: A Textbook*, 1st edn. Springer, Cham, Switzerland (2018)
- [6] Tang, J., Deng, C., Huang, G.-B.: Extreme learning machine for multilayer perceptron. *IEEE Trans. Neural Networks Learn. Syst.* **27**(4), 809–821 (2016)
- [7] Butcher, J., Verstraeten, D., Schrauwen, B., Day, C., Haycock, P.: Extending reservoir computing with random static projections: a hybrid between extreme learning and rc. (2010)
- [8] Rudin, C., Radin, J.: Why are we using black box models in ai when we don't need to? a lesson from an explainable ai competition. *Harvard Data Science Review* **1**(2) (2019). <https://doi.org/10.1162/99608f92.5a8a3a3d>. <https://hdsr.mitpress.mit.edu/pub/f9kuryi8>
- [9] Rinderle, M., Gagliardi, A.: Machine learning amp; multiscale simulations: toward fast screening of organic semiconductor materials. In: 2021 International Conference on Numerical Simulation of Optoelectronic Devices (NUSOD) (2021). <https://doi.org/10.1109/NUSOD52207.2021.9541414>
- [10] Djordjevic, K., Galovic, S., Jordovic Pavlovic, M., Nestic, M., Popovic, M., Cojbasic, Z., Markushev, D.: Photoacoustic optical semiconductor characterization based on machine learning and reverse-back procedure. *Optical and Quantum Electronics* **52** (2020). <https://doi.org/10.1007/s11082-020-02373-x>

- [11] Jordovic Pavlovic, M., Kupusinac, A., Djordjevic, K., Galovic, S., Markushhev, D., Nestic, M., Popovic, M.: Computationally intelligent description of a photoacoustic detector. *Optical and Quantum Electronics* **52** (2020). <https://doi.org/10.1007/s11082-020-02372-y>
- [12] Nelles, O.: *Nonlinear System Identification: From Classical Approaches to Neural Networks, Fuzzy Models, and Gaussian Processes*, p. 1203. Springer, Cham, Switzerland (2020). <https://doi.org/10.1007/978-3-030-47439-3>
- [13] Rubio, J.d.J.: Stability analysis of the modified levenberg–marquardt algorithm for the artificial neural network training. *IEEE Transactions on Neural Networks and Learning Systems* **32**(8), 3510–3524 (2021). <https://doi.org/10.1109/TNNLS.2020.3015200>
- [14] Khafaf, N., El-Hag, A.: Bayesian regularization of neural network to predict leakage current in a salt fog environment. *IEEE Transactions on Dielectrics and Electrical Insulation* **25**, 686–693 (2018). <https://doi.org/10.1109/TDEI.2017.006936>
- [15] Khadse, C., Chaudhari, M., Borghate, V.B.: Electromagnetic compatibility estimator using scaled conjugate gradient backpropagation based artificial neural network. *IEEE Transactions on Industrial Informatics* **PP**, 1–1 (2016). <https://doi.org/10.1109/TII.2016.2605623>
- [16] Huang, G., Huang, G.-B., Song, S., You, K.: Trends in extreme learning machines: A review. *Neural Networks* **61**, 32–48 (2015). <https://doi.org/10.1016/j.neunet.2014.10.001>
- [17] Jaeger, H.: The "echo state" approach to analysing and training recurrent neural networks-with an erratum note'. Bonn, Germany: German National Research Center for Information Technology GMD Technical Report **148** (2001)
- [18] Bala, A., Ismail, I., Ibrahim, R., Sait, S.: Applications of metaheuristics in reservoir computing techniques: A review. *IEEE Access* **6**, 58012–58029 (2018). <https://doi.org/10.1109/ACCESS.2018.2873770>
- [19] Shao, X., Shi, Y.: Neural-network-based constrained output-feedback control for mems gyroscopes considering scarce transmission bandwidth. *IEEE Transactions on Cybernetics*, 1–13 (2021). <https://doi.org/10.1109/TCYB.2021.3070137>
- [20] Drzewietzki, L., Thè, G.A.P., Gioannini, M., Breuer, S., Montrosset, I., Elsässer, W., Hopkinson, M., Krakowski, M.: Theoretical and experimental investigations of the temperature dependent continuous wave lasing characteristics and the switch-on dynamics of an inas/ingaas quantum-dot

semiconductor laser. *Optics Communications* **283**(24), 5092–5098 (2010).
<https://doi.org/10.1016/j.optcom.2010.07.013>

- [21] Shumway, R.H., Stoffer, D.S.: *Time Series Analysis and Its Applications: with R Examples*. Springer, New York, USA (2017)
- [22] Chouikhi, N., Ammar, B., Rokbani, N., Alimi, A.M.: Pso-based analysis of echo state network parameters for time series forecasting. *Applied Soft Computing* **55**, 211–225 (2017). <https://doi.org/10.1016/j.asoc.2017.01.049>
- [23] Amaya, E., Alvares, A.: Prognostic of rul based on echo state network optimized by artificial bee colony. *International Journal of Prognostics and Health Management* **7** (2016). <https://doi.org/10.36001/ijphm.2016.v7i1.2359>