Deep Learning-Driven Extraction of Superluminescent Diodes Parameters

Andrea Marchisio
Politecnico di Torino
andrea_marchisio@polito.it

Vittorio Curri
Politecnico di Torino
vittorio.curri@polito.it

Andrea Carena
Politecnico di Torino
andrea.carena@polito.it

Paolo Bardella
Politecnico di Torino
paolo.bardella@polito.it

***Abstract*—We present a deep learning-based method for the automatic extraction of physical parameters from optical spectra and power values of a chirped, tapered, dual-section quantum dot superluminescent diode. The neural network is able to estimate a set of parameters that are capable of reproducing the behavior of the target device with high accuracy.**

***Index Terms*—Superluminescent diodes, deep learning, parameter extraction, time-domain traveling wave, quantum dots.**

I. Introduction

Superluminescent diodes (SLEDs) are a prominent solution for sources in a wide range of applications, due to their high output power and spectral brightness. In particular, chirped and tapered Quantum Dot (QD) SLEDs are characterized by a large bandwidth, as well as a tunable spectral asymmetry between the two facets [1], which could introduce the possibility of multiplexing their outputs for flexible tuning of the device output and bandwidth. The characterization of such devices is not straightforward due to the number of parameters required to properly reproduce the experimental behavior in stationary and dynamic conditions, the latter for the amplification of optical pulses: brute-force approaches to extract these parameters are not recommended because of the high computational cost and time of those methods. A much more effective approach for the characterization or inverse design of devices relies on machine learning solutions [2]. We hereby present a Deep Learning (DL) approach for the extraction of device parameters of a chirped, tapered, and dual-section QD SLED, starting from its power and spectral data. The accuracy of the Deep Neural Network (DNN) is tested by comparing the predictions with some reference curves similar to those experimentally reported in [1], here obtained with Time-Domain Traveling Wave (TDTW) simulations.

II. Reference device

The device parameters are presented in [1]. The active material contains 10 chirped InAs QD layers: three layers with ground state (GS) emission λ_{GS} centered at $1243\,\mathrm{nm}$, three layers with $\lambda_{\mathrm{GS}} = 1211\,\mathrm{nm}$, and four with $\lambda_{\mathrm{GS}} = 1285\,\mathrm{nm}$. The device is $6\,\mathrm{mm}$ long, with an initial $14\,\mu\mathrm{m}$ wide straight ridge waveguide (rear of the device), followed by a first tapered section with a full taper angle of $3°$ and a second tapered section with a full taper angle of $0.8°$, which is equivalent to a total width of $110\,\mu\mathrm{m}$ at the front facet. At $1.875\,\mathrm{mm}$ from the rear facet, an insulation trench is etched, creating two bias sections that can be driven independently for greater flexibility.

III. TDTW model for data-set generation

The TDTW model [3] takes into account the chirped nature of the active QD material (different layers with different ground state emission wavelengths) and the stimulated emission from the second QD excited state.

The carrier dynamics is described by a set of multi-population rate equations (MPREs), which take into account the scattering and recombination phenomena graphically summarized in Fig. 1. The population dynamics in the conduction and valence bands are assumed to be equivalent (excitonic approximation) to speed up the simulation. The propagation of the forward and backward field components $E^{\pm}$ is described with two wave equations that, under the slowly varying envelope approximation (SVEA), read [4]:

$$\frac{1}{v_{\mathrm{g}}}\frac{\partial E^{\pm}}{\partial t} \pm \frac{\partial E^{\pm}}{\partial z} = -\frac{1}{2}(\alpha_{\mathrm{i}}^{\pm}(z) + \alpha_{\mathrm{p}})E^{\pm}(z,t) - \\ -\mathrm{j}\frac{\omega_0 {\Gamma_{\mathrm{xy}}}^{\pm}(z)}{2cn_{\mathrm{eff}}\epsilon_0}P^{\pm}(z,t) + {S_{\mathrm{sp}}}^{\pm}(z,t) \tag{1}$$

with v_{g} group velocity, $E^{\pm}$ progressive and regressive components of the field, α_{p} plasma losses, ω_0 reference angular frequency, n_{eff} effective index. The intrinsic waveguide losses $\alpha_{\mathrm{i}}^{\pm}(z)$ and the transverse confinement factor ${\Gamma_{\mathrm{xy}}}^{\pm}(z)$ vary with respect to the position and have different behaviors depending on the propagation direction in the tapered structure due to the weak confinement of the field in the transverse direction. $P^{\pm}(z,t)$ is the macroscopic polarization, linked to the stimulated response of the system, while ${S_{\mathrm{sp}}}^{\pm}(z,t)$ describes the spontaneous emission process.

Finally, boundary conditions are required in order to take into account the residual facet reflectivities r_0 and r_{L} [5]:

$$E^{+}(t,0) = r_0 E^{-}(t,0) \tag{2}$$

$$E^{-}(t,L) = r_{\mathrm{L}} E^{+}(t,L) \tag{3}$$

The model is solved numerically with a finite difference scheme and it is employed to generate a training data-set for

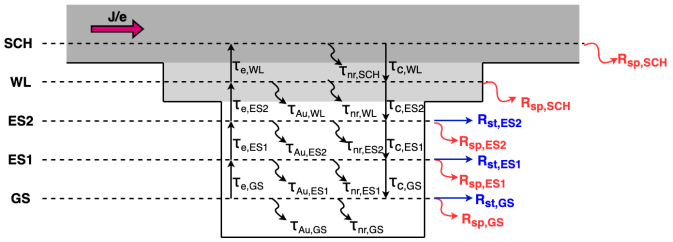

Fig. 1. Scheme of the considered levels and the scattering and recombination phenomena that affect them.

TABLE I
INVESTIGATED PARAMETERS AND CORRESPONDING RANGES.

Parameter	Range	Parameter	Range	Parameter	Range
Progr. losses α_i^+ $[\mathrm{cm}^{-1}]$	2.2–5.0	GS gain g_{GS} $[10^{-16}\,\mathrm{cm}^2\,\mathrm{eV}]$	0.2–0.6	GS Auger rec. time τ_{Aug}^{GS} [ns]	0.1–5.0
Regr. losses α_i^- $[\mathrm{cm}^{-1}]$	0.3–2.4	ES_1 gain g_{ES1} $[10^{-16}\,\mathrm{cm}^2\,\mathrm{eV}]$	0.3–0.8	ES_1 Auger rec. time τ_{Aug}^{ES1} [ns]	0.1–5.0
Injection efficiency η_i [ ]	0.5–0.8	ES_2 gain g_{ES2} $[10^{-16}\,\mathrm{cm}^2\,\mathrm{eV}]$	0.4–1	ES_2 Auger rec. time τ_{Aug}^{ES2} [ns]	0.1–5.0

DL. This is achieved by randomly generating combinations of the nine target parameters shown in Table I, as they play a fundamental role in the definition of the static and dynamical behavior of SOA. The parameters' values are randomly generated according to a uniform distribution over the associated ranges, listed in Table I, the only additional constraint being that $g_{ES2} > g_{ES1} > g_{GS}$. For each parameter set, the rear current is fixed to $I_R = 100\,\mathrm{mA}$, while the front current I_F is changed from 1 A to 5 A. A single simulation lasts 5 ns. Each entry in the data-set contains the front and rear facet spectra (102 equally spaced points each) and the front and rear facet average output powers for the five current combination. Due to the high computational cost, training is done on a data-set of 1000 samples (60% training, 20% validation, and 20% testing).

IV. Deep Learning Scheme

The deep neural network (DNN) used for parameter extraction is developed in TensorFlow™ with the Keras API. It is a single fully connected DNN that employs the ADAM optimizer, with step-decaying learning rate, and Leaky ReLU as activation of the layers. To prevent overfitting, batch normalization and L2 regularization are considered. To maximize the accuracy of the prediction, different combinations of hyperparameters are evaluated during the training process. In particular, we tested multiple values of number of hidden layers ([2, 3, 4, 5, 6, 7]), number of neurons per layer ([20, 30, 50, 100]) and batch size ([10, 20, 50, 100, 150, 200]). During this optimization process, early stopping is implemented to prevent overfitting and to speed up the process itself. Two combinations of hyperparameters are selected since they are able to minimize the loss (mean squared error) and the average error of the predictions on the test set, defined as

$$\overline{\Delta e} = \frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{X}_i - \mathbf{X}_{\mathrm{targ}_i}}{\mathbf{X}_{\mathrm{targ}_i}} \tag{4}$$

where $N = 9$ is the number of target parameters, $\mathbf{X}_{\mathrm{targ}_i}$ is the vector of values for the i-th parameter of the test set, and $\mathbf{X}_i$ is the vector of associated predictions. We selected 3 hidden layers, 100 neurons per layer, and 50 batch size (minimum loss) and 5 hidden layers, 100 neurons per layer, and 100 batch size (minimum average error). Finally, the reference curves are fed to the trained networks and their predictions are averaged for better accuracy (model averaging). The obtained parameters are simulated once again with the TDTW solver in order to compare the spectra and output powers.

V. Results Discussion

Fig. 2 shows the comparison of the reference spectra (solid lines) and those obtained by simulating the extracted parameters (circles), at different values of front section current. The predicted parameters yield spectral curves that are in good agreement with the target. Indeed, they are characterized by the same spectral components with very similar profiles, with a limited decrease in accuracy at larger currents. In general, due to the limited size of the data-set, the predicted normalized curves present a lower peak at $\lambda = 1240\,\mathrm{nm}$ with respect to the reference cases; we expect to solve this issue by increasing the size of the data-set.

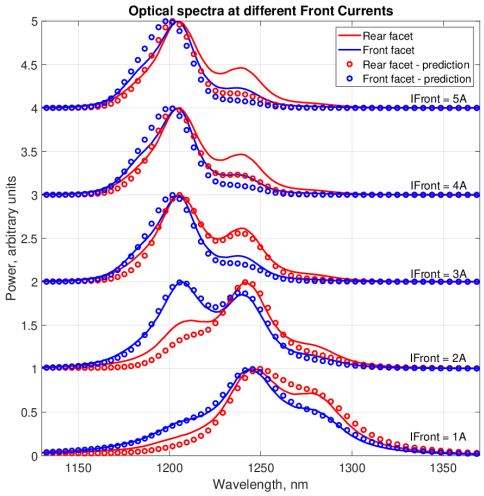

Fig. 2. Comparison of the reference and predicted spectra.

VI. Conclusions

We proposed a DL approach to the extraction of parameters from the spectral data of superluminescent diodes. While accuracy improvements can be achieved by increasing the number of entries in the data-set, the proposed DNN is able to extract a set of nine parameters that reproduce the behavior of the target device, thus providing an accurate starting point for studying the device dynamic behavior in the presence of external optical pulses.

References

[1] A. F. Forrest *et al.*, "Wide and tunable spectral asymmetry between narrow and wide facet outputs in a tapered quantum-dot superluminescent diode," *Optics express*, vol. 28, no. 2, pp. 846–859, 2020.

[2] Y. Xu *et al.*, "Interfacing photonics with artificial intelligence: an innovative design strategy for photonic structures and devices based on artificial neural networks," *Photon. Res.*, vol. 9, no. 4, pp. B135–B152, Apr 2021.

[3] A. Marchisio *et al.*, "Time-domain travelling-wave analysis of semiconductor optical amplifiers based on chirped quantum dot materials," in *CLEO proceedings*, 2023.

[4] M. Rossetti *et al.*, "Time-domain travelling-wave model for quantum dot passively mode-locked lasers," *IEEE Journal of Quantum Electronics*, vol. 47, no. 2, p. 139 – 150, 2011, cited by: 66.

[5] A. F. Forrest *et al.*, "Numerical and experimental characterization of chirped quantum dot-based semiconductor optical amplifiers," in *NUSOD proceedings*, 2021, pp. 17–18.