Proceedings of the 10th World Congress on Electrical Engineering and Computer Systems and Sciences (EECSS'24)
Barcelona, Spain - August 19 - 21, 2024
Paper No. MVML 103
DOI: 10.11159/mvml24.103

Physics-Informed Clustering For Massive Single Photon Emitter Hyperspectral Data Analysis

Juan Francisco Martínez[1], José Manuel Llorens[1,2], Enrique Navarro[1], Yolanda González[2] and Benito Alén[1,2]
[1]g2-Zero,
Calle de las Dos Doncellas, 6, Getafe, Madrid E-28906, Spain
juanf.martinez@g2-zero.com
[2] Instituto de Micro y Nanotecnología, IMN-CNM, CSIC (CEI UAM+CSIC), Isaac Newton, 8, Tres Cantos, Madrid E-28760, Spain
mailto:Third.Author@uottawa.ca

Extended Abstract

In recent decades, the escalating interest in quantum information has led to the proliferation of a wide range of quantum technologies. Notably, Single Photon Sources (SPS) based on single III-V Semiconductor Quantum Dots (QD) have emerged as a viable means to implement Quantum Key Distribution protocols as well as various other applications [1,2]. QDs are superb single photon emitters at low temperatures, showing simultaneously record brightness, purity and indistinguishability if compared with other SPS technologies [1,2]. Leveraging on well-known III-V semiconductor microfabrication techniques, QDs are the best candidates to supply the future demand of SPS sources in different quantum applications markets which g2-Zero and other companies are addressing with different proprietary designs [3].

A distinguishing and unwanted feature of QDs is their self-assembly nature during growth, which implies that their positions and spectral features cannot be precisely determined before growth. This poses an additional challenge in the SPS device fabrication, compared to classical optoelectronic devices like LEDs or VCSELs. As in any scalable development of semiconductor devices, systematic execution of physical measurements over samples and massive data analysis are indispensable, and thus the use of machine learning techniques. Hence, it is essential to devise automation software to facilitate systematic QD characterization at a single level, as well as software capable of efficiently analysing substantial volumes of data. Due to the self-assembly nature of the growth of the QDs, the primary objective of the data analysis pipeline is to process hyperspectral data produced by the device to identify the positions of the QDs and their respective emission spectra. The input data comprises the emission intensity at each point of a 2-dimensional sample within a certain wavelength range. The output information about the positions of QDs and their respective properties would serve as feedback to decide which QDs merit further exploration.

A high-spectral-resolution single-dot spectrum shows a multitude of sharp lines, resembling a barcode uniquely identifying the nanostructure [4]. The main challenge in the data analysis lies in discerning the spectral lines of the target QD from those of nearby QDs, given the presence of crosstalk. This phenomenon arises due to the optical resolution of state of art cryogenic optical microscopes typically used for the hyperspectral data analysis. Consequently, for an accurate hyperspectral characterization, it becomes necessary to distinguish genuine spectral lines of the target QD from those attributable to crosstalk. An additional request is that the algorithm should be reliable and operate at real-time during the data acquisition.

In this work, we present a *python* set of physics-informed routines designed to execute the analysis task, i.e., identifying the positions of QDs while effectively filtering crosstalk interference. This task can be regarded as a clustering routine, since its aim is to assign each of the spectral lines measured in a certain region to the QD to which it belongs. Our algorithm is grounded in the premise that spectral lines genuinely associated with a certain QD attain their maximum intensity value when that QD is at the centre of the Point Spread Function of the microscope. Conversely, spectral peaks induced by crosstalk will

reach their maximum at a different position, corresponding to a different QD. Therefore, this criterion enables to discriminate the positions and the real spectra of the QDs, with a highly satisfactory result.

Additionally, we conduct a comparative analysis of the outcomes produced by our algorithm against those obtained through unsupervised clustering algorithms, like k-means and Gaussian Mixture. However, the results obtained in this way are not consistent with a detailed inspection nor with our physics-informed algorithm.

In summary, our physics-informed clustering outperforms unsupervised algorithms such as k-means and Gaussian Mixture Models. It appears that the specific challenges posed by crosstalk interference demand a domain-knowledge-driven approach, making generic unsupervised classification algorithms insufficient for this task. This work bridges the gap between manual microscopic characterization of single QDs and fully automated cryogenic hyperspectral microscopy systems, paving the way to scale-up the fabrication of semiconductor SPS.

References

[1] C. Couteau, S. Barz, T. Durt, T. Gerrits, J. Huwer, R. Prevedel, J. Rarity, A. Shields and G. Weihs, "Applications of single photons to quantum communication and computing," *Nat Rev Phys*., vol. 5, no. 6, pp. 326–338, 2023.

[2] C. Couteau, S. Barz, T. Durt, T. Gerrits, J. Huwer, R. Prevedel, J. Rarity, A. Shields and G. Weihs, "Applications of single photons in quantum metrology, biology and the foundations of quantum physics," *Nat Rev Phys*., vol. 5, no. 6, pp. 354-363, 2023

[3] B. Alén, D. Fuster, Y. González and L. González, "Device for emitting single photons or entangled photon pairs," EP3361516B1, December 18, 2019

[4] V. Mlinar and A. Zunger, "Spectral Barcoding of Quantum Dots: Deciphering Structural Motifs from the Excitonic Spectra," *Phys. Rev. B*, vol. 80, 035328, 2009