

**Dmitriy A. Poteryayev<sup>1,2</sup>**

Pietro Novelli<sup>3</sup>, Annalisa Coriolano<sup>1,3</sup>, Riccardo Dettori<sup>4</sup>, Valentina Tozzini<sup>5,2</sup>, Fabio Beltram<sup>2</sup>, Massimiliano Pontil<sup>3</sup>, Antonio Rossi<sup>1</sup>, Stiven Forti<sup>1</sup>, Camilla Coletti<sup>1</sup>

<sup>1</sup>Center for Nanotechnology Innovation @NEST, Italian Institute of Technology, Pisa 56127, Italy

<sup>2</sup>NEST Laboratory, Scuola Normale Superiore, Pisa 56127, Italy

<sup>3</sup>Computational Statistics and Machine Learning, Italian Institute of Technology, Genova 16163, Italy

<sup>4</sup>Department of Physics, University of Cagliari, Cagliari 09042, Italy

<sup>5</sup>Nanoscience Institute, National Research Council (CNR-NANO), Pisa 56127, Italy

Contact@E-mail [dmitriy.poteryayev@sns.it](mailto:dmitriy.poteryayev@sns.it)

## Revealing Hidden Raman Signatures Through Attention-Based Spectral Unmixing

Raman spectroscopy is a widely used tool for nanoscale materials characterization, yet weak spectral features are often obscured by strong and spatially variable background signals. This challenge is particularly severe in interfacial and low-dimensional systems, where dominant substrate responses make conventional reference-based subtraction unreliable. Here, we introduce a transformer-based deep learning framework for reference-free spectral unmixing that reconstructs substrate contributions directly from partially observed spectra. By exploiting self-attention mechanisms to capture non-local spectral correlations, the model learns complex background signatures without requiring dedicated reference measurements. Subtraction of the reconstructed background enables the recovery of weak, previously inaccessible spectral features. We demonstrate the approach on buffer layer graphene grown on silicon carbide (SiC), a prototypical background-dominated system, where the model reveals vibrational signatures of the buffer layer otherwise hidden by the substrate response. The extracted features are validated against ab initio calculations, confirming their physical origin. Beyond this specific case, the framework provides a generalizable strategy for robust, automated spectral unmixing, compatible with real-time acquisition and closed-loop, Artificial Intelligence assisted experimental workflows.

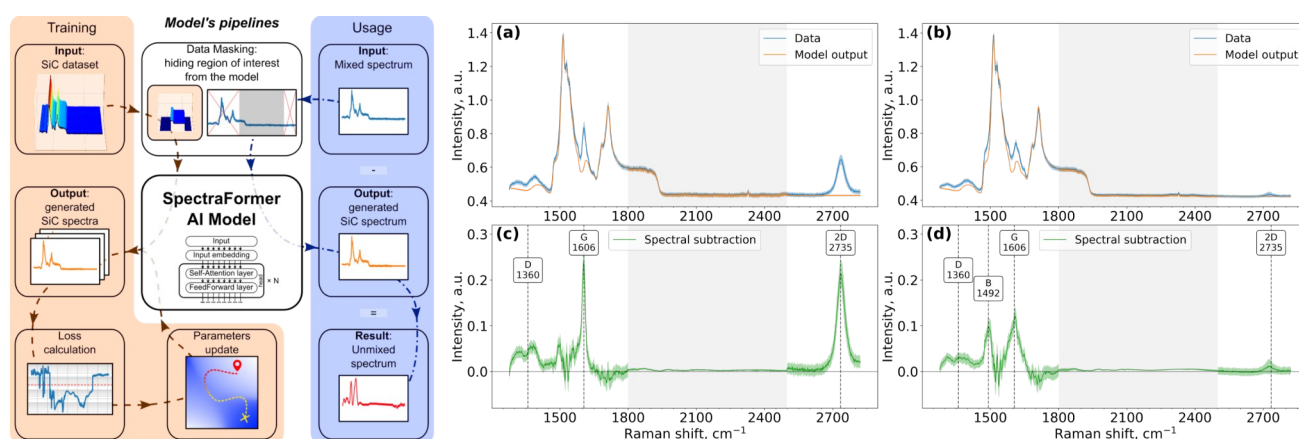


**Figure 1:** Model's pipelines for training and usage cases, and examples of model usage on monolayer graphene on SiC and buffer layer graphene on SiC. The model successfully reconstructs only the substrate signal, which enables unmixing.