

RESEARCH ARTICLE OPEN ACCESS

Revealing Hidden Raman Signatures Through Attention-Based Spectral Unmixing

Dmitriy A. Poteryayev^{1,2}  | Pietro Novelli³  | Annalisa Coriolano^{1,3}  | Riccardo Dettori⁴  | Valentina Tozzini⁵  | Fabio Beltram²  | Massimiliano Pontil³  | Antonio Rossi¹  | Stiven Forti¹  | Camilla Coletti¹ 

¹Center for Nanotechnology Innovation @NEST, Istituto Italiano di Tecnologia, Pisa, Italy | ²NEST Laboratory, Scuola Normale Superiore, Pisa, Italy | ³Computational Statistics and Machine Learning, Istituto Italiano di Tecnologia, Genova, Italy | ⁴Department of Physics, University of Cagliari, Cagliari, Italy | ⁵Istituto Nanoscienze of the National Research Council, NEST Lab-SNS, Pisa, Italy

Correspondence: Antonio Rossi (antonio.rossi@iit.it) | Stiven Forti (stiven.forti@iit.it) | Camilla Coletti (camilla.coletti@iit.it)

Received: 24 June 2026 | **Revised:** 3 August 2026 | **Accepted:** 26 August 2026

Keywords: ab initio | artificial intelligence | deep learning | graphene | Raman spectroscopy | silicon carbide | spectral line | subtraction | transformer

ABSTRACT

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 nonlocal 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, 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.

1 | Introduction

Quantitative spectral analysis is a cornerstone of modern materials characterization methods, yet the extraction of weak signals from complex, stochastically variable backgrounds remains a fundamental challenge. In techniques such as Raman spectroscopy, the physical signature of a studied material is often obscured by dominant substrate responses, fluorescence, or environmental noise. Traditional unmixing methods typically rely on the subtraction of a dedicated reference spectrum or various sophisticated background-fitted curves that may not be physics-driven. However, these approaches are frequently compromised by minor experimental fluctuations in focus, laser fluence, or local substrate

heterogeneity. This leads to artifacts and hinders the reliability of automated or real-time analysis.

The emergence of artificial intelligence (AI) offers a data-driven strategy to overcome these limitations. While machine learning and deep learning frameworks have been increasingly applied to Raman spectroscopy for qualitative spectral classification, material identification, and pattern recognition [1, 2], leveraging data-driven architectures for reference-free quantitative signal unmixing remains a far less explored challenge. Specifically, transformer architectures [3], originally designed for natural language processing [4], have demonstrated a unique capacity for sequence modeling by leveraging the self-attention mechanism.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Advanced Intelligent Systems* published by Wiley-VCH GmbH.

Unlike traditional convolution or linear models, transformers can identify nonlocal correlations across the entire spectral domain. This makes them particularly well-suited for spectroscopic problems in which weak physical signatures are embedded across complex and variable spectral backgrounds.

The key advantage of the attention mechanism over principal component analysis (PCA) lies in how they construct and utilize high-dimensional spaces to capture relationships. While both methods operate in high-dimensional environments, they do so with entirely different conceptual foundations. In the attention mechanism, data exists within a learned embedding space. The core foundation of this space is the semantic meaning and context. Tokens (such as words or pieces of data) are mapped as dense vectors, where vectors that are geometrically close to one another share similar meanings or contextual usage. Within this embedding space, the model uses learned Query, Key, and Value matrices. These matrices dynamically project the existing embeddings to calculate the real-time, context-dependent relevance between all tokens in a sequence. In contrast, PCA projects data into a space defined by principal components. The core foundation of this space is statistical variance, not semantics. PCA analyzes the original features of a dataset and constructs a new set of orthogonal (perpendicular) axes. These axes are mathematically ordered to capture the maximum possible spread of the data, completely independent of any sequential or contextual relationships between the data points.

This lack of contextual awareness makes static methods like PCA not suited for complex physical systems where subtle features are deeply intertwined with a dominant background. A particularly challenging example of this scenario is represented by Raman spectroscopy of graphene-related interfaces on silicon carbide (SiC), where weak interfacial vibrational features are buried under the substrate response. Graphene grown by thermal decomposition of SiC provides a scalable platform for graphene electronics and quantum metrology [5–7]. During growth, silicon sublimation from the SiC(0001) surface [8] leaves behind an interfacial reconstruction known as buffer layer or zero-layer graphene (ZLG), a layer where carbon atoms are arranged in a honeycomb lattice while being partially covalently bonded to the SiC substrate [9–11]. Unlike monolayer graphene (MLG), ZLG exhibits an electronic bandgap of 0.4–0.6 eV [12], making it particularly attractive for electronic and optoelectronic applications. The recent demonstration of its excellent transport properties [13] further highlights the growing interest in ZLG for scalable electronics. Also, ZLG has emerged as a platform for stabilizing confined 2D materials, including 2D gold [14] and other 2D metals [15].

Despite its potential, the controlled synthesis of high-quality ZLG still remains challenging. Optimization of ZLG growth and the realization of homogeneous layers critically depend on the availability of a straightforward and reliable characterization method. Raman spectroscopy would be an ideal metrological tool due to its speed, robustness, and ease of use [16], as well as its compatibility with closed-loop AI-assisted growth optimization workflows [17]. However, the vibrational fingerprint of ZLG is obscured by the dominant SiC background. As a consequence, Raman-based identification of ZLG often relies on weak and indirectly extracted spectral features [18], making the analysis slow, operator-dependent, and poorly suited for high-throughput

or real-time characterization workflows. This limitation is particularly critical in the context of machine learning-driven approaches for autonomous materials optimization [19], where rapid and reliable spectroscopic feedback is essential.

In this work, we introduce SpectraFormer, a transformer-based AI model framework for reference-free spectral unmixing of the ZLG Raman signal from the SiC background. The model treats Raman spectra as an ordered sequence of intensity values and is trained to reconstruct the SiC response directly from partially masked spectra and to ultimately extract ZLG signatures. We compare the unmixed ZLG spectra with *ab initio* calculations based on a high-fidelity supercell model, obtaining excellent agreement and thus validating the physical reliability of the proposed approach. In addition, SpectraFormer reveals physically meaningful vibrational features overlooked in previous theoretical predictions [20]. This approach offers a new route for real-time and robust Raman analysis, bypassing the limitations of conventional reference-based subtraction and enabling integration into closed-loop AI-driven synthesis workflows. More broadly, our results illustrate how the tight interplay between machine learning, experimental data, and theoretical simulations can unlock new capabilities in materials characterization [21].

2 | Results

To clarify the physical origin of the spectral unmixing problem addressed in this work, we first consider the structural and spectroscopic evolution of graphene grown on SiC during thermal decomposition. Figure 1a presents a schematic ball-and-stick representation of the investigated system: a 4H-SiC(0001) substrate is deprived of its Si atoms by thermal annealing. In particular, when three SiC bilayers are decomposed, the excess carbon atoms rearrange on the surface and form the ZLG. By further annealing at higher temperatures, deeper layers of the substrate lose their Si atoms, and a new ZLG forms, transforming the former one into monolayer graphene. The latter one is now only weakly interacting with the substrate, represented by the ZLG/SiC(0001), and the topmost monolayer therefore develops all the electronic properties typical of graphene. In the ZLG, instead, one carbon atom out of three is covalently bound to the substrate. Hence, the interatomic hopping between carbon atoms is strongly suppressed, and so are the π -bands, which indeed are not developed in the ZLG [22]. The vibrational properties of the MLG and ZLG are substantially different as well.

Representative Raman spectra are acquired from three distinct sample configurations: bare SiC, ZLG/SiC, and MLG/ZLG/SiC (Figure 1b). In all cases, the Raman response is dominated by the intense second-order features of the SiC substrate, which extend over a broad spectral range and mask weaker carbon layer contributions. As the ZLG forms, additional spectral components emerge; however, their direct identification remains challenging due to the overwhelming SiC signal and the partial overlap of vibrational modes. Figure 1c,d focus on spectral regions where carbon layer contributions are most apparent. With a graphene monolayer, the characteristic Raman 2D peak, a second-order double-resonant vibrational mode of graphene, remains clearly observable. This enables straightforward identification of MLG

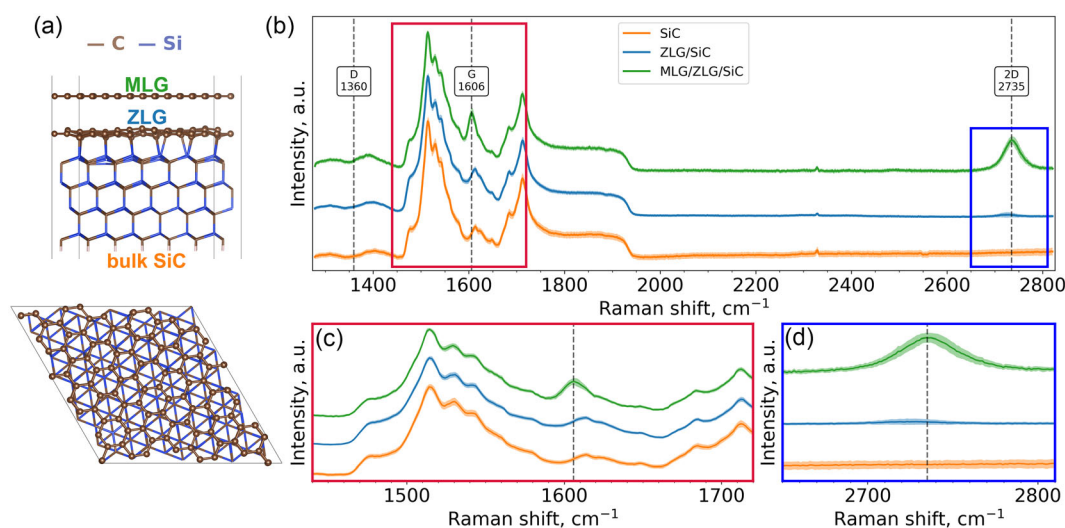


FIGURE 1 | (a) Ball-and-stick atomic structure model of investigated materials (SiC substrate, ZLG and MLG); (b) Raman spectroscopy data of three different samples: bare SiC substrate, SiC substrate with ZLG, and SiC substrate with both ZLG and free standing MLG (each solid line is the mean across the dataset for the corresponding sample, the shaded area around each solid line is the $\pm$ standard deviation). The insets (c,d) are taken from the corresponding color-coded rectangles in panel (b). They highlight the spectral ranges where sample-specific features emerge. Measurement conditions are consistent for all represented spectra: 5 mW nominal laser power, 5 s acquisition time, 1 accumulation per spectrum.

despite the dominant substrate signal overlapping with the D and G peak positions, which are the defect-activated six-atom ring breathing mode and E_{2g} phonon mode, respectively [16]. In contrast, for the ZLG, only a weak residual 2D Raman mode-like signal can be detected, which is attributed to the small monolayer graphene inclusions [23]. The vibrational signatures of the ZLG are instead masked by the dominant and broadband Raman response of the SiC substrate. Importantly, this masking effect is not fixed but varies between measurements as a result of slight changes in experimental conditions, such as focus, alignment, and acquisition parameters, as well as local crystalline defects, to which the SiC band is associated in that spectral region [24]. As a consequence, reference-based background subtraction or local fitting strategies become unreliable and difficult to apply in a systematic or automated manner. Resolving the substrate signal overlap requires an analysis strategy that does not rely on a predefined reference spectrum and does not assume that spectral information can be extracted from local regions alone. Instead, the relevant information is distributed across the entire Raman spectrum and across many measurements and must be inferred from their collective statistical structure.

Transformer-based models treat each Raman spectrum as an ordered sequence of intensity values to quantify how variations at one Raman shift correlate with all others. In practice, this is achieved by first mapping the chosen raw spectral intensities into a learned representation space, which allows the model to encode complex spectral patterns beyond simple peak positions or amplitudes. The model then applies the self-attention mechanism, which quantifies how variations in intensity at one Raman shift are correlated with variations at all other Raman shifts within the same spectrum. Through multiple stacked layers of this mechanism, the model progressively builds a global representation of the substrate Raman response, capturing both strong features and subtle correlations that are not apparent from local inspection. By exploiting these nonlocal correlations, the transformer is able to infer the most probable contribution of the SiC substrate

even in spectral regions that are intentionally masked out. The ability to infer masked content fundamentally distinguishes the approach from local interpolation, polynomial fitting, or convolution-based methods, which primarily rely on neighboring spectral points. As a result, transformer-based models are particularly effective at disentangling weak, overlapping interfacial contributions embedded within a strong and structured background, enabling robust spectral unmixing under variable experimental conditions [25].

The SpectraFormer training and inference pipelines are schematically illustrated in Figure 2. The two pipelines share the same forward pass through the transformer architecture but differ in whether model parameters are updated. Optimal beginning model architecture parameters that were found during several training sessions are listed in Table S1. During training, batches of experimentally acquired SiC Raman spectra are first partially masked along the Raman shift axis. This masking strategy prevents the model from trivially reproducing the input signal and instead forces it to infer the missing spectral content from the remaining visible portions of the spectrum. The masked spectra are then passed through the SpectraFormer model, which generates a reconstructed SiC spectrum over the full Raman shift range. The predicted output is compared to the original, unmasked spectrum to compute the loss function. The loss is evaluated across both the Raman shift and sample index dimensions, ensuring that the model learns not only individual spectral features but also their statistical variability across the full dataset (see Figure S5). For clarity, Figure 2 illustrates the loss calculation along the Raman shift axis for a single spectrum.

A mean Gamma deviance loss function is used for model optimization, as it is well suited for positive-valued spectroscopic signals and facilitates convergence toward a lower minimum on the loss hyperplane [26] (the full loss aggregation procedure is described in the Supporting Information, Note 2, along with the training curve on Figure S6 and intermediate outputs on Figure S7):

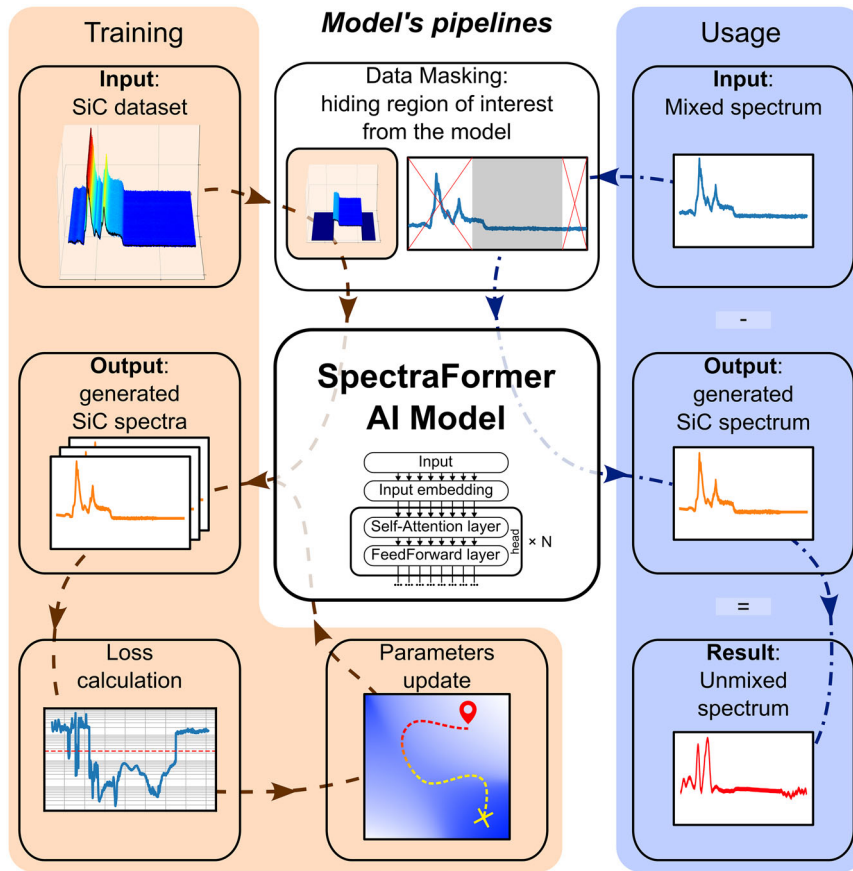


FIGURE 2 | Model's pipelines for training and usage cases. For both approaches, the model generates a bare SiC substrate signal reconstruction based on masked input, but it differs in the input type (for training—SiC spectra, for usage—mixed spectra) and whether model parameters are being updated or not.

$$L = \sum_{i=1}^N \frac{[L_i(I_i^{\text{true}}, I_i^{\text{prediction}})]}{N} = \left(\frac{I_i^{\text{true}}}{I_i^{\text{prediction}}} - 1 \right) - \ln \left(\frac{I_i^{\text{true}}}{I_i^{\text{prediction}}} \right) \quad (1)$$

Transformer-based unmixing applied to experimental spectra reveals carbon layer features by capturing the substrate contribution with high accuracy. Figure 3 shows the result of transformer-based unmixing applied to Raman spectra of graphene-related layers on SiC and demonstrates that this approach enables a controlled isolation of the SiC contribution, thereby exposing spectral features that are obscured in conventional measurements. The model application to SiC data only shows no prominent feature (see Figure S8).

Reconstructed SiC spectra are consistent over the entire Raman shift range for both sample configurations (Figure 3a for MLG/ZLG/SiC and Figure 3b for ZLG/SiC), indicating that the model captures the substrate contribution independently of the overlying carbon layers. After the subtraction of the reconstructed SiC signal, the residual spectra contain only non-SiC contributions. Once the dominant substrate background is removed, spectral features associated with carbon layers become clearly visible, most prominently in the low Raman shift region. These features are consistent with Raman signatures previously predicted to belong to the ZLG [18, 20] and are revealed without manual background modeling or peak selection. In particular, the most prominent of these modes

are known as B and L, which have been assigned to benzene-like modes and shorter bond-localized vibrations, respectively [20]. While they are known, other peaks are also observed in our measurement and are better discussed in the following section.

3 | Discussion

The unmixing procedure reveals vibrational features previously obscured by the background (Figure 3). This process avoids the need for the reference measurement or manual background modeling. The SpectraFormer reconstructs the bulk SiC Raman response using only the spectral region between 1800 and 2500 cm^{-1} . Spectral unmixing is then performed by subtracting the reconstructed SiC spectrum from the measured mixed spectrum, thereby making it possible to isolate the Raman contributions arising from the carbon layers.

3.1 | Bulk Signal Reconstruction Method Applied to MLG/ZLG/SiC and ZLG/SiC Mixed Systems

A direct comparison of the two configurations provides internal validation of the unmixing. In the ZLG/SiC case (Figure 3d), the residual spectrum shows a weak graphene 2D peak that is attributed to a minor signal originating from the initial stages of

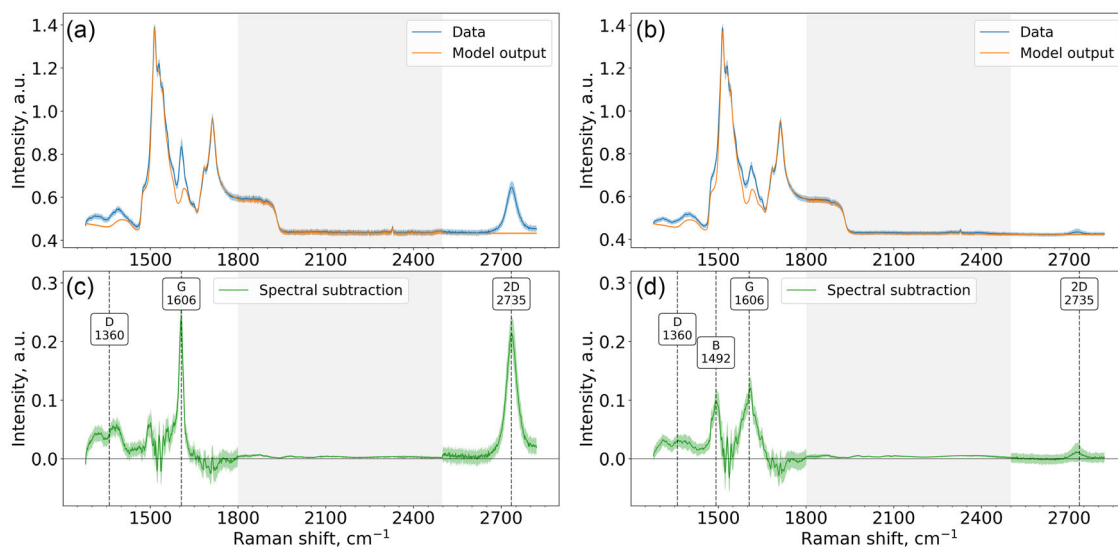


FIGURE 3 | Model performance after the training with different inputs (each solid line is the mean across the dataset for the corresponding sample, the shaded area around each solid line is the $\pm$ standard deviation): Raman spectra of (a,c) MLG/ZLG/SiC and (b,d) ZLG/SiC samples are provided, allowing to reveal targeted features by subtraction of generated SiC spectra; middle gray shaded region from 1800 to 2500 cm^{-1} is the region of data available to the model. Intensity values at (c,d) can be interpreted as normalized to the SiC peak at 1514 cm^{-1} , while those at (a,b) are also shifted by +0.4 a.u. (for details of data preprocessing, see Supporting Information, Note 2).

graphene growth, consistent with the absence of a fully developed MLG. In contrast, clear, well-known, and expected G and 2D [16] contributions appear in the MLG/ZLG/SiC case (Figure 3c), reflecting the presence of MLG. The relative intensity of 2D and G features differs from the expected pristine undoped graphene reference, where the I_{2D}/I_G is 3 times higher, and their position is redshifted with respect to our measurement [27]. However, this behavior is consistent with the strong charge doping known for epitaxial graphene on SiC [17], for which these values can vary over a wide range [27]. Noticeably, G and 2D peaks of the MLG are recovered only when a graphene monolayer is present, while the absence of a 2D peak in the ZLG-only case demonstrates that the model does not artificially introduce MLG-associated features.

3.2 | Revealed ZLG Features Physical Origin Validation With Ab Initio Calculations

To validate the physical origin of the spectral features revealed by SpectraFormer, we performed density functional theory (DFT)-based lattice dynamics calculations of the ZLG/SiC interface (Figure 4). Following the approach recently introduced by Rattke and Lazzeri [20], we extracted the one-phonon Raman response of the ZLG using a Raman-like proxy obtained by projecting the $\mathbf{q} = 0$ vibrational modes onto the Raman-active graphene E_{2g} mode ($\omega_{\text{DFT}} = 1538.2 \text{ cm}^{-1}$). The calculated spectrum reproduces the main experimentally observed B, L, and G features, with good agreement between predicted and measured peak positions (Table S2). The small residual deviations are consistent with the known tendency of semilocal GGA functionals to underestimate high-frequency C—C stretching modes [28, 29]. These results confirm that the spectral components recovered by SpectraFormer correspond to genuine vibrational signatures of the ZLG.

Because the Raman-like intensity is obtained from the projection of ZLG vibrational modes onto the graphene E_{2g} mode, the calculated spectrum can be decomposed into local atomic contributions. The resulting real-space maps (Figure 4b) reveal the microscopic origin of the recovered B, L, and G features. The B peak is associated with collective vibrations of the superperiodic crest network, the L peak originates from localized strained sp^2 regions, and the G contribution is distributed across the extended sp^2 network, reflecting a more canonical in-plane stretching character [20].

Unlike the B/L/G features, the D band cannot be captured within a purely Γ -point Raman-like description because it originates from disorder-activated finite- $\mathbf{q}$ phonons [30–32]. We therefore analyzed the full vibrational density of states (vDOS), which is dominated by buffer-layer vibrations above 1000 cm^{-1} (Figures S9 and S10). In the D-band region, the calculations reproduce a lower-frequency D_1 component and a broader D_2 contribution that can be resolved into multiple subcomponents (D_{2a} – D_{2c}), in good agreement with the experiment (Table S2). Real-space atom-resolved maps (Figure 4c) indicate that D_1 arises from more extended buffer-layer vibrations, whereas the higher-frequency D_2 components become progressively more localized and are associated with strained environments near tile boundaries and Si-bonded regions. The combined analysis supports the assignment of the experimentally recovered features to distinct vibrational motifs of the ZLG/SiC interface [16, 30, 31, 33].

Overall, our results demonstrate that transformer-based spectral unmixing provides reliable access to weak vibrational signatures that are otherwise obscured by dominant substrate backgrounds. By reconstructing and removing the SiC contribution directly from experimental data, SpectraFormer reveals spectral features that belong solely to the MLG and ZLG. The more elusive vibrational modes of ZLG are further validated through first-principles calculations. The agreement between the experiment

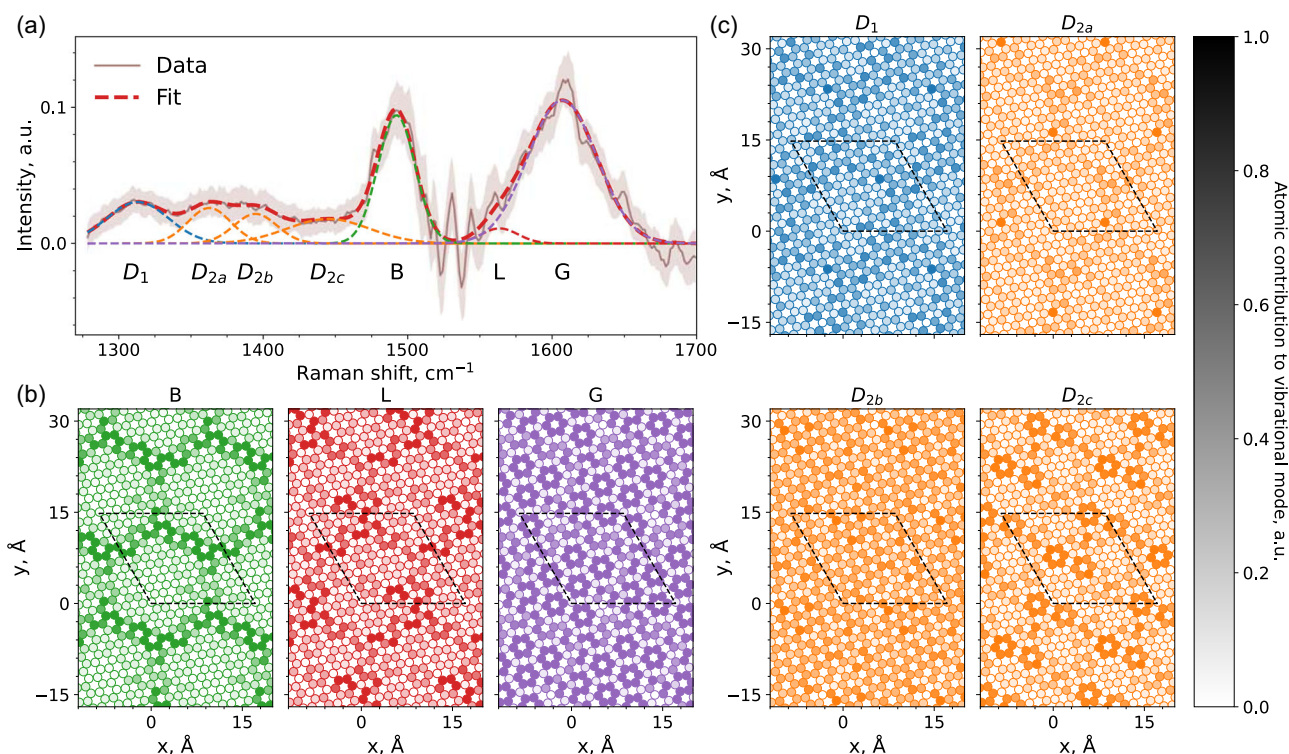


FIGURE 4 | (a) Zoomed region of experimental data in Figure 3d with fit (each solid line is the mean across the dataset for the corresponding sample, the shaded area around each solid line is the $\pm$ standard deviation), where each component is taken from ab initio calculations; (b,c) Spatial maps of atomic contribution to each of vibrational modes used (with shared gray color bar).

and theory confirms the physical origin of the recovered signals and resolves long-standing ambiguities in the Raman interpretation of the ZLG/SiC interface. Experimental comparisons with representative classical background subtraction methods widely used in chemometric analyses of vibrational spectroscopy [34–36] background subtraction methods, including direct averaged reference spectrum subtraction, Non-negative Matrix Factorization (NMF), Principal Component Analysis (PCA), and Partial Least Squares (PLS) regression, are provided in the Supplementary Information. Among the evaluated classical methods, SpectraFormer consistently achieved the lowest reconstruction error in the shown silent region (1800–2500 cm^{-1}), indicating more accurate background removal while preserving the sample vibrational features.

4 | Conclusions

This work demonstrates that transformer-based spectral reconstruction provides a robust and generalizable solution to the problem of substrate-dominated spectroscopy. By decoupling spectral analysis from the need for explicit reference measurements, the model enables the consistent extraction of weak features under variable experimental conditions, addressing a key limitation of conventional approaches. While demonstrated here for Raman spectroscopy of ZLG on SiC, the underlying framework is broadly applicable to other spectroscopic modalities and material systems in which structured backgrounds obscure relevant signals, including heterostructures, interfacial layers, and complex composite materials. The physical reliability of the extracted signals is supported by atomistic simulations, which

enable direct assignment of the recovered spectral features and clarify their microscopic origin. This combined data-driven and first-principles approach not only validates the method but also provides access to vibrational signatures that are otherwise experimentally inaccessible, thereby enabling a more complete understanding of the underlying physics of the system. Importantly, the method is inherently compatible with automated and real-time data acquisition as it operates directly on measured spectra without additional calibration steps. This opens the door to integration within closed-loop, AI-driven experimental platforms, where rapid and reliable feedback from spectroscopic measurements is essential. More generally, the combination of attention-based models with physically grounded validation provides a pathway toward data-driven yet interpretable characterization tools, enabling more efficient exploration, optimization, and control of nanoscale material systems.

5 | Materials and Methods

5.1 | Materials Growth and Characterization

ZLG is grown through thermal decomposition, starting from a 6H-SiC substrate. Commercial SiC wafers are first thoroughly cleaned using acetone and isopropyl alcohol in an ultrasonic bath and then subjected to oxygen plasma treatment, followed by Piranha and HF acidic baths to eliminate any organic residues from the surface. Subsequently, the SiC(0001) wafers are exposed to high-temperature hydrogen gas inside a furnace to remove polishing scratches, a procedure known as hydrogen etching. Once

this step is completed, the ZLG growth process begins. All growth procedures are conducted using an AIXTRON Black Magic cold-wall reactor in an Ar environment.

Samples of SiC with the grown ZLG were further characterized using X-ray photoemission spectroscopy (XPS), confirming the presence of ZLG-related features (for details, see Supporting Information, Note 1, Figure S1). XPS spectra are obtained with a SPECS XR-50 Al $K\alpha$ X-ray source, AstraioS190 analyser, entrance slit 1.5×3 mm, pass energy 10 eV, and 2D CMOS detector.

All Raman spectroscopy data were acquired using a Renishaw inVia micro-Raman spectrometer. Since the region of interest for ZLG features is located in the range $(1000\text{--}3000)\text{ cm}^{-1}$, datasets were acquired on both 6H and 4H-SiC(0001) polytypes in that range centered at different positions (centers were chosen in the range from 1800 to 2300 cm^{-1} with 100 cm^{-1} step) with a span of $1500\text{--}1600\text{ cm}^{-1}$ for each spectrum (1015 data points per spectrum). We used a $100\times$ objective with 0.85 NA, a 532 nm excitation laser, and a 1800 lines/mm grating. To achieve a balanced training dataset and broaden the model's spectral knowledge, different measurement parameters were used. In particular, different combinations of acquisition time, nominal laser power, and number of accumulations per spectrum (Figure S2).

5.2 | Transformer Model Training

Model training was conducted on 4 NVIDIA Tesla V100 16Gb GPUs using the Franklin HPC infrastructure of IIT. Model was trained on arithmetically preprocessed Raman spectroscopy data (preprocess pipeline is described in Supporting Information, Note 2). Data preprocessing was required to remove spectroscopic outliers and establish robustness across the computational pipeline. After preprocessing, data kept its statistical properties (see Figures S3 and S4).

5.3 | Ab Initio Calculations

We model the ZLG/SiC interface using the reduced commensurate supercell often denoted as the $\sqrt{31} \times \sqrt{31}$ R8.95° reconstruction of the Si-terminated SiC surface, which corresponds to a 7×7 graphene-like overlayer rotated by 21.787° . In this work, we adopt the same structure considered in [11, 37]: the *Sh* stacking, characterized by the presence of at least one “hollow” site (a surface Si atom located beneath the center of a buffer hexagon). The initial configuration is generated by placing a flat hexagonal carbon layer above the Si-terminated surface at a separation slightly larger than the expected Si–C bond length and then fully relaxing atomic positions (see Computational details below). The reduced supercell enables calculations at significantly lower computational cost than the larger $6\sqrt{3}$ -based model, while retaining the key local structural motifs and registry patterns relevant to the ZLG. The substrate is modeled as a four-layer slab of cubic SiC, with the bottom surface passivated by hydrogen atoms to remove spurious dangling-bond states.

All calculations are performed within DFT, with the same setup adopted in previous works [11, 38] using ultrasoft RKKJ

pseudopotentials [39] and the PBE exchange-correlation functional [40]. Dispersion interactions are included through the semi-empirical Grimme D2 correction (PBE-D2) [41]. The plane-wave kinetic energy cutoff is set to 30 Ry, with a charge-density cutoff of 300 Ry, and self-consistency is converged to 10^{-8} .

Structural relaxations are carried out with a quasi-Newton BFGS optimizer [42] using standard convergence thresholds of 10^{-3} a.u. on forces and 10^{-4} a.u. on total energy. Geometry optimizations and SCF calculations to extract the forces from the atomic displacements were performed at the Γ point. A Gaussian smearing of 0.01 Ry was used throughout.

The in-plane lattice vectors are fixed to those of a relaxed SiC slab, while the out-of-plane supercell length is set to $31.8\text{ \AA}$ to avoid interactions between periodic replicas. Calculations were performed with QUANTUM ESPRESSO [43] (version 7.2.0).

Interatomic force constants and vibrational modes have been obtained using the ALAMODE [44] code within the frozen-phonon approach. Since the ZLG hosts in-gap and near-gap electronic states that are spatially localized on specific atomic motifs (crest regions, benzene-like patches, etc.), we limited atomic displacements to these regions while keeping the remainder of the simulation cell fixed to reduce the computational cost. See Supporting Information, Note 4, for more details.

Author Contributions

A.R., S.F., and C.C. conceived the project. A.R. supervised the execution of the work. D.A.P. developed and trained the transformer model, analyzed the data, and wrote the manuscript. P.N. implemented the deep learning architecture. A.C. and D.A.P. performed the Raman measurements. R.D. and V.T. carried out the DFT and Raman mode calculations. F.B., M.P., and C.C. provided financial support. All authors discussed the results and contributed to the editing and revision of the manuscript.

Acknowledgments

We acknowledge the project PNRR MUR Project PE000013 CUP J53C22003010006 Future Artificial Intelligence Research (FAIR) and PNRR MUR Project PE0000023—National Institute of Quantum Science and Technology (NQSTI) funded by the European Union—NextGenerationEU.

We acknowledge that the research activity herein was carried out using the IIT Franklin HPC infrastructure; we gratefully acknowledge the Data Science and Computation Facility and its Support Team for their support and assistance on the IIT High Performance Computing Infrastructure.

We acknowledge the CINECA award under the ISCRA initiative, for the availability of high-performance computing resources and support under project IsCd0_SPHERE HP10CSHBY3.

Open access publishing facilitated by Istituto Italiano di Tecnologia, as part of the Wiley - CRUI-CARE agreement.

Funding

This study was supported by PNRR MUR (Grants PE000013 and PE0000023), and CINECA (Grant IsCd0_SPHERE HP10CSHBY3).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data used for SpectraFormer model training process are available at [<https://huggingface.co/dpoteryayev/SiC-high-f>] [45].

Code Availability

The code for SpectraFormer model, along with detailed instructions, is available at [<https://github.com/pietronvil/SpectraFormer>].

References

1. J. Qi, Y. Liu, D. Wang, X. Liu, and J. Duan, "Machine Learning-Enhanced Raman Spectroscopy for Rapid Identification of Aviation Coating Damage," *Corrosion Science* 260 (2026): 113584, <https://doi.org/10.1016/j.corsci.2025.113584>.
2. Y. Liu, Y. Wu, J. Wang, J. Qi, C. Zhou, and Y. Xue, "Recent Advances in Raman Spectral Classification With Machine Learning," *Sensors* 26 (2026): 341, <https://doi.org/10.3390/s26010341>.
3. R. E. Turner, "An Introduction to Transformers," (2023), <https://doi.org/10.48550/arXiv.2304.10557>.
4. A. Vaswani, N. Shazeer, N. Parmar, et al., "Attention Is All You Need," (2023), <https://doi.org/10.48550/arXiv.1706.03762>.
5. A. Tzalenchuk, S. Lara-Avila, A. Kalaboukhov, et al., "Towards a Quantum Resistance Standard Based on Epitaxial Graphene," *Nature Nanotechnology* 5 (2010): 186–189, <https://doi.org/10.1038/nnano.2009.474>.
6. T. J. B. M. Janssen, A. Tzalenchuk, S. Lara-Avila, S. Kubatkin, and V. I. Fal'ko, "Quantum Resistance Metrology Using Graphene," *Reports on Progress in Physics* 76 (2013): 104501, <https://doi.org/10.1088/0034-4885/76/10/104501>.
7. X. Wan, X. Fan, C. Zhai, et al., "Contrasting Transport Performance of Electron and Hole-Doped Epitaxial Graphene for Quantum Resistance Metrology," *Chinese Physics Letters* 40 (2023): 107201, <https://doi.org/10.1088/0256-307X/40/10/107201>.
8. A. J. Van Bommel, J. E. Crombeen, and A. Van Tooren, "LEED and Auger Electron Observations of the SiC(0001) Surface," *Surface Science* 48 (1975): 463–472, [https://doi.org/10.1016/0039-6028\(75\)90419-7](https://doi.org/10.1016/0039-6028(75)90419-7).
9. K. V. Emtsev, A. Bostwick, K. Horn, et al., "Towards Wafer-Size Graphene Layers by Atmospheric Pressure Graphitization of Silicon Carbide," *Nature Materials* 8 (2009): 203–207, <https://doi.org/10.1038/nmat2382>.
10. S. Goler, C. Coletti, V. Piazza, et al., "Revealing the Atomic Structure of the Buffer Layer Between SiC(0001) and Epitaxial Graphene," *Carbon* 51 (2013): 249–254, <https://doi.org/10.1016/j.carbon.2012.08.050>.
11. T. Cavallucci and V. Tozzini, "Intrinsic Structural and Electronic Properties of the Buffer Layer on Silicon Carbide Unraveled by Density Functional Theory," *Scientific Reports* 8 (2018): 13097, <https://doi.org/10.1038/s41598-018-31490-7>.
12. M. N. Nair, I. Palacio, A. Celis, et al., "Band Gap Opening Induced by the Structural Periodicity in Epitaxial Graphene Buffer Layer," *Nano Letters* 17 (2017): 2681–2689, <https://doi.org/10.1021/acs.nanolett.7b00509>.
13. J. Zhao, P. Ji, Y. Li, et al., "Ultrahigh-Mobility Semiconducting Epitaxial Graphene on Silicon Carbide," *Nature* 625 (2024): 60–65, <https://doi.org/10.1038/s41586-023-06811-0>.
14. S. Forti, S. Link, A. Stöhr, et al., "Semiconductor to Metal Transition in Two-Dimensional Gold and Its Van der Waals Heterostack With Graphene," *Nature Communications* 11 (2020): 2236, <https://doi.org/10.1038/s41467-020-15683-1>.
15. N. Briggs, B. Bersch, Y. Wang, et al., "Atomically Thin Halfvan der Waals Metals Enabled by Confinement Heteroepitaxy," *Nature Materials* 19 (2020): 637–643, <https://doi.org/10.1038/s41563-020-0631-x>.
16. A. C. Ferrari and D. M. Basko, "Raman Spectroscopy as a Versatile Tool for Studying the Properties of Graphene," *Nature Nanotechnology* 8 (2013): 235–246, <https://doi.org/10.1038/nnano.2013.46>.
17. L. Sabattini, A. Coriolano, C. Casert, et al., "Towards AI-Driven Autonomous Growth of 2D Materials Based on a Graphene Case Study," *Communications Physics* 8 (2025): 180, <https://doi.org/10.1038/s42005-025-02086-1>.
18. F. Fromm, M. H. Oliveira Jr, A. Molina-Sánchez, et al., "Contribution of the Buffer Layer to the Raman Spectrum of Epitaxial Graphene on SiC(0001)," *New Journal of Physics* 15 (2013): 043031, <https://doi.org/10.1088/1367-2630/15/4/043031>.
19. F. Zhang, R. Tamura, F. Zeng, D. Kozawa, and R. Kitaura, "Bayesian Optimization for Controlled Chemical Vapor Deposition Growth of WS₂," *ACS Applied Materials & Interfaces* 16 (2024): 59109–59115, <https://doi.org/10.1021/acsami.4c15275>.
20. G. Radtke and M. Lazzeri, "Vibrational Properties of Epitaxial Graphene Buffer Layer on Silicon Carbide," *Physical Review B* 111 (2025): L220303, <https://doi.org/10.1103/xk69-4n46>.
21. L. Bonati, P. Novelli, and A. Rossi, "From Trial-and-Error to Intelligent Workflows: Machine Learning and Simulations for Scalable 2D Material Synthesis," *2D Materials* 13 (2025): 013001, <https://doi.org/10.1088/2053-1583/ae2692>.
22. K. V. Emtsev, F. Speck, T. Seyller, L. Ley, and J. D. Riley, "Interaction, Growth, and Ordering of Epitaxial Graphene on SiC{0001} Surfaces: A Comparative Photoelectron Spectroscopy Study," *Physical Review B* 77 (2008): 155303, <https://doi.org/10.1103/PhysRevB.77.155303>.
23. S. Fiori, Y. Murata, S. Veronesi, A. Rossi, C. Coletti, and S. Heun, "Li-Intercalated Graphene on SiC(0001): An STM Study," *Physical Review B* 96 (2017): 125429, <https://doi.org/10.1103/PhysRevB.96.125429>.
24. P. Wang and S. Wang, "Raman Scattering Study of Silicon Carbide Irradiated With 1.25 MeV Si Ions," *IOP Conference Series: Materials Science and Engineering* 493 (2019): 012092, <https://doi.org/10.1088/1757-899X/493/1/012092>.
25. P. Ghosh, S. K. Roy, B. Koirala, B. Rasti, and P. Scheunders, "Hyperspectral Unmixing Using Transformer Network," *IEEE Transactions on Geoscience and Remote Sensing* 60 (2022): 1–16, <https://doi.org/10.1109/TGRS.2022.3196057>.
26. P. McCullagh and J. Nelder, *Generalized Linear Models*, 2nd ed. (Chapman and Hall, 1989), 532.
27. M. Bruna, A. K. Ott, M. Ijäs, D. Yoon, U. Sassi, and A. C. Ferrari, "Doping Dependence of the Raman Spectrum of Defected Graphene," *ACS Nano* 8 (2014): 7432–7441, <https://doi.org/10.1021/nn502676g>.
28. N. Mounet and N. Marzari, "First-Principles Determination of the Structural, Vibrational and Thermodynamic Properties of Diamond, Graphite, and Derivatives," *Physical Review B* 71 (2005): 205214, <https://doi.org/10.1103/physrevb.71.205214>.
29. K. Nakano, T. Morresi, M. Casula, R. Maezono, and S. Sorella, "Atomic Forces by Quantum Monte Carlo: Application to Phonon Dispersion Calculations," *Physical Review B* 103 (2021): L121110, <https://doi.org/10.1103/physrevb.103.L121110>.
30. F. Tuinstra and J. L. Koenig, "Raman Spectrum of Graphite," *The Journal of Chemical Physics* 53 (1970): 1126–1130, <https://doi.org/10.1063/1.1674108>.
31. P. Venzuela, M. Lazzeri, and F. Mauri, "Theory of Double-Resonant Raman Spectra in Graphene: Intensity and Line Shape of Defect-Induced and Two-Phonon Bands," *Physical Review B* 84 (2011): 035433, <https://doi.org/10.1103/PhysRevB.84.035433>.

32. A. C. Ferrari and J. Robertson, "Interpretation of Raman Spectra of Disordered and Amorphous Carbon," *Physical Review B* 61 (2000): 14095–14107, <https://doi.org/10.1103/physrevb.61.14095>.
33. J. Maultzsch, S. Reich, and C. Thomsen, "Double-Resonant Raman Scattering in Graphite: Interference Effects, Selection Rules, and Phonon Dispersion," *Physical Review B* 70 (2004): 155403, <https://doi.org/10.1103/PhysRevB.70.155403>.
34. R. Luce, P. Hildebrandt, U. Kuhlmann, and J. Liesen, "Using Separable Nonnegative Matrix Factorization Techniques for the Analysis of Time-Resolved Raman Spectra," *Applied Spectroscopy* 70 (2016): 1464–1475, <https://doi.org/10.1177/0003702816662600>.
35. M. D. Peris-Díaz and A. Krężel, "A Guide to Good Practice in Chemometric Methods for Vibrational Spectroscopy, Electrochemistry, and Hyphenated Mass Spectrometry," *TrAC Trends in Analytical Chemistry* 135 (2021): 116157, <https://doi.org/10.1016/j.trac.2020.116157>.
36. M. Lasalvia, V. Capozzi, and G. Perna, "A Comparison of PCA-LDA and PLS-DA Techniques for Classification of Vibrational Spectra," *Applied Sciences* 12 (2022): 5345, <https://doi.org/10.3390/app12115345>.
37. T. Cavallucci, Y. Murata, M. Takamura, H. Hibino, S. Heun, and V. Tozzini, "Unraveling Localized States in Quasi Free Standing Monolayer Graphene by Means of Density Functional Theory," *Carbon* 130 (2018): 466–474, <https://doi.org/10.1016/j.carbon.2018.01.027>.
38. T. Cavallucci and V. Tozzini, "Multistable Rippling of Graphene on SiC: A Density Functional Theory Study," *The Journal of Physical Chemistry C* 120 (2016): 7670–7677, <https://doi.org/10.1021/acs.jpcc.6b01356>.
39. A. M. Rappe, K. M. Rabe, E. Kaxiras, and J. D. Joannopoulos, "Optimized Pseudopotentials," *Physical Review B* 41 (1990): 1227–1230, <https://doi.org/10.1103/PhysRevB.41.1227>.
40. J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized Gradient Approximation Made Simple," *Physical Review Letters* 77 (1996): 3865–3868, <https://doi.org/10.1103/PhysRevLett.77.3865>.
41. S. Grimme, "Semiempirical GGA-Type Density Functional Constructed With a Long-Range Dispersion Correction," *Journal of Computational Chemistry* 27 (2006): 1787–1799, <https://doi.org/10.1002/jcc.20495>.
42. S. R. Billeter, A. Curioni, and W. Andreoni, "Efficient Linear Scaling Geometry Optimization and Transitionstate Search for Direct Wavefunction Optimization Schemes in Density Functional Theory Using a Plane-Wave Basis," *Computational Materials Science* 27 (2003): 437–445, [https://doi.org/10.1016/S0927-0256\(03\)00043-0](https://doi.org/10.1016/S0927-0256(03)00043-0).
43. P. Giannozzi, S. Baroni, N. Bonini, et al., "QUANTUM ESPRESSO: A Modular and Opensource Software Project for Quantum Simulations of Materials," *Journal of Physics: Condensed Matter* 21 (2009): 395502, <https://doi.org/10.1088/0953-8984/21/39/395502>.
44. T. Tadano, Y. Gohda, and S. Tsuneyuki, "Anharmonic Force Constants Extracted From First-Principles Molecular Dynamics: Applications to Heat Transfer Simulations," *Journal of Physics: Condensed Matter* 26 (2014): 225402, <https://doi.org/10.1088/0953-8984/26/22/225402>.
45. D. Poteryayev, A. Coriolano, and C. Coletti, Hugging Face SpectraFormer Dataset (Revision 858c2b1) (2026), <https://doi.org/10.57967/hf/8904>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.