






RESEARCH ARTICLE | JUNE 28 2024

Retrieving genuine nonlinear Raman responses in ultrafast spectroscopy via deep learning

Giuseppe Fumero ; Giovanni Batignani ; Edoardo Cassetta; Carino Ferrante ; Stefano Giagu ; Tullio Scopigno 

 Check for updates

APL Photonics 9, 066119 (2024)

<https://doi.org/10.1063/5.0198013>



View
Online



Export
Citation



APL Energy

Latest Articles Online!

Read Now

Retrieving genuine nonlinear Raman responses in ultrafast spectroscopy via deep learning

Cite as: APL Photon. 9, 066119 (2024); doi: 10.1063/5.0198013

Submitted: 15 January 2024 • Accepted: 10 June 2024 •

Published Online: 28 June 2024



View Online



Export Citation



CrossMark

Giuseppe Fumero,^{1,a)}  Giovanni Batignani,^{1,2}  Edoardo Cassetta,¹ Carino Ferrante,^{1,3}  Stefano Giagu,¹ 
and Tullio Scopigno^{1,2,4,b)}

AFFILIATIONS

¹Dipartimento di Fisica, Sapienza Università di Roma, Roma, Italy

²Istituto Italiano di Tecnologia, Center for Life Nano Science @ Sapienza, Roma, Italy

³CNR-SPIN, c/o Dipartimento di Scienze Fisiche e Chimiche, Via Vetoio, Coppito (AQ), Italy

⁴Istituto Italiano di Tecnologia, Graphene Labs, Genova, Italy

^{a)}**Current address:** Physical Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA and Department of Physics and Astronomy, West Virginia University, Morgantown, WV 26506-6315, USA.

^{b)}**Author to whom correspondence should be addressed:** tullio.scopigno@uniroma1.it

ABSTRACT

Noise manifests ubiquitously in nonlinear spectroscopy, where multiple sources contribute to experimental signals generating interrelated unwanted components, from random point-wise fluctuations to structured baseline signals. Mitigating strategies are usually heuristic, depending on subjective biases such as the setting of parameters in data analysis algorithms and the removal order of the unwanted components. We propose a data-driven frequency-domain denoiser based on a convolutional neural network to extract authentic vibrational features from a nonlinear background in noisy spectroscopic raw data. The different spectral scales in the problem are treated in parallel by means of filters with multiple kernel sizes, which allow the receptive field of the network to adapt to the informative features in the spectra. We test our approach by retrieving asymmetric peaks in stimulated Raman spectroscopy, an ideal test-bed due to its intrinsic complex spectral features combined with a strong background signal. By using a theoretical perturbative toolbox, we efficiently train the network with simulated datasets resembling the statistical properties and lineshapes of the experimental spectra. The developed algorithm is successfully applied to experimental data to obtain noise- and background-free stimulated Raman spectra of organic molecules and prototypical heme proteins.

© 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0198013>

I. INTRODUCTION

Nonlinear optics has enabled and fostered the application of spectroscopy to ultrashort time scales. Thanks to the developments of photonic techniques for femtosecond pulse generation and shaping, nonlinear spectroscopy addressed the interdisciplinary studies of ultrafast phenomena across a wide energy range,¹ advancing, among others, the understanding of many-body interactions in structured and strongly correlated systems,^{2–4} carrier dynamics and e–ph couplings upon excitation,^{5–8} photochemical reactions, and vibronic and non-adiabatic effects in molecules.^{9–14} Similar advancements have also become possible in the closely related field of microscopy.^{15–18} A common approach consists of using multiple ultrashort pulses, shaped in frequency and time, to resolve the induced modifications to a certain optical observable in a differential

manner when one of the pulses is switched on and off. Dynamical insights can be obtained by adding a photochemical actinic pump and/or by tuning the pulses in resonance with the electronic absorption edges from which excited-state relaxation occurs. In particular, coherent resonant Raman spectroscopies exploit stimulated Raman scattering (SRS) to probe the structural response of the system undergoing ultrafast dynamics, thanks to their sensitivity to both the electronic and vibrational degrees of freedom.^{19–26}

Even if nonlinearity is pivotal for accessing ultrafast dynamics over multiple time and energy scales, it is accompanied by two major complications in the data analysis and interpretation: (1) the signal generated by the nonlinear process is usually low compared to the residual, non-interacting light or the fluctuations in the laser source, leading to point-wise noisy fluctuations in the measured spectra; and (2) multiple nonlinear processes are usually generated

by the same experimental layout, leading to the need for *a posteriori* protocols to isolate the desired spectroscopic information from an unwanted baseline background and signal distortions due to overlying competitive effects.^{27–30} Post-processing routines are usually strongly dependent on the specific sample and the experimental parameters. The sources of noise are not quantitatively known, and often the exact lineshapes of both the baseline and signal cannot be predicted in advance.^{31,32} Moreover, point-wise denoising and baseline subtraction are generally heavily correlated operations that cannot be factorized, particularly when the target clean spectrum contains asymmetric lineshapes or components with largely different relative intensities.³³ The practical solution is often delegated to the experienced eye of the spectroscopist, a strategy that hampers automatization of the routine and may lead to sub-optimal resolutions, ambiguity, and human biases.

To overcome these critical issues and enhance the sensitivity of nonlinear Raman techniques beyond the limitations induced by the background and low signal-to-noise conditions, we have devised and trained a deep neural network (NN) based on multiple convolutional layers operating in parallel for denoising and baseline removal of raw SRS spectra. By designing the network architecture and choosing a suitable loss function and optimization strategy during training, we show how to perform the two tasks in parallel and avoid the difficulties that hinder the application of standard data processing algorithms.

Applications of deep learning to nonlinear spectroscopy are still in their infancy but have already demonstrated great potential.^{34–42} NN has also been applied to the preprocessing of spontaneous Raman data.^{43–48} The lineshapes in these techniques are always positive, and the luminescence background is usually not structured. More importantly, the removal of the luminescent background and the denoising are typically tackled separately. In the case of nonlinear Raman spectroscopy, due to the asymmetric and complex lineshapes that are typical of SRS signals,^{49–51} there are no optimal methods to disentangle the two tasks and perform them sequentially. Supervised training of the algorithms is further complicated by the absence of large enough labeled datasets to ensure a statistically relevant representation of the diverse SRS baseline and peak structures. To address these challenges, we combine a multi-parallel convolutional NN architecture with supervised training built on a theoretical toolbox based on the density matrix perturbative expansion for accurate modeling of the spectroscopic signals and their characteristic noise.

II. SIGNAL AND NOISE IN STIMULATED RAMAN SCATTERING

SRS is a third order nonlinear optical effect that can be generated in the sample by the joint action of a broadband femtosecond

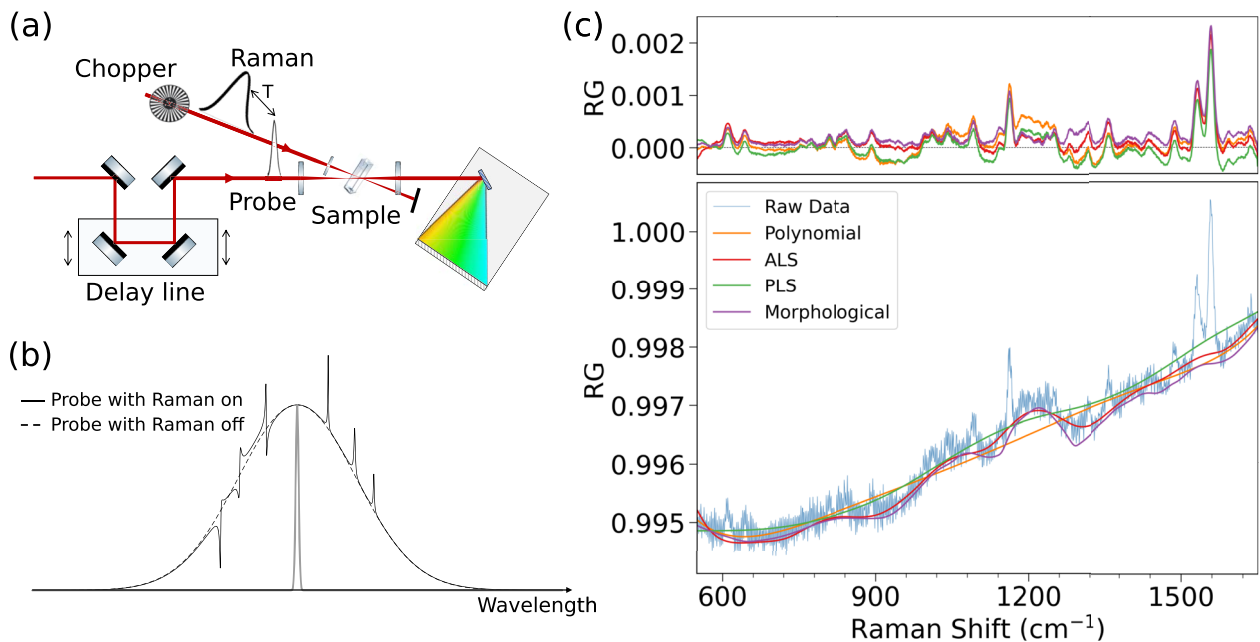


FIG. 1. (a) Experimental scheme of SRS spectroscopy. A femtosecond broadband probe is focused on the sample together with a narrowband picosecond Raman pulse. The probe is then spectrally dispersed and detected. The relative arrival time of the pulses can be varied by a delay line. A mechanical chopper allows for the differential detection of the signal with and without the Raman pulse. (b) Spectral envelope of the probe pulse detected after the sample in the presence (black line) and absence (black dashed line) of the Raman pulse (gray line). SRS features are obtained at the top of the probe spectrum when the two pulses interact with the sample. (c) Example of the ambiguities that can arise from baseline removal in SRS red side data measured on a fluorescent protein (wt-GFP). The application of different algorithms—polynomial fitting, Asymmetric Least Squares (ALSs), Penalized Least Squares (PLSs), and iterative morphological fitting—determines the baseline differently (bottom panel). This results in differences in the retrieved SRS spectra (top panel).

probe pulse (PP) and a narrowband picosecond Raman pulse (RP) overlapped in time¹⁹ and then heterodyne detected for spectroscopic applications to access the vibrational structure of the system under investigation [Fig. 1(a)]. The signal is coherently stimulated when the energy difference between the two laser pulses matches a Raman active molecular transition. Similarly to the spontaneous case, where the Raman peaks are spectrally located at the red and blue shifted sides of the excitation wavelength, the SRS signal is positively and negatively offset with respect to the RP. SRS features, however, arise as peaks, dips, or even dispersive signals generated on top of the broadband spectral envelope of the probe, which is spectrally dispersed by means of a monochromator and detected [Fig. 1(b)]. SRS data usually display the normalized difference between the PP spectra recorded after the sample with the RP switched on and off [Raman Gain (RG)] as a function of the detected frequency shift with respect to the RP (Raman shift). If the RP wavelength is tuned in resonance with the absorption of the sample, the transmission of the PP itself can be modulated even in the absence of stimulated Raman due to the fast electronic response of the sample [Transient Absorption (TA) effect].⁵² This, together with additional nonlinear processes, including but not limited to solvent effects⁵³ and cross-phase modulation,²⁸ causes the presence of an unwanted baseline, usually broader than the Raman features, which needs to be removed from the SRS raw data to extract the Raman spectra and correctly retrieve the vibrational information.

The treatment of raw SRS data are further complicated by the phase and pulse-to-pulse instabilities of the laser sources and by the noise associated with the detection process.^{54–56} This latter is typically due to the electronic fluctuations of the photogenerated carriers, the error related to the readout process, and the intrinsic shot noise limit caused by the quantum nature of light. As expected in any type of heterodyne-detected spectroscopy, these sources of noise combine with the inherent ambiguity of the baseline subtraction procedure, resulting in an overall signal-to-noise ratio that is often considerably lower than the nominal sensitivity of the technique. This is exemplified in Fig. 1(c), in which we show the outcomes of different baseline estimation techniques applied to a typical SRS spectrum of the wild-type green fluorescent protein. Without any assumption on the nonlinear dynamics and the Raman peaks, lineshapes, and spectral positions, the estimated baseline is highly dependent on the subtraction algorithm. Consequently, it is often not possible to infer any conclusive statement from such ambiguously extracted features.

III. METHODS

A. Simulated training datasets

In order to simulate SRS spectra closely resembling the data obtained by the experiments, we used a perturbative framework based on the density matrix expansion.⁵⁷ We simulated two datasets with different levels of noise, the *high noise* (HN) and *low noise* (LN) datasets. Each dataset consists of 5000 raw SRS spectra of 801 points each, with noise and baseline associated with the corresponding clean spectra, serving as ground truth (GT). 80% of the samples (4000 simulated spectra for each dataset) were used during the training phase. Additional details on the datasets are given in

Sec. I B of the [supplementary material](#). Since we adopted the SRS clean spectra as GTs, we expect the NN to preserve any asymmetry due to resonance conditions, contributions to excited state vibrations, or dynamics-induced effects in the retrieved peak lineshapes. This choice maintains the related microscopic information accessible for further analysis of the experimental data and avoids biases toward specific models. In settings where the asymmetry is known and detrimental for a specific purpose, for example, in microscopy applications, its source can be included in the simulation toolbox before training the NN, or an asymmetry-free GT dataset can be used instead while keeping the same NN architecture.

The HN and LN datasets have been simulated using nonlinear response perturbation theory. In this theoretical framework, the nonlinear signal is obtained from the n -th order nonlinear optical polarization $P^{(n)}$, which consists of the convolution between matter correlation functions and the electromagnetic fields. The radiation–matter interaction is treated perturbatively, and the density matrix is expanded in the power of the fields, applying many-body Green function techniques in the Liouville space. Diagrammatic representations are exploited to isolate all the relevant terms in the expansion and calculate the non-equilibrium expectation values of the correlation functions.^{58,59} Different sets of diagrams are associated with the SRS signal and the TA baseline,⁶⁰ which can be simulated separately (see Sec. I A of the [supplementary material](#)). Shot noise is also included by means of fluctuations scaling as the square root of the number of detected photons, and the associated uncertainty is propagated through the spectroscopic signal.⁶¹ Consequently, all three different spectral scales that are peculiar to the SRS data are present in the simulations: long-scale baseline variations, point-wise noise fluctuations, and, in between these two extrema, the vibrational features. Within each dataset, half of the samples have been generated by simulated signals on the red side of the SRS spectrum with respect to the narrowband RP and half on the blue side. All the molecular and experimental parameters have been varied randomly through the datasets and sampled from a uniform distribution within the boundaries summarized in Table S1 of the [supplementary material](#). The main difference between the two datasets consists of the different number of averaged acquisitions per sample, $N_{\text{acquisition}}$. Averaging impacts the overall noise of the sample as a multiplicative factor inversely proportional to the squared root of the number of acquisitions. Each sample in the dataset resulted from the average of a $N_{\text{acquisition}}$ noisy replica of the experiment. $N_{\text{acquisition}}$ was set to 1 and 100 for the HN and LN datasets, respectively. To avoid biases due to the selection of a particular dynamical model or pulse synthesis method, no explicit dynamics is considered, and the pulses are modeled as Gaussian temporal profiles with random duration. In particular, the probe pulse duration is kept constant since any effect due to the temporal dispersion of uncompressed probe pulses is not expected to impact the signal lineshapes or the outcome of the denoising procedure (see Sec. I B of the [supplementary material](#)). A variety of possible lineshapes and spectral properties are sampled by randomly selecting the experimental and molecular parameters. Major deviations from these parameters, such as strongly asymmetrical temporal profiles of the Raman pulses or sample-specific dynamics, can be treated by a re-training or a fine tuning of the neural network, starting from the pretrained weights provided in this work.⁶²

B. Neural network architecture and supervised training

Considering the different spectral scales present in the problem, we adopted a convolutional neural network architecture^{63–65} and combined different kernel sizes in parallel to reach different receptive fields at the same time. The general architecture of the model is shown in Fig. 2 and is based on parallel linear and nonlinear transformations consisting of zero padded convolutions with different kernel sizes and nonlinear activation functions. The network takes as an input the SRS spectra sampled at n_{input} points in frequency. The input is replicated to feed N_{kernel} parallel branches, which are composed of N_{layers} convolutional blocks each and return feature maps of different dimensions due to the different number of output channels. In particular, each block performs c_i convolutions with a different, randomly initialized filter of size k_i , followed by a Rectified Linear Unit (ReLU) activation function.⁶⁶ Both the number of filters c_i and the kernel size k_i are branch-dependent and fixed along the branch. After the last convolution block, each branch returns a feature map of dimension $b + \{c_i, (1, n_{input})\}$, where b is the size of the training batches. The features from different branches are concatenated along the channel dimension after the last convolutional block of each branch and feed a single final branch of N_{conv} convolutional blocks with ReLU activation. In the final branch, the kernel size is fixed to one, and the number of filters downscales by half each block. In such a way, we obtain a parametric linear combination of the features obtained from the different branches to be optimized by training. The last convolution has n_{input} filters so that the output channel dimension matches the input to recover feature interpretability and precedes a final residual step in which the

input is subtracted from the feature map. The residual layer allows the neural model to learn only the structure of the noise, reducing the number of required labeled samples and the overall complexity of the training phase. The network’s output is $y = f_{\theta,w,b}(x)$, where the parametric nonlinear map $f_{\theta,w,b}$ depends on the hyperparameters $\{\theta\}$, summarized in Table I, and on the weights w and biases b of the convolutional layers. The task of the training step is to learn the set of $\{w, b\}$ that solves the minimization problem,⁶⁷

$$\arg \min_{\{w,b\}} \sum_k L(f_{\theta,w,b}(x_k), y_k^{GT}) \tag{1}$$

being the hyperparameters $\{\theta\}$ fixed before the training. In Eq. (1), $\{(x_1, y_1^{GT}), \dots, (x_L, y_L^{GT})\}$ is the training set of simulated noisy raw

TABLE I. Hyperparameters for the HN and LN networks obtained by training with the HN and LN datasets and optimization with a grid search over the network architecture shown in Fig. 2.

$\{\theta\}$	Network HN	Network LN
N_{conv}	0	4
N_{param}	11k	10k
N_{kernel}	63	21
$N_{batch\ size}$	32	32
N_{epoch}^0	25	25
N_{epoch}^1	200	200
W_{grad}	0.6	0.6
ℓ	2	2

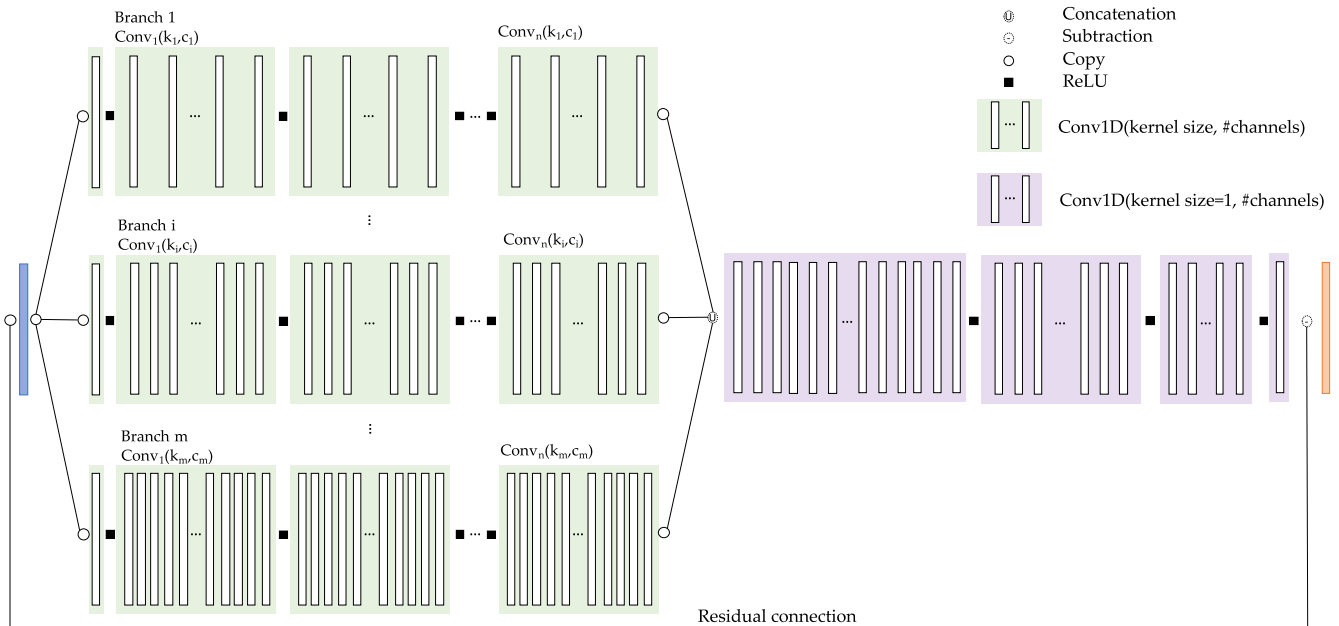


FIG. 2. Architecture of the neural network. The input data (blue box) feed multiple convolutional branches, which operate in parallel (green boxes). Within each branch, data are transformed by convolution layers ($Conv_1, \dots, Conv_n$) with the same kernel size and number of channels, connected by nonlinear activation functions (ReLU, black squares). Concatenated output data from the convolutional branches feed a series of 1D convolutional layers. The output (orange box) is obtained after the last residual layer.

spectra (x_k) and corresponding clean ones (y_k^{GT}), i.e., the ground truth (GT). Both x_k and y_k^{GT} are the 1D vectors of size n_{input} , functions of the sampled frequencies $\omega_1, \dots, \omega_{n_{input}}$. We performed stochastic gradient descent⁶⁸ to minimize the custom loss function L ,

$$L(y, y^{GT}) = (1 - W_{grad}) \|y - y^{GT}\|^2 + W_{grad} \mathcal{N} \|\nabla y - \nabla y^{GT}\|^\ell, \quad (2)$$

where $\|\cdot\|^\ell$ indicates the ℓ norm and $y = f_{\theta,w,b}(x)$. L contains a reconstruction term and a derivative term depending on the gradients of the reconstructed spectrum and GT. The two terms are activated at different epochs during training; initially, only the reconstruction term is adopted for N_{epoch}^0 , allowing the algorithm to perform first estimations of possible baselines and noise without heavily impacting on the latter. Then the gradient term is included in the loss calculation, and the model is trained for additional N_{epoch}^1 epochs. The gradient term impacts heavily on the removal of the point-wise fluctuating noise, while the reconstruction term is sensitive to both the noise and corrections to the broadband baseline. The parallel optimization of these two terms is the key to obtain more accurate results with respect to performing these two tasks sequentially. The balance weight between the reconstruction and the gradient term is controlled by the hyperparameter $W_{grad} \in (0, 1)$, which has been optimized by looking at the network performances and fixed to the value reported in Table I. \mathcal{N} is a normalization factor fixed once per training set only to regularize the loss decay during training so that the value of the total loss does not experience an abrupt change during the first epoch after the gradient term is switched on. The final model architecture was selected after cross-validating all the model hyperparameters by means of grid-search, including N_{layers} , N_{conv} , the kernel sizes, learning rate, batch sizes, number of training epochs, and the relative weight between the gradient and reconstruction loss terms. The learning rate was also dynamically decreased during training, according to the scheduler. This also allowed for balancing the effects of the training with the full loss function in Eq. (2), which occurs with a smaller learning rate. Batch normalization and regularization techniques were not adopted since overfitting was not observed and they decreased the performance of the neural net. The kernel sizes have been varied linearly across the branches, from a minimal size of 5 to a maximal size of M_k , which was chosen by a hyperparameter sweep and then fixed to $M_k = 88$. The number of filters and, hence, the output channels of each convolutional block were chosen in a kernel-size dependent way to obtain the same number of trainable parameters, N_{param} , in each convolutional branch. In addition, N_{param} has been chosen by a hyperparameter sweep in order to find a trade-off between the needs of computational feasibility and performance. Backpropagation with the Adam optimizer was used to train the model.⁶⁶ We found that the optimal architecture depends on the level of variability of the signal-to-noise in the training set. For a network trained on the dataset HN with a high level of noise, the best architecture (network HN) was obtained for the value of the hyperparameters reported in the second column of Table I. For the LN training set with a lower maximum value of noise, we obtained the best performances with the network specified by the hyperparameters in the third column of Table I (network LN). The choice of the best hyperparameters to be used on experimental data can be performed using a discriminator on the signal-to-noise level of the data, as detailed in

Sec. II D of the [supplementary material](#). Code with a minimal example implementation, datasets, and pretrained weights discussed here are available online.⁶²

IV. RESULTS AND DISCUSSION

A. Validation on simulated data

We tested the networks trained on the HN and LN datasets on test sets containing 20% of the samples of the corresponding datasets, previously unseen by the networks during training. In Fig. 3, we present the result for five typical spectra extracted from the HN (panel a) and LN (panel b) test sets. The spectra reconstructed by the NN (green lines) show excellent agreement with the clean GT spectra (black dotted lines) for very different baseline shapes, which are shown by the raw data in red in the top panels. These examples are compared to the results obtained by an iterative spline (iSpline) procedure for baseline removal⁶⁹ followed by Savitzky–Golay (SG) filtering. We note that the parameters of the SG filter have been chosen *a posteriori* by using the GT in order to obtain a trade-off between an optimal smoothing of the point-wise noise and the preservation of the linewidth of the peaks. Consequently, this is an optimal procedure that cannot be met in a real case scenario, i.e., when the GT is not available. Notwithstanding, the reconstruction of the lineshape, linewidth, and relative intensity of the peaks and the overall removal of the baseline and the noisy point-wise fluctuations are qualitatively better when using the NN.

For a quantitative comparison, it is necessary to define indicators that measure the capability of a given algorithm to perform a specific task. We combined selected metrics from the literature with custom-defined ones to evaluate the performance of the NN for tasks more closely related to spectroscopic purposes, namely, the identification of peak positions and lineshapes, evaluation of relative intensities, separation of overlapped components, and minimization of false-peak predictions. In particular, we computed the Structural Similarity Index Measure (SSIM), which evaluates the overall resemblance of the processed Raman spectra to the GT, the Normalized Mean Absolute and Mean Squared Errors (NMAEs and NMSEs), as defined in Sec. III B of the [supplementary material](#). In addition to these standard metrics, we developed a custom edge finder that determines the spectral positions of the positive and negative peaks from an input spectrum resulting from the processing of the noisy raw data and compares them to the associated GT spectrum. We used the edge finder to count the correctly assigned edges [true positive (TP)] and the errors induced by the data processing [false positive and false negative (FP and FN)]. The edges are considered TP if their spectral position deviates from the corresponding GT by more than a tunable value of tolerance. In the following, we set the tolerance to one pixel. This allows us to redefine this problem as a classification task and calculate standard metrics as the *F1-score*,

$$F1 = \frac{2TP}{2TP + FP + FN} \quad (3)$$

and *precision*,

$$\text{precision} = \frac{TP}{TP + FP}. \quad (4)$$

Precision counts the number of correctly retrieved edges scaled by the total number of edges found, while the F1-score measures the

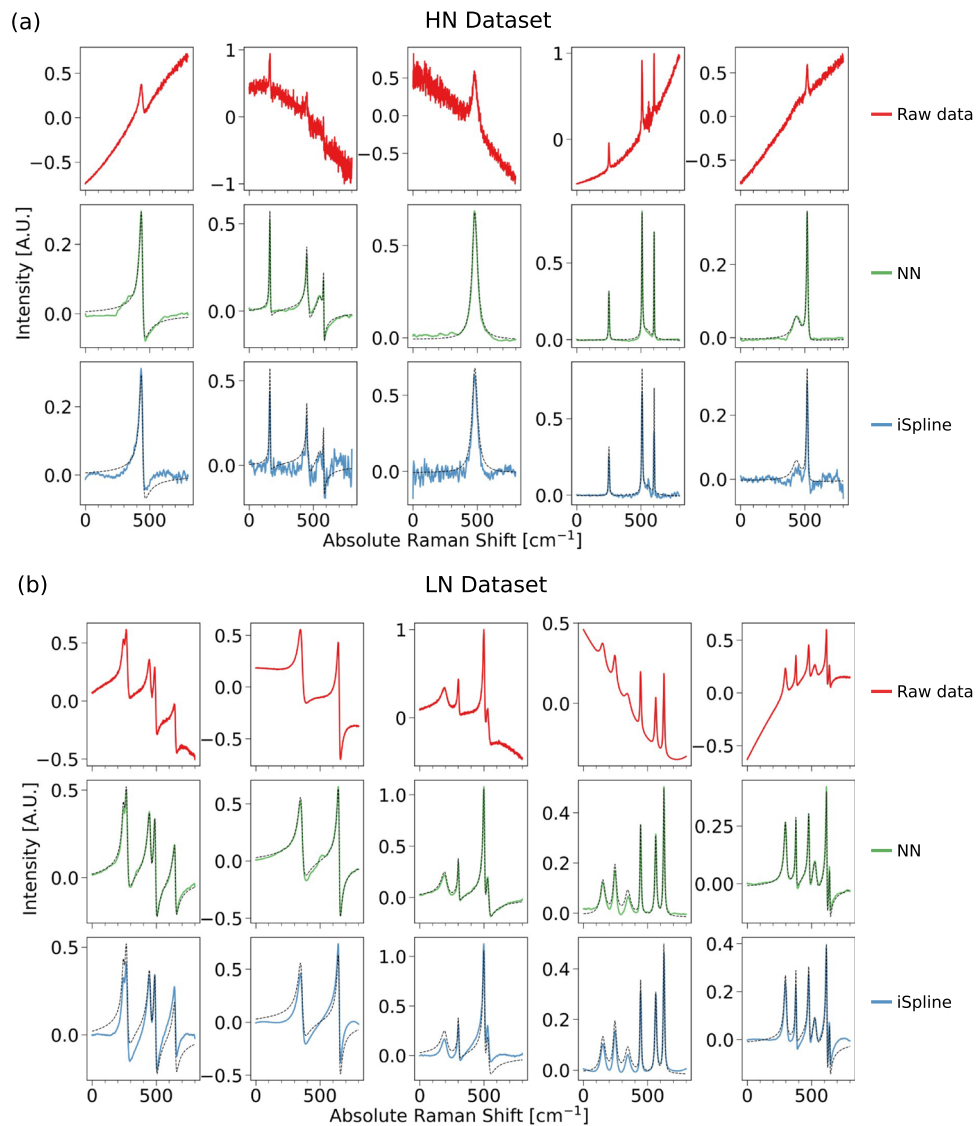


FIG. 3. Evaluation of representative simulated test samples from the HN (a) and LN (b) datasets. For each sample, noisy raw data are shown in the top panel (red lines). The central and bottom panels show the SRS spectra obtained by using the NN (green) or the iterative spline procedure for baseline subtraction followed by Savitzky–Golay filtering (blue). GTs are also reported for comparison (black dotted lines).

ability of each algorithm to deliver correct predictions, balancing the possible predicting biases toward FPs and FNs. Both metrics range from zero, indicating the worst performance, to a best value of one. Finally, we defined custom Signal-to-Noise Ratio (SNR) metrics as

$$\text{SNR} = \frac{\text{Area}(I_y) - \langle N_y \rangle}{\|N_y - N_{y,GT}\|^2}, \quad (5)$$

$$I = y(\omega_v),$$

$$N_f = f(\omega \neq \omega_v) \text{ with } f = y, y^{GT},$$

where $\omega_v = \{\omega_i, \forall i : |\omega_i - \omega_{peak}^{GT}| < \varepsilon\}$ and $\varepsilon = 80$ pixels, which was set taking into consideration the typical width of the Raman feature. This metric measures the capability of an algorithm to obtain a clean and smooth baseline with respect to the area of the main peaks in the spectral position defined by the GT.

We have compared the results obtained by the NN with those obtained by four different traditional algorithms for baseline removal: third-order modified polynomial (iModPoly),⁷⁰ iterative spline, SG filter, and peak-screened Asymmetric Least Squares (ALSs).⁷¹ For each algorithm, baseline removal was followed by an additional SG filter for point-wise denoising and smoothing (see Sec. III A of the [supplementary material](#) for additional details).

The results obtained for the HN test set are shown in panels (a) and (b) of Fig. 4, while the corresponding results for the LN dataset are reported in panels (c) and (d). For both datasets, the NN outperforms the standard algorithms in identifying all the edges in the SRS spectrum. In particular, for the HN dataset, the NN achieves a full precision score in 63% of the test samples, compared to 42% of the best standard algorithm (the iSpline+SG filter). More importantly, the distribution of the F1-score achieved by the NN is narrower and shifted to higher values, with a mean of 0.86 and a standard deviation of 0.18 to be compared with the mean of 0.70 and a standard deviation of 0.30 of the standard data processing routine. This is shown by the histograms of the distribution of F1-scores reported in Fig. S2 of the [supplementary material](#), while, in the whisker box plots of Fig. 4(a), we isolate the quartile of the test set in which each of the two algorithms achieved the worst results to highlight the differences in their performances. We note that the values of precision depend on the amount of tolerance on the peak position discrepancy between the GT and reconstructed data, but this dependence does not alter the results presented here (details are reported in Sec. II B of the [supplementary material](#)). In addition for the other tasks

that we have considered, the NN shows superior performances, as illustrated by the analysis of the associated metrics in Fig. 4(b). In particular, for the metrics NMAE, NMSE, and 1/SNR, which project optimal results toward zero, the medians relative to the NN are lower than the first quartile of all the distributions of the standard methods. There is also a large improvement in the SSIM metrics—optimal at SSIM = +1—for which the median of the NN is larger than the 95th percentile of the best standard method. This is correlated with the ability of the NN to also recover the overall lineshape of the Raman features, in addition to the signal-to-noise, intensity, and peak positions, which are best measured by the other metrics. The performances of the standard methods are comparable to each other, with the polynomial and iterative spline algorithms giving the best results. Similar results are also obtained on the LN dataset, where all the methods achieve better results overall, with a gap between the best metrics of the NN and those of the standard algorithms, as shown in Figs. 4(c) and 4(d) and Sec. II C of the [supplementary material](#). In particular, the mean of the F1-score achieved by the NN is 0.92 with a standard deviation of 0.14, while the iSpline with the SG filter obtained a mean of 0.84 with a standard deviation of 0.24.

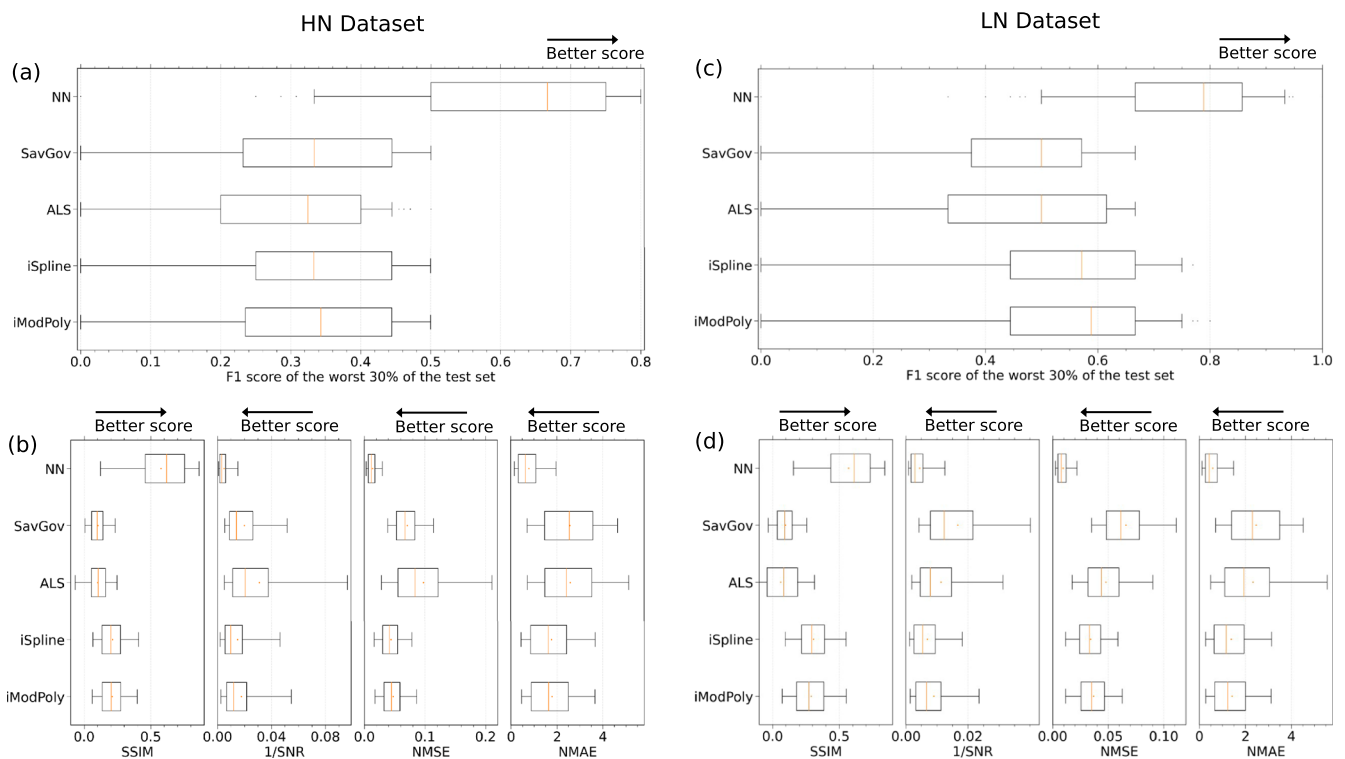


FIG. 4. Comparison of the statistical analysis over selected metrics between the NN method and multiple non-data driven routines applied to the HN [panels (a) and (b)] and LN [panels (c) and (d)] test sets. As indicated by the black arrows, for the F1 and SSIM metrics, the best scores correspond to higher values on the x-axis (optimal performance is at one). For all the other metrics, the best scores correspond to lower values on the x-axis (optimal performance is at zero). [(a) and (c)] Results of the classification metrics obtained by means of the edge finder algorithm. The whisker plots report on the F1-score relative to the portion of 30% of the test set for which each method has the worst score. [(b) and (d)] Whisker plots of the results obtained by the different methods measured by the SSIM, SNR, NMSE, and NMAE metrics. For all the whisker plots, the box covers from the first to the third quartile, while the whiskers extend from the box to the 5th and 95th percentiles. The orange line and dot indicate the median and the mean, respectively. Black dots indicate values that are past the end of the whiskers.

B. Validation on experimental data

To validate the NN denoiser in a real case scenario, we have applied it to the resonant SRS spectra measured on equine heart deoxy Myoglobin (Mb) dissolved in pH 7.4 buffer and Cresyl Violet (CV) dissolved in methanol, using the hyperparameters found by training with the LN dataset, whose signal-to-noise ratios are comparable to the ones obtained for these particular samples with our SRS setup. In Fig. 5(a), we report the raw experimental SRS spectra of Mb pumped across the Soret absorption band, with the RP wavelength tuned at 447 and 472 nm for the red and blue sides of the spectrum, respectively. The resonant condition impacts differently on the two sides of the spectrum and makes the Raman lineshapes highly dependent on the RP wavelength and, for the blue side, on the frequency of the normal mode.⁴⁹ In the central and lower panels, we show the corresponding spectra processed by the NN and the polynomial algorithm. The results are in good agreement with

the literature,^{69,72} and the best results were obtained by traditional processing routines for both the gain and dispersive lineshapes.

We then tested the algorithm in the more demanding case of CV undergoing photoexcitation due to a resonant, high-fluence RP tuned at 580 nm. In these conditions, the RP also acts as an actinic pump, inducing electronically excited-state populations in the sample by means of two additional light-matter interactions preceding the SRS process.⁷⁴ The efficiency of this higher-order nonlinear process depends on the time delay between the Raman and the probe. By tuning the delay, it is possible to control the amount of excited-state population induced by the Raman pulse. Here, a positive (negative) sign indicates that the PP follows (precedes) the RP. At large negative delays, the Raman scattering mainly involves molecules initially in the electronically ground state. At positive delays, SRS can probe vibrational transitions originating from a mixture of excited and ground-state populations, with the cost of a lower signal-to-

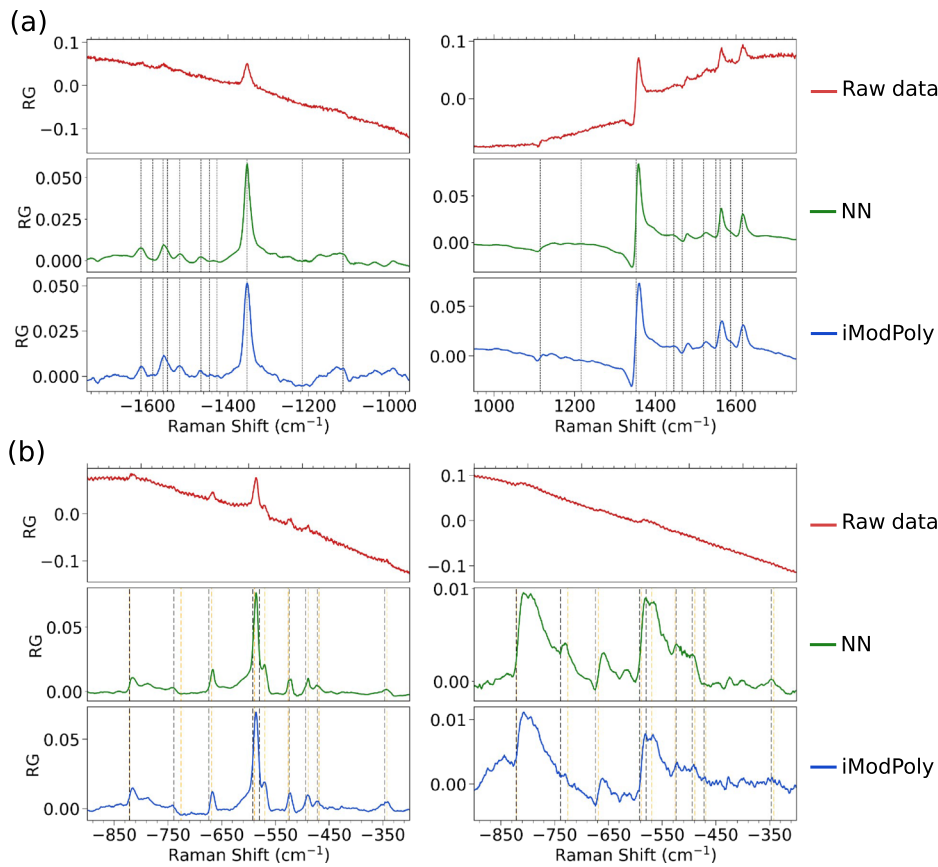


FIG. 5. Application to SRS experimental data. (a) SRS spectra of deoxy Mb. Left panels: red side Raman spectra obtained with RP tuned at 447 nm. Right Panels: blue side Raman spectra obtained with the RP tuned at 472 nm. Raw data are indicated by red lines, while green and blue lines represent processed spectra predicted by the NN network and by the iModPoly algorithm for baseline removal, followed by a SG noise filter. Vertical dashed lines indicate the spectral position of the Raman modes of deoxy Mb for resonant excitation in the Soret absorption band⁶⁹ (reported in Table S2 of the [supplementary material](#)). (b) SRS spectra of Cresyl Violet obtained upon resonant excitation with a RP tuned at 580 nm at two different delays between the RP and PP: -0.8 ps (PP preceding RP, left panels) and $+1.2$ ps (PP following RP, right panels). Raw data are indicated by red lines, while green and blue lines represent processed spectra predicted by the NN network and by the iModPoly algorithm for baseline removal, followed by a SG noise filter. Vertical dashed lines indicate the spectral position of the Raman modes of the ground (in black)^{73–76} and the excited states (in orange),^{73–75} which are reported in Table S3 of the [supplementary material](#).

noise ratio, spectral distortions, and an overall increased complexity of the baseline and lineshapes. We considered the SRS spectra for two different delays between the Raman and probe pulses: -0.8 ps (PP preceding RP) and $+1.2$ ps (PP following RP). In Fig. 5(b), we report the raw SRS spectra at the two delays and the corresponding ones processed by the NN and by the iModPoly algorithm, followed by SG filtering. For the negative delay (left panel), the processed spectra present sharp and intense features, with a high signal-to-noise ratio and spectral positions matching the frequencies of both the electronically excited and ground-state Raman modes of CV, due to the resonant RP wavelength. For the positive delay (right panel), as a consequence of the superposition of the Raman processes originating from a system initially prepared in the ground or the excited-state population,⁷⁴ spectral features appear broadened and red-shifted, their lineshapes become dispersive and asymmetric, and the signal-to-noise ratio decreases drastically due to the effective time window during which the vibrational coherences are sampled.⁶⁹ The increased complexity has a different impact on the two methods used to process the raw spectra. Notably, the combined iModPoly and SG algorithms are not able to resolve the excited-state peak at 733 cm^{-1} , which is conversely retrieved by the NN. The polynomial processing also shows parasitic broad peaks between 850 and 900 cm^{-1} and poorly captures the dispersive line shape at 680 cm^{-1} . Moreover, the two approaches give different relative intensities between the peaks, as also observed for the simulated datasets. Even if a microscopic model of the excited state effects causing the changes in the Raman peaks was not included in the simulated dataset, during the training phase, the NN was exposed to a statistical population of lineshapes, intensities, and spectral positions that was large enough to allow it to interpolate among a wider range of peak characteristics. These results confirm that the NN algorithm achieves higher performances in the case of complex experimental spectra.

V. CONCLUSION

Low signal-to-noise ratios and structured spectral lineshapes in nonlinear Raman spectroscopy cause loss of information and increase the complexity of performing the measurements due to the need for longer exposure times for averaging and extensive data analysis. We have demonstrated that wisely designed deep neural networks can overcome these limitations and achieve background removal and denoising of stimulated Raman spectra. By means of multiple kernel sizes operating in parallel within a convolutional residual neural architecture, it is possible to adapt the receptive field of the network to the informative features in the spectra and treat the multiple spectral scales present in the SRS data, related to the complex Raman lineshapes, to the background, and to point-wise noisy fluctuations. We have shown that such architecture is able to adapt to different levels of noise and prominence of the Raman bands from the baseline by training on datasets simulated through the diagrammatic theory and characterized by noise and material parameters resembling those present in experimental conditions. Depending on the level of noise, the NN demonstrated itself to be comparable to or superior to the standard algorithms commonly used for SRS data post-processing. The advantages are particularly evident in the presence of multiple features with asymmetric lineshapes and intensities, that are weak compared to the noise and the other bands in the

same spectral region. The NN algorithm was able to identify the Raman bands, reconstruct the correct lineshapes and relative intensities and enhance the spectral resolution by resolving the vibrational frequencies and bandwidths for close or overlapping Raman features. These abilities are pivotal for the interpretation of experiments leveraging resonant optical excitations to detect cooperative mechanisms between coherent vibrations and electronic excitations. Once trained, the network generalizes to experimental spectra obtained on different samples, preserving its high performance. It can be adopted as a post-processing routine to access the wealth of information concealed by the data complexity in SRS and combined with optimization methods that operate during the data acquisition to enhance the instrumental SNR of the technique. We note that SRS is an optimal test-bed for investigating deep learning applications to nonlinear spectroscopy, given its spectral complexity and variety of lineshapes for both the Raman features and background contributions. For these reasons, we anticipate the possibility of extending the use of the proposed NN architecture to different linear and nonlinear spectroscopic techniques that are affected by similar noise effects by means of fine tuning of the final NC layers and transfer learning techniques.

SUPPLEMENTARY MATERIAL

See the [supplementary material](#) for the following: (I) Generation of the simulated datasets, (II) analysis of the performance of HN and LN networks on the simulated datasets, (III) standard post-processing methods and evaluation metrics, (IV) experimental methods, and (V) frequencies of the Raman modes in the experimental spectra.

ACKNOWLEDGMENTS

G.F. thanks Marco Fumero for many fruitful and enjoyable discussions. This project has received funding from the PRIN 2020 Project, Grant No. 2020HTSXMA-PSIMOVIE (G.B.), from the European Union's Horizon 2020 research and innovation program Graphene Flagship under Grant Agreement No. 881603 (T.S.), and has been partially supported by PNRR MUR Project No. PE0000013-FAIR (S.G.). G.B. and T.S. acknowledge the "Progetti di Ricerca Medi 2020," the "Progetti di Ricerca Medi 2021," and the "Progetti di Ricerca Medi 2022" grants by Sapienza Università di Roma. G.F. acknowledges the support from the grant "Avvio alla Ricerca 2022" by Sapienza Università di Roma.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Giuseppe Fumero: Conceptualization (equal); Data curation (lead); Formal analysis (lead); Investigation (lead); Methodology (lead); Project administration (equal); Software (lead); Supervision (equal);

Validation (equal); Visualization (equal); Writing – original draft (lead); Writing – review & editing (equal). **Giovanni Batignani**: Conceptualization (equal); Data curation (supporting); Investigation (supporting); Validation (equal); Writing – original draft (supporting); Writing – review & editing (equal). **Edoardo Cassetta**: Data curation (supporting); Formal analysis (supporting); Software (supporting). **Carino Ferrante**: Investigation (supporting); Validation (equal); Writing – review & editing (equal). **Stefano Giagu**: Conceptualization (equal); Funding acquisition (equal); Methodology (supporting); Project administration (equal); Resources (equal); Supervision (equal); Validation (equal); Writing – original draft (supporting); Writing – review & editing (equal). **Tullio Scopigno**: Conceptualization (equal); Funding acquisition (equal); Project administration (equal); Resources (equal); Supervision (equal); Validation (equal); Writing – original draft (supporting); Writing – review & editing (equal).

DATA AVAILABILITY

Code with minimal example implementation, datasets and pretrained weights that support the findings of this study are openly available online, in a dedicated GitHub repository at <https://github.com/Gifum/SRSdenoiser>. The SRS experimental data on CV and Mb are available from the corresponding author upon reasonable request.

REFERENCES

- M. Maiuri, M. Garavelli, and G. Cerullo, “Ultrafast spectroscopy: State of the art and open challenges,” *J. Am. Chem. Soc.* **142**, 3–15 (2019).
- Y. Han, W. Liang, X. Lin, Y. Li, F. Sun, F. Zhang, P. C. Sercel, and K. Wu, “Lattice distortion inducing exciton splitting and coherent quantum beating in CsPbI₃ perovskite quantum dots,” *Nat. Mater.* **21**, 1282–1289 (2022).
- J. Bloch, A. Cavalleri, V. Galitski, M. Hafezi, and A. Rubio, “Strongly correlated electron–photon systems,” *Nature* **606**, 41–48 (2022).
- P. Malý, J. Lüttig, P. A. Rose, A. Turkin, C. Lambert, J. J. Krich, and T. Brixner, “Separating single- from multi-particle dynamics in nonlinear spectroscopy,” *Nature* **616**, 280–287 (2023).
- R. Ulbricht, E. Hendry, J. Shan, T. F. Heinz, and M. Bonn, “Carrier dynamics in semiconductors studied with time-resolved terahertz spectroscopy,” *Rev. Mod. Phys.* **83**, 543–586 (2011).
- F. Provencher, N. Bérubé, A. W. Parker, G. M. Greetham, M. Towrie, C. Hellmann, M. Côté, N. Stingelin, C. Silva, and S. C. Hayes, “Direct observation of ultrafast long-range charge separation at polymer–fullerene heterojunctions,” *Nat. Commun.* **5**, 4288 (2014).
- T. Zhu, L. Yuan, Y. Zhao, M. Zhou, Y. Wan, J. Mei, and L. Huang, “Highly mobile charge-transfer excitons in two-dimensional WS₂/tetracene heterostructures,” *Sci. Adv.* **4**, eaao3104 (2018).
- C. Ferrante, G. Di Battista, L. E. P. López, G. Batignani, E. Lorchat, A. Virga, S. Berciaud, and T. Scopigno, “Picosecond energy transfer in a transition metal dichalcogenide–graphene heterostructure revealed by transient Raman spectroscopy,” *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2119726119 (2022).
- A. J. Musser, M. Liebel, C. Schnedermann, T. Wende, T. B. Kehoe, A. Rao, and P. Kukura, “Evidence for conical intersection dynamics mediating ultrafast singlet exciton fission,” *Nat. Phys.* **11**, 352–357 (2015).
- S. P. Neville, M. Chergui, A. Stolow, and M. S. Schuurman, “Ultrafast X-ray spectroscopy of conical intersections,” *Phys. Rev. Lett.* **120**, 243001 (2018).
- K. S. Zinchenko, F. Ardana-Lamas, I. Seidu, S. P. Neville, J. van der Veen, V. U. Lanfaloni, M. S. Schuurman, and H. J. Wörner, “Sub-7-femtosecond conical-intersection dynamics probed at the carbon K-edge,” *Science* **371**, 489–494 (2021).
- H. Kuramochi, S. Takeuchi, K. Yonezawa, H. Kamikubo, M. Kataoka, and T. Tahara, “Probing the early stages of photoreception in photoactive yellow protein with ultrafast time-domain Raman spectroscopy,” *Nat. Chem.* **9**, 660–666 (2017).
- G. Fumero, C. Schnedermann, G. Batignani, T. Wende, M. Liebel, G. Bassolino, C. Ferrante, S. Mukamel, P. Kukura, and T. Scopigno, “Two-dimensional impulsively stimulated resonant Raman spectroscopy of molecular excited states,” *Phys. Rev. X* **10**, 011051 (2020).
- J. D. Gaynor, J. Sandwisch, and M. Khalil, “Vibronic coherence evolution in multidimensional ultrafast photochemical processes,” *Nat. Commun.* **10**, 5621 (2019).
- P. T. C. So, C. Y. Dong, B. R. Masters, and K. M. Berland, “Two-photon excitation fluorescence microscopy,” *Annu. Rev. Biomed. Eng.* **2**, 399–429 (2000).
- W. Min, C. W. Freudiger, S. Lu, and X. S. Xie, “Coherent nonlinear optical imaging: Beyond fluorescence microscopy,” *Annu. Rev. Phys. Chem.* **62**, 507–530 (2011).
- C. H. Camp, Jr. and M. T. Cicerone, “Chemically sensitive bioimaging with coherent Raman scattering,” *Nat. Photonics* **9**, 295–305 (2015).
- D. Polli, V. Kumar, C. M. Valensise, M. Marangoni, and G. Cerullo, “Broadband coherent Raman scattering microscopy,” *Laser Photonics Rev.* **12**, 1800020 (2018).
- R. C. Prince, R. R. Frontiera, and E. O. Potma, “Stimulated Raman scattering: From bulk to nano,” *Chem. Rev.* **117**, 5070–5094 (2016).
- C. R. Hall, J. Conyard, I. A. Heisler, G. Jones, J. Frost, W. R. Browne, B. L. Feringa, and S. R. Meech, “Ultrafast dynamics in light-driven molecular rotary motors probed by femtosecond stimulated Raman spectroscopy,” *J. Am. Chem. Soc.* **139**, 7408–7414 (2017).
- M. S. Barclay, M. Caricato, and C. G. Elles, “Femtosecond stimulated Raman scattering from triplet electronic states: Experimental and theoretical study of resonance enhancements,” *J. Phys. Chem. A* **123**, 7720–7732 (2019).
- S. Tahara, H. Kuramochi, S. Takeuchi, and T. Tahara, “Protein dynamics preceding photoisomerization of the retinal chromophore in bacteriorhodopsin revealed by deep-UV femtosecond stimulated Raman spectroscopy,” *J. Phys. Chem. Lett.* **10**, 5422–5427 (2019).
- C. Ferrante, G. Batignani, E. Pontecorvo, L. C. Montemiglio, M. H. Vos, and T. Scopigno, “Ultrafast dynamics and vibrational relaxation in six-coordinate heme proteins revealed by femtosecond stimulated Raman spectroscopy,” *J. Am. Chem. Soc.* **142**, 2285–2292 (2020).
- C. Fang and L. Tang, “Mapping structural dynamics of proteins with femtosecond stimulated Raman spectroscopy,” *Annu. Rev. Phys. Chem.* **71**, 239–265 (2020).
- L. M. Malard, L. Lafeta, R. S. Cunha, R. Nadas, A. Gadelha, L. G. Cançado, and A. Jorio, “Studying 2D materials with advanced Raman spectroscopy: CARS, SRS and TERS,” *Phys. Chem. Chem. Phys.* **23**, 23428–23444 (2021).
- G. Batignani, C. Ferrante, G. Fumero, M. Martinati, and T. Scopigno, “Femtosecond stimulated Raman spectroscopy,” *Nat. Rev. Methods Primers* **4**, 34 (2024).
- Y. Hontani, M. Klotz, T. Polívka, M. K. Shukla, R. Sobotka, and J. T. M. Kennis, “Molecular origin of photoprotection in cyanobacteria probed by watermarked femtosecond stimulated Raman spectroscopy,” *J. Phys. Chem. Lett.* **9**, 1788–1792 (2018).
- G. Batignani, G. Fumero, E. Pontecorvo, C. Ferrante, S. Mukamel, and T. Scopigno, “Genuine dynamics vs cross phase modulation artifacts in femtosecond stimulated Raman spectroscopy,” *ACS Photonics* **6**, 492–500 (2019).
- R. Ranjan, G. Costa, M. A. Ferrara, M. Sansone, and L. Sirlito, “Noises investigations and image denoising in femtosecond stimulated Raman scattering microscopy,” *J. Biophot.* **15**, e202100379 (2022).
- L. Genchi, S. P. Laptchenk, and C. Liberale, “Background signals in stimulated Raman scattering microscopy and current solutions to avoid them,” *Adv. Phys.: X* **8**(1), 2176258 (2023).
- R. Gautam, S. Vanga, F. Ariese, and S. Umopathy, “Review of multidimensional data processing approaches for Raman and infrared spectroscopy,” *EPJ Tech. Instrum.* **2**, 8 (2015).
- X. Shen, S. Ye, L. Xu, R. Hu, L. Jin, H. Xu, J. Liu, and W. Liu, “Study on baseline correction methods for the Fourier transform infrared spectra with different signal-to-noise ratios,” *Appl. Opt.* **57**, 5794 (2018).

- ³³M. Kloz, R. van Grondelle, and J. T. Kennis, "Wavelength-modulated femtosecond stimulated Raman spectroscopy—Approach towards automatic data processing," *Phys. Chem. Chem. Phys.* **13**, 18123 (2011).
- ³⁴W. J. Thrift, S. Ronaghi, M. Samad, H. Wei, D. G. Nguyen, A. S. Cabuslay, C. E. Groome, P. J. Santiago, P. Baldi, A. I. Hochbaum, and R. Ragan, "Deep learning analysis of vibrational spectra of bacterial lysate for rapid antimicrobial susceptibility testing," *ACS Nano* **14**, 15336–15348 (2020).
- ³⁵C. M. Valensise, A. Giuseppi, F. Vernuccio, A. De la Cadena, G. Cerullo, and D. Polli, "Removing non-resonant background from CARS spectra via deep learning," *APL Photonics* **5**, 061305 (2020).
- ³⁶A. Hosseinizadeh, N. Breckwoldt, R. Fung, R. Sepehr, M. Schmidt, P. Schwander, R. Santra, and A. Ourmazd, "Few-fs resolution of a photoactive protein traversing a conical intersection," *Nature* **599**, 697–701 (2021).
- ³⁷C. M. Valensise, A. Giuseppi, G. Cerullo, and D. Polli, "Deep reinforcement learning control of white-light continuum generation," *Optica* **8**, 239 (2021).
- ³⁸Z. Chen, N. Andrejevic, N. C. Drucker, T. Nguyen, R. P. Xian, T. Smidt, Y. Wang, R. Ernstorfer, D. A. Tennant, M. Chan, and M. Li, "Machine learning on neutron and x-ray scattering and spectroscopies," *Chem. Phys. Rev.* **2**, 031301 (2021).
- ³⁹A. Bresci, M. Guizzardi, C. M. Valensise, F. Marangi, F. Scotognella, G. Cerullo, and D. Polli, "Removal of cross-phase modulation artifacts in ultrafast pump-probe dynamics by deep learning," *APL Photonics* **6**, 076104 (2021).
- ⁴⁰M. Stanfield, J. Ott, C. Gardner, N. F. Beier, D. M. Farinella, C. A. Mancuso, P. Baldi, and F. Dollar, "Real-time reconstruction of high energy, ultrafast laser pulses using deep learning," *Sci. Rep.* **12**, 5299 (2022).
- ⁴¹F. Vernuccio, A. Bresci, V. Cimini, A. Giuseppi, G. Cerullo, D. Polli, and C. M. Valensise, "Artificial intelligence in classical and quantum photonics," *Laser Photonics Rev.* **16**, 2100399 (2022).
- ⁴²C. D. Rankine and T. J. Penfold, "Accurate, affordable, and generalizable machine learning simulations of transition metal x-ray absorption spectra using the XANESNET deep neural network," *J. Chem. Phys.* **156**, 164102 (2022).
- ⁴³J. Wahl, M. Sjö Dahl, and K. Ramser, "Single-step preprocessing of Raman spectra using convolutional neural networks," *Appl. Spectrosc.* **74**, 427–438 (2020).
- ⁴⁴S. Barton, S. Alakkari, K. O'Dwyer, T. Ward, and B. Hennelly, "Convolution network with custom loss function for the denoising of low SNR Raman spectra," *Sensors* **21**, 4623 (2021).
- ⁴⁵P. Abdolghader, A. Ridsdale, T. Grammatikopoulos, G. Resch, F. Légaré, A. Stolow, A. F. Pegoraro, and I. Tamblin, "Unsupervised hyperspectral stimulated Raman microscopy image enhancement: Denoising and segmentation via one-shot deep learning," *Opt. Express* **29**, 34205 (2021).
- ⁴⁶M. T. Gebrekidan, C. Knipfer, and A. S. Braeuer, "Refinement of spectra using a deep neural network: Fully automated removal of noise and background," *J. Raman Spectrosc.* **52**, 723–736 (2021).
- ⁴⁷J. Shen, M. Li, Z. Li, Z. Zhang, and X. Zhang, "Single convolutional neural network model for multiple preprocessing of Raman spectra," *Vib. Spectrosc.* **121**, 103391 (2022).
- ⁴⁸R. Luo, J. Popp, and T. Bocklitz, "Deep learning for Raman spectroscopy: A review," *Analytica* **3**, 287–301 (2022).
- ⁴⁹G. Batignani, E. Pontecorvo, G. Giovannetti, C. Ferrante, G. Fumero, and T. Scopigno, "Electronic resonances in broadband stimulated Raman spectroscopy," *Sci. Rep.* **6**, 18445 (2016).
- ⁵⁰A. L. Dobryakov, O. A. Krohn, M. Quick, I. N. Ioffe, and S. A. Kovalenko, "Positive and negative signal and line shape in stimulated Raman spectroscopy: Resonance femtosecond Raman spectra of diphenylbutadiene," *J. Chem. Phys.* **156**, 084304 (2022).
- ⁵¹G. Batignani, G. Fumero, E. Mai, M. Martinati, and T. Scopigno, "(INVITED) Stimulated Raman lineshapes in the large light-matter interaction limit," *Opt. Mater.: X* **13**, 100134 (2022).
- ⁵²R. Berera, R. van Grondelle, and J. T. M. Kennis, "Ultrafast transient absorption spectroscopy: Principles and application to photosynthetic systems," *Photosynth. Res.* **101**, 105–118 (2009).
- ⁵³P. Roy, F. Al-Kahtani, A. N. Cammidge, and S. R. Meech, "Solvent tuning excited state structural dynamics in a novel bianthryl," *J. Phys. Chem. Lett.* **14**, 253–259 (2023).
- ⁵⁴K. E. H. Anderson, S. L. Sewall, R. R. Cooney, and P. Kambhampati, "Noise analysis and noise reduction methods in kilohertz pump-probe experiments," *Rev. Sci. Instrum.* **78**, 073101 (2007).
- ⁵⁵X. Audier, S. Heuke, P. Volz, I. Rimke, and H. Rigneault, "Noise in stimulated Raman scattering measurement: From basics to practice," *APL Photonics* **5**, 011101 (2020).
- ⁵⁶K. C. Robben and C. M. Cheatum, "Edge-pixel referencing suppresses correlated baseline noise in heterodyned spectroscopies," *J. Chem. Phys.* **152**, 094201 (2020).
- ⁵⁷S. Mukamel, *Principles of Nonlinear Optical Spectroscopy, Oxford Series in Optical and Imaging Sciences* (Oxford University Press, New York, 1999).
- ⁵⁸K. E. Dorfman, B. P. Fingerhut, and S. Mukamel, "Broadband infrared and Raman probes of excited-state vibrational molecular dynamics: Simulation protocols based on loop diagrams," *Phys. Chem. Chem. Phys.* **15**, 12348 (2013).
- ⁵⁹G. Fumero, G. Batignani, K. E. Dorfman, S. Mukamel, and T. Scopigno, "On the resolution limit of femtosecond stimulated Raman spectroscopy: Modelling fifth-order signals with overlapping pulses," *ChemPhysChem* **16**, 3438–3443 (2015).
- ⁶⁰M. Kowalewski, B. P. Fingerhut, K. E. Dorfman, K. Bennett, and S. Mukamel, "Simulating coherent multidimensional spectroscopy of nonadiabatic molecular processes: From the infrared to the X-ray regime," *Chem. Rev.* **117**, 12165–12226 (2017).
- ⁶¹The electronic noise due to the detection process is normally negligible with respect to the shot noise when using single shot detection with a large dynamic range CCD camera, which is the common acquisition scheme for SRS. In different contexts, it can be included by means of a random additive noise.
- ⁶²SRSdenoiser repository, Software available from <https://github.com/Gifum/SRSdenoiser>, 2023.
- ⁶³Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.* **1**, 541–551 (1989).
- ⁶⁴M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Computer Vision – ECCV 2014* (Springer International Publishing, 2014), pp. 818–833.
- ⁶⁵K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.* **26**, 3142–3155 (2017).
- ⁶⁶I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, 2016), <https://www.deeplearningbook.org>.
- ⁶⁷Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436–444 (2015).
- ⁶⁸S. Ruder, "An overview of gradient descent optimization algorithms," [arXiv:1609.04747](https://arxiv.org/abs/1609.04747) (2017).
- ⁶⁹C. Ferrante, G. Batignani, G. Fumero, E. Pontecorvo, A. Virga, L. C. Montemiglio, G. Cerullo, M. H. Vos, and T. Scopigno, "Resonant broadband stimulated Raman scattering in myoglobin," *J. Raman Spectrosc.* **49**, 913–920 (2018).
- ⁷⁰J. Zhao, H. Lui, D. I. McLean, and H. Zeng, "Automated autofluorescence background subtraction algorithm for biomedical Raman spectroscopy," *Appl. Spectrosc.* **61**, 1225–1232 (2007).
- ⁷¹V. I. Korepanov, "Asymmetric least-squares baseline algorithm with peak screening for automatic processing of the Raman spectra," *J. Raman Spectrosc.* **51**, 2061–2065 (2020).
- ⁷²T. G. Spiro, *Biological Applications of Raman Spectroscopy: Resonance Raman Spectra of Heme and Metalloproteins* (John Wiley & Sons, 1988), Vol. 3.
- ⁷³C. Fitzpatrick, J. H. Odhner, and R. J. Levis, "Spectral signatures of ground- and excited-state wavepacket interference after impulsive excitation," *J. Phys. Chem. A* **124**, 6856–6866 (2020).
- ⁷⁴G. Batignani, C. Ferrante, and T. Scopigno, "Accessing excited state molecular vibrations by femtosecond stimulated Raman spectroscopy," *J. Phys. Chem. Lett.* **11**, 7805–7813 (2020).
- ⁷⁵J. Lu, Y. Lee, and J. M. Anna, "Extracting the frequency-dependent dynamic Stokes shift from two-dimensional electronic spectra with prominent vibrational coherences," *J. Phys. Chem. B* **124**, 8857–8867 (2020).
- ⁷⁶G. Batignani, C. Sansone, C. Ferrante, G. Fumero, S. Mukamel, and T. Scopigno, "Excited-state energy surfaces in molecules revealed by impulsive stimulated Raman excitation profiles," *J. Phys. Chem. Lett.* **12**, 9239–9247 (2021).