Deep-Learned Time-Resolved Coherent Raman Spectroscopy

Yuchen Ma,[1] Zehua Han,[2] Haotian Song,[1,3] Tao Peng,[2,*] Zhedong Zhang,[4,†] and Marlan O. Scully[2,5,6]

[1]*School of Physics, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China*
[2]*Texas A&M University, College Station, Texas 77843, USA*
[3]*College of Physics & Astronomy, University of Manchester, Manchester M13 9PL, UK*
[4]*Department of Physics, City University of Hong Kong, Kowloon, Hong Kong SAR*
[5]*Baylor University, Waco, Texas, 76706, USA*
[6]*Princeton University, Princeton, New Jersey 08544, USA*

(Dated: October 23, 2021)

We propose an ultrafast frequency-resolved coherent Raman spectroscopy with deep learning algorithm. The deep-learned signal is capable of time and frequency resolutions that are not accessible by conventional Raman spectroscopy. We develop a schematic based on deep learning for Raman spectroscopy, and testify the feasibility of revealing the electronic coherence dynamics typically within 50fs. The heterodyne-detected deep-learned Raman signal is further developed to capture the phases of electronic coherence and emission in real-time domain. The paradigm in our work can be extended to other spectroscopic signals, e.g., pump-probe and photon-coincidence counting, illuminating a great potential of reconstructing the microscopic dynamics of molecules.

Introduction.–Ultrafast measurement [1], retrieves both time- and frequency-resolved photoelectron spectroscopy that helps observe dynamics in real time including structural dynamics in molecules as well as electronic dynamics on surfaces. The much shorter gate compared to the traced process enables us to yield the temporal evolution of the process. However, Heisenberg conjugation restricts ultrafast Spectroscopic resolution on frequency and temporal domain. Much delicated information in microscopic like the structure and dynamics of electronic process in molecules tends to be mixed and difficult to separate.

Recent works on stimulated Raman and pump-probe spectroscopies with entangled photon pairs theoretically derived the resolution enhancement results [2–4]. However, it should be noted that the entangled light from nonlinear crystal sources is difficult to prepare and manipulate. Other theoretical proposals focusing on temporal and spectral gates manipulation revealed the superresolved excitonic relaxation pathways and charge fluctuation [5–7]. It turns out the superresolution in spectroscopies is always an worthwhile route to explore because of the promising ultrafast dynamics applications in photophysical systems including low-dimensional semiconductor materials, exciton polaritons, and nano-plasmonics.

Machine learning, an exceptional approach that imitates multiple learning modes in human's brain, has been seeing a prodigious rise in last decades, and perceived as one of the main disruptive technologies impacting physical world [8]. It is broadly used in speeding up many-body simulations and classifying many-body quantum phases [9–14], improving imaging quality in computational imaging systems [15–18], and machine-learned quantum information and quantum computing[19–21]. Typically, recent work related to deep speckle correlation utilized neural network to retrieve the objects' information from groups of images smeared by dynamic turbid media [17]. This work inspires that deep-learning-based technique possesses an ability to extract information from physical systems, which lead to resolution [22–24], sensitivity [25], and efficiency [26] enhancement.

In this Letter, we apply this technique on ultrafast coherent Raman spectroscopy with classical pulses to obtain the time-frequency resolutions. Results under the FAST CARS (Femtosecond adaptive spectroscopic technique for coherent anti-Stokes Raman scattering) regime via deep learning clearly reflect the dynamics of electronic coherence. Here, we show how deep learning is integrated into the physical process and breaks Heisenberg limit.

Spectro-Net and its training.–The deep neural network (DNN), basically, should be trained by some pre-generated dataset in which each signal contains Raman resonances defined by randomized values and randomized damping rates, correspondingly. As is shown in Fig. 1 (a), the signals retrieved by classical ultrafast pulse are totally blurred both in temporal and frequency domain. Nevertheless, by making some difference in Raman signal expression which will be shown later, we could theoretically get the ideal Raman signal as the training labels. Through inputting these data and labels into the untrained DNN, in Fig. 1 (b), the neural networks will be trained to own the capability to identify spectral signals (in and out of dataset) and retrieve their time and frequency resolution. Therefore, owning experimental Raman results generated by ultrafast pulse, the well-trained DNN can read them, extract temporal and frequency information, and then predict the sub-Heisenberg Raman results as desired (Fig. 1 (c)). In this section, we schedule the network regime of deep-learned FAST CARS and describe the mechanism of FAST CARS with classical ultrafast pulse, in detail.

To begin with, in the following part of training process in deep learning, random Raman resonances and dephasing rates are generated in groups. Using these parameters and theoretically derived homodyne and heterodyne signals, we generate training inputs and ideal signal results as labels. The proposed scheme, as shown in Fig 1 (d), obeys the Unet architecture [27], which performs excellent in Semantic segmentation. Input of the Unet is a FAST CARS two-dimensional (2D) signal with unresolved frequency components and temporal evolution. Next, the input goes through the encoding path consist

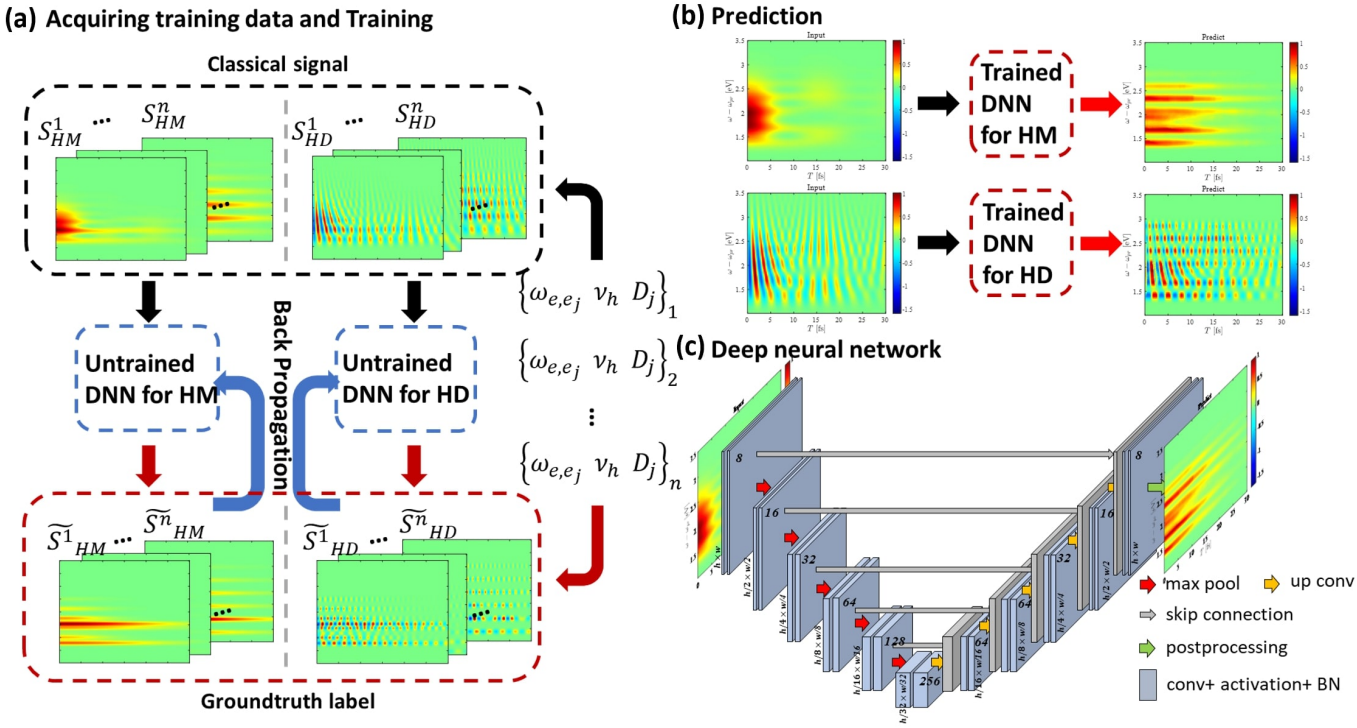

FIG. 1: Schematic of deep-learned time-resolved coherent Raman spectroscopy. (a) Generation of training and label dataset with random parameters. Then the DNN is trained with HM and HD signal dataset, respectively. (b) A well-trained DNN can transform unresolved FAST CARS signal with classical pulses into time- and frequency-resolved coherent Raman spectroscopy. (c) Architecture of the DNN. Higher-dimensional information of input image is gradually extracted along the encoding path. The decoder reverse this process by reorganizing those information.

of convolutional layers, maxpooling layers, and batch normalization layers; then goes through the decoding path consist of transpose convolutional layers, maxpooling layers, and batch normalization layers. The postprecessing step contains a convolutional layer with one-channel output and some operation based on physical priori knowledge, for example, normalization and signal intensity decrease with time. We use ReLu as activition function expect the input layer in DNN for HD, where we use tanh instead. The output of the Unet is a 2D signal having the same size with input signal. Typically, to make the neural network better at resolving frequency information, we proposed a mixed loss function $\mathrm{LOSS_{mix}}$ as

$$\begin{aligned}\mathrm{LOSS_{mix}} &= \mathrm{MSE}(S_{\mathrm{GT}} - S_{\mathrm{Pred}}) \\ &+ \lambda_1 \mathrm{MSE}(\mathcal{F}_{2D}(S_{\mathrm{GT}}) - \mathcal{F}_{2D}(S_{\mathrm{Pred}}))\end{aligned} \tag{1}$$

$\mathcal{F}_{2D}$ is the 2D fast Fourier transform (FFT). The first term of our mixed loss is traditional mean squared error (MSE). The second term, MSE of feature maps after 2D FFT, is proposed to strengthen the network's ability at eliminating beating in the input figure. λ_1 is coefficient, taken 0.01 in the training process. Namely, we stress this unique mixed loss function, Eq. 1, in which the MSE from signal domain and signal spectrum domain are combined. It is intuitive to find that the Fourier spectrum of the 2D signal tend to has good performance in training because the Raman signal generated by ultrafast pulse always maintains oscillation physically, though they usually combined with each other to make blurring in signal domain. By adding the MSE of Fourier spectrum of the 2D signal, therefore, the DNN performs much better in splitting Raman resonances under ultrafast dynamic situations. The training process was performed on the Google Colab Pro with GPU (NIVIDIA Tesla P100) using Keras/Tensorflow.

To test its feasibility for monitoring the ultrafast dynamics within rapid decay, we consider the ultrafast coherent Raman spectroscopy with 4-oriented amino-4'nitrostilbene for excited states and methane (CH_4) for ground states. The schematic is separated into two parts – training, and test. To begin with, for training process we prepare the training datasets based on the well-established theoretical FAST CARS signal expression of these two typical types of molecules. Both the excited-state Raman signal for 4-oriented amino-4'nitrostilbene and the FAST CARS signal for methane read [4]

$$S_{\mathrm{HM}}(\omega, T) = N(N-1)|\varepsilon_\omega|^4 |\sum_j \alpha^*_{e,e_j} f_{e,e_j}(T)|^2 \tag{2}$$

$$S_{\mathrm{HD}}(\omega, T) = N|\varepsilon_\omega|^2 \sum_j \mathrm{Im}[\alpha^*_{e,e_j} E^*_{LO}(\bar{\omega}) f_{e,e_j}(T)] \tag{3}$$

where $S_{\mathrm{HM}}(\omega, T)$ and $S_{\mathrm{HD}}(\omega, T)$ represent the homodyne and heterodyne signal, respectively. α_{e,e_j} is the Raman polarizability and the field $E^*_{LO}(\bar{\omega})$ contains a certain phase of local oscillator. The parameter $\varepsilon_\omega = \sqrt{2\pi\omega/V}$ out of summation, and N denotes the number of choices for electrons transition between ground and excited states. The Raman line-shape function of these two types molecules, $f_j(T)$, can be expressed as

$$f_j(T) = 2\pi \int d\tau \rho_j(\tau) E(\tau - T) \tag{4}$$

where the Gaussian pulse shape $E(\omega - \omega_0) = E_0 e^{-(\omega-\omega_0)^2/2\sigma_0^2}$ in Fourier spectrum. Up till now, two types of molecules share the theoretical algorithms with generality. The difference between them is the vibrational coherence mode ρ_j. For 4-oriented amino-4'nitrostilbene, the vibrational coherence mode can be described as

$$\rho_{e,e_j}(T) = \rho_{e,e_j}(0) \sum_{n=0}^{\infty} (e^{-F_j} \frac{F_j^n}{n!}) e^{-i(\tilde{\omega}_{e,e_j} + n\nu_h)T - D_j T^2} \tag{5}$$

Nevertheless, the vibrational coherence mode of methane should be

$$\rho_{bg}(T) = \rho_{bg}(0) e^{-(i\omega_{bg} + \gamma_{bg})T} \tag{6}$$

$(e^{-F_j} \frac{F_j^n}{n!})$ is the Franck-Condon factor. F_j and D_j quantifies the coupling of excitons to low- and high-frequency vibrations, respectively [28]. γ_{bg}^{-1} quantifies the dephasing of the vibrational coherence. The multiple Raman resonance $\omega - \omega_{pr} = \omega_{bg} = \tilde{\omega}_{e,e_j} + n\nu_h$ in which $n\nu_h$ is the vibration splitting and the intensity is governed by the Franck-Condon factor quantifying the vibronic interaction. Inserting these vibrational coherence modes into Raman line-shape function,

$$f^{(am)}_{e,e_j}(T) = \frac{2\pi E_0 \sigma_0}{\sqrt{\sigma_0^2 + 2D_j}} \sum_{n=0}^{\infty} \rho^{(n)}_{e,e_j}(T) e^{-\frac{(\omega - \omega_0 - \tilde{\omega}_{e,e_j} - n\nu_h + 2iD_j T)^2}{2(\sigma_0^2 + 2D_j)}} \tag{7}$$

$$f^{(mt)}_{bg}(T) = 2\pi \sum_b \rho_{bg}(T) e^{-\frac{(\omega - \omega_{pr} - \omega_{bg} - i\gamma_{bg})^2}{2\sigma_0^2}} \tag{8}$$

We finally get the theoretical homodyne and heterodyne signal with respect to frequency and dephasing parameters. Looking at Eq. 7 and Eq. 8, we find that the Raman line-shape functions are Gaussian with respect to $\omega - \omega_{pr}$, and exponentially decay with time evolution. The neighborhood rasonances overlap with each other and go beyond Rayleigh limit. Thus, the oscillating signal produces beat frequencies, which leads to insufficient resolution in the time and frequency domains. However, if we manually separate them by calculating each Raman resonance one by one and merely sum up all of peak data with Gaussian bandwidth equivalent to T_s^{-1} ($T_s(T_i)$ is the time delay between the photons in s (reference) arm and the SPDC pump field, due to the group velocity dispersion in the nonlinear crystal), we will achieve an ideally time and frequency superresolved Raman spectroscopy.

Simulation results. - In order to justify the feasibility of SpectroNet, we simulate, we simulate the ultrafast coherent Raman spectroscopy with 4-oriented amino-4'nitrostilbene for excited states and methane for ground state. Analog to the training process, we input the classical FAST CARS results into the SpectroNet, and the predicted time- and frequency-resolved spectra are shown in Fig. 2 (a), (b) and Fig. 3 (a), (b). It should be noted that due to the wide probe bandwidth compared to energy level differences, the classical FAST CARS frequency information stick to each other, as is shown in Fig. 2 (c), (d) and Fig. 3 (c), (d). Both HM and HD signals are blurred in frequency domain and beat in time domain. Nevertheless, when we input them into SpectroNet, the superresolved spectral are achieved without frequency overlap and beating.

To indicate the dynamics of electronics evolution, we present the frequency slices in Fig. 2 (e), (f) and Fig. 3 (e), (f). The predicted results are ploted in blue lines, and the ground truths generated in the same way as labels are shown in red lines. Namely, from the HD and HM slices we claim that the SpectroNet owns great ability at retrieving electronics coherence and damping rates, respectively.

To see the precision of frequency prediction, we plot the time slice at $T = 0, 5, 10, 15, 20, 25, 30 fs$ for 4-oriented amino-4'nitrostilbene and $T = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 ps$ for methane in Fig. 4 with dash lines. The ground truths in solid lines are generated by the same process as labels in training as references. The predicted FAST CARS results via SpectroNet are perfectly fit with labels, particularly for 4-oriented amino-4'nitrostilbene whose electronic coherence are more complex and almost stick to each other. This indicates that the SpectroNet owns potential to predict spectroscopic signals for molecules with more complicated electronic coherence systems.

EXPERIMENT

(HZH)

SUMMARY

X. N., Y. M., and Z. H. contribute equally to this work.

* taopeng@tamu.edu
† zzhan26@cityu.edu.hk

[1] F. Krausz and M. Ivanov, Reviews of modern physics **81**, 163 (2009).
[2] O. Roslyak, C. A. Marx, and S. Mukamel, Physical Review A **79**, 033832 (2009).
[3] K. E. Dorfman, F. Schlawin, and S. Mukamel, The journal of physical chemistry letters **5**, 2843 (2014).

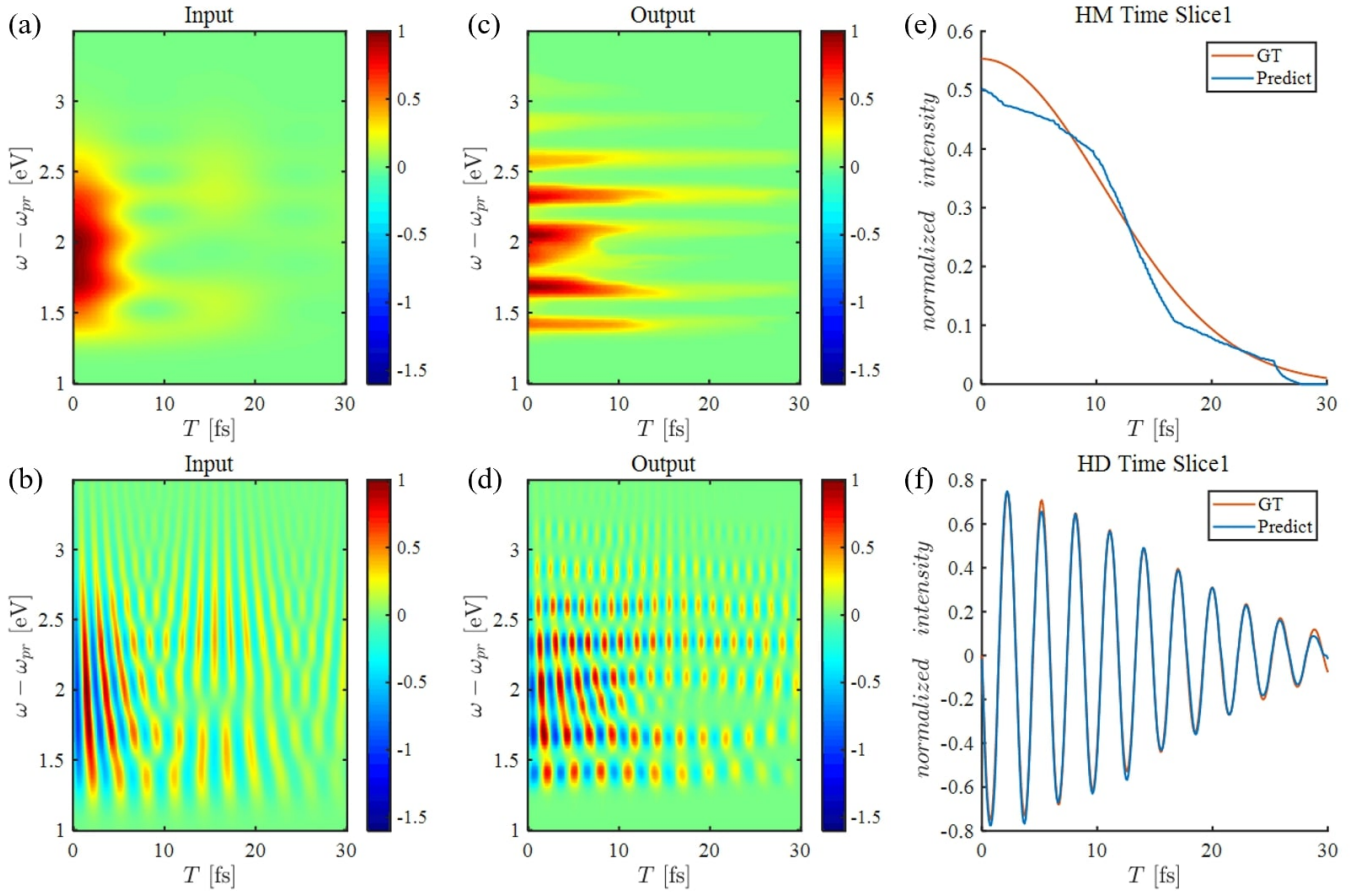

FIG. 2: (a) Homodyne-detected FAST CARS without SpectroNet for time-evolving electronic coherence versus the delay T between the pump-probe pulses. (b) same as (a) but for heterodyne detection. (c,d) Same as (a,b) but using SpectroNet. (e,f) 1D plot on one frequency $\omega - \omega_{pr} = 2.32eV$ slice to see the time-evolution of electronic coherence. Parameters are taken from 4-oriented amino-4'nitrostilbene, i.e., $\omega_e = 7.1$eV, $\omega_{e_1} = 5.3$eV, $\omega_{e_2} = 5.7$eV, $v_h = 0.26$eV, $F_1 = 2.2$, $F_2 = 1.3$, $D_1^{-1/2} = 30$fs, and $D_1^{-1/2} = 20$fs.

[4] Z. Zhang, T. Peng, X. Nie, G. S. Agarwal, and M. O. Scully, arXiv preprint arXiv:2106.10988 (2021).
[5] Z. Zhang, P. Saurabh, K. E. Dorfman, A. Debnath, and S. Mukamel, The Journal of chemical physics **148**, 074302 (2018).
[6] S. Asban, K. E. Dorfman, and S. Mukamel, Proceedings of the National Academy of Sciences **116**, 11673 (2019).
[7] L. Ye, J. R. Rouxel, D. Cho, and S. Mukamel, Proceedings of the National Academy of Sciences **116**, 395 (2019).
[8] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Reviews of Modern Physics **91**, 045002 (2019).
[9] L.-F. Arsenault, A. Lopez-Bezanilla, O. A. von Lilienfeld, and A. J. Millis, Physical Review B **90**, 155136 (2014).
[10] X. Gao and L.-M. Duan, Nature communications **8**, 1 (2017).
[11] G. Carleo and M. Troyer, Science **355**, 602 (2017).
[12] D.-L. Deng, Physical review letters **120**, 240402 (2018).
[13] F. Noé, S. Olsson, J. Köhler, and H. Wu, Science **365** (2019).
[14] A. A. Gentile, B. Flynn, S. Knauer, N. Wiebe, S. Paesani, C. E. Granade, J. G. Rarity, R. Santagati, and A. Laing, Nature Physics , 1 (2021).
[15] D. Pile, Nature Photonics **11**, 539 (2017).
[16] A. Sinha, J. Lee, S. Li, and G. Barbastathis, Optica **4**, 1117 (2017).
[17] Y. Li, Y. Xue, and L. Tian, Optica **5**, 1181 (2018).
[18] G. Barbastathis, A. Ozcan, and G. Situ, Optica **6**, 921 (2019).
[19] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Nature **549**, 195 (2017).
[20] V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, Nature **567**, 209 (2019).
[21] L. Hu, S.-H. Wu, W. Cai, Y. Ma, X. Mu, Y. Xu, H. Wang, Y. Song, D.-L. Deng, C.-L. Zou, *et al.*, Science advances **5**, eaav2761 (2019).
[22] R. Strack, Nature methods **15**, 403 (2018).
[23] Y. Li, Y. Ni, R. A. Croft, T. Di Matteo, S. Bird, and Y. Feng, Proceedings of the National Academy of Sciences **118** (2021).
[24] H. Lin, H. J. Lee, N. Tague, J.-B. Lugagne, C. Zong, F. Deng, J. Shin, L. Tian, W. Wong, M. J. Dunlop, *et al.*, Nature communications **12**, 1 (2021).
[25] R. Santagati, A. A. Gentile, S. Knauer, S. Schmitt, S. Paesani, C. Granade, N. Wiebe, C. Osterkamp, L. P. McGuinness, J. Wang, *et al.*, Physical Review X **9**, 021019 (2019).
[26] A. Radovic, M. Williams, D. Rousseau, M. Kagan, D. Bonacorsi, A. Himmel, A. Aurisano, K. Terao, and T. Wongjirad,

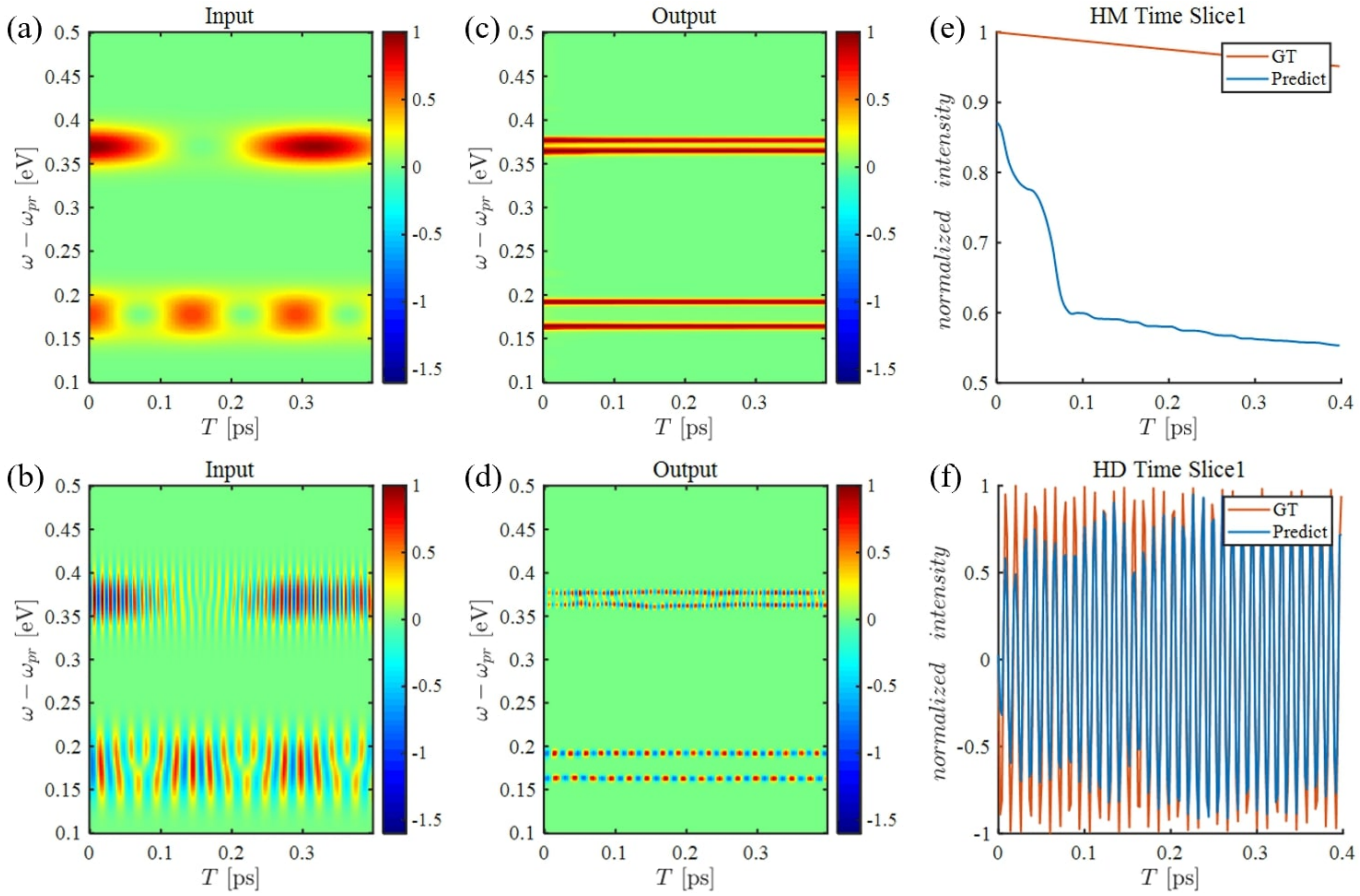

FIG. 3: (a) Homodyne-detected FAST CARS without SpectroNet for time-evolving electronic coherence versus the delay T between the pump-probe pulses. (b) same as (a) but for heterodyne detection. (c,d) Same as (a,b) but using SpectroNet. (e,f) 1D plot on one frequency $\omega - \omega_{pr} = xxxeV$ slice to see the time-evolution of electronic coherence. Parameters are taken from methane, i.e., the wave numbers are $\lambda_1^{-1} = 2914cm^{-1}$, $\lambda_2^{-1} = 1533cm^{-1}$, $\lambda_3^{-1} = 2914cm^{-1}$, and $\lambda_4^{-1} = 2914cm^{-1}$; the dephasing of vibrational coherences are $\gamma_1^{-1} = 16ps$, $\gamma_2^{-1} = 20ps$, $\gamma_3^{-1} = 15ps$, and $\gamma_4^{-1} = 22ps$.

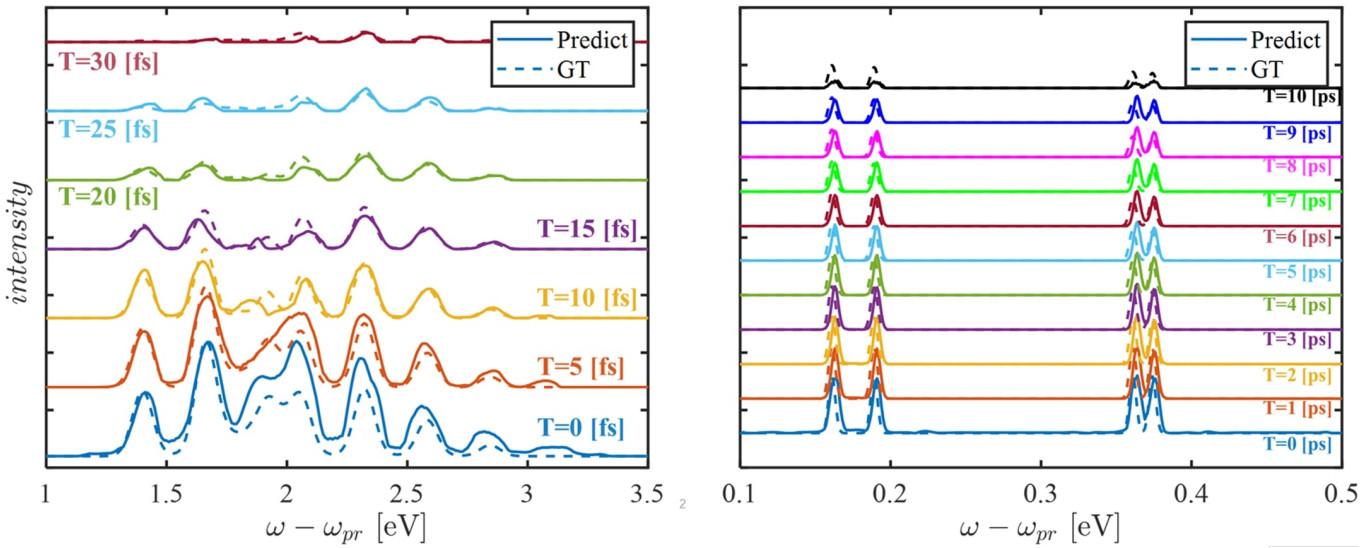

FIG. 4: (a) 1D plot of 4-oriented amino-4'nitrostilbene at $T = 0, 5, 10, 15, 20, 25, 30fs$. (b) 1D plot of methane at $T = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10ps$. The solid lines represent the predicted FAST CARS signals via SpectroNet while the dash lines indicate the reference signals which we define as ground truths/labels.

Nature **560**, 41 (2018).

[27] O. Ronneberger, P. Fischer, and T. Brox, in *International Conference on Medical image computing and computer-assisted intervention* (Springer, 2015) pp. 234–241.

[28] P. F. Barbara, T. J. Meyer, and M. A. Ratner, The Journal of Physical Chemistry **100**, 13148 (1996).