

Machine learning approach to reconstruct dephasing time from solid HHG spectra

Jiahao Liu(刘佳豪), Xi Zhao(赵曦), Jun Wang(王俊), and Songbin Zhang(张松斌)
Citatio: Chin. Phys. B, 2025, 34 (**9**): 097804. DOI: 10.1088/1674-1056/ade064
http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Twin pulses of THz generation from a nonlinear crystalline quartz by femtosecond laser pulse

Xiangmei Dong(董祥美), Dan-Ni Li(李丹妮), Zuan-Ming Jin(金钻明), Hui-Ping Zhang(张慧萍), Hong-Guang Li(李宏光), Shao-Hui Wu(吴少晖), Yan Peng(彭滟), Yiming Zhu(朱亦鸣), and Songlin Zhuang(庄松林)
Chin. Phys. B, 2025, 34 (**9**): 094211. DOI: 10.1088/1674-1056/addbcd

Strain modulation of second harmonic generation in new tetrahedral transition metal dichalcogenide monolayers

Hu Chen(陈虎), Shi-Qi Li(李仕琪), Yuqing Wu(吴雨晴), Xiaozhendong Bao(鲍晓振东), Zhijie Lei(雷志杰), Hongsheng Liu(柳洪盛), Yuee Xie(谢月娥), Junfeng Gao(高峻峰), Yuanping Chen(陈元平), and Xiaohong Yan(颜晓红)
Chin. Phys. B, 2025, 34 (**8**): 084206. DOI: 10.1088/1674-1056/addaa3

Simultaneous second and third harmonics generation in periodically poled lithium niobate: Coupling and competition

Junming Liu(刘峻铭), Liqiang Liu(刘励强), Lihong Hong(洪丽红), and Zhiyuan Li(李志远)
Chin. Phys. B, 2025, 34 (**8**): 084203. DOI: 10.1088/1674-1056/add24a

Heterogeneous integration of silicon nitride photonics and CVD-grown WS_2 for second harmonic generation

Xuhang Jia(贾旭航), Bangren Xu(许邦仁), Jieyi Luo(罗劼叙), Ning Liu(刘宁), Yuhang Wang(王宇航), Biyuan Zheng(郑弼元), Wei Xu(徐威), and Ken Liu(刘肯)
Chin. Phys. B, 2025, 34 (**7**): 074205. DOI: 10.1088/1674-1056/adc7f7

Multifunctional disk device for optical switch and temperature sensor

Bian Zhen-Yu (卞振宇), Liang Rui-Sheng (梁瑞生), Zhang Yu-Jing (张郁靖), Yi Li-Xuan (易丽璇), Lai Gen (赖根), Zhao Rui-Tong (赵瑞通)
Chin. Phys. B, 2015, 24 (**10**): 107801. DOI: 10.1088/1674-1056/24/10/107801

SPECIAL TOPIC — Ultrafast physics in atomic, molecular and optical systems

Machine learning approach to reconstruct dephasing time from solid HHG spectra

Jiahao Liu(刘佳豪)[1], Xi Zhao(赵曦)[1,†], Jun Wang(王俊)[2,‡], and Songbin Zhang(张松斌)[1,§]

[1] *School of Physics and Information Technology, Shaanxi Normal University, Xi'an* 710062, *China*
[2] *Institute of Atomic and Molecular Physics, Jilin University, Changchun* 130012, *China*

(Received 22 February 2025; revised manuscript received 1 May 2025; accepted manuscript online 4 June 2025)

The dephasing time T_2 is a fundamental parameter that characterizes the coherence of electronic states and electron–phonon interactions in condensed matter physics. Accurate measurement of T_2 is essential for elucidating ultrafast electronic and phononic processes, which are crucial for the development of advanced electronic, optoelectronic, and quantum devices. However, due to the complexity of solid-state systems with their intricate band structures and strong many-body interactions, reconstructing T_2 remains a long-term challenge for both condensed matter physics and optical science. In this work, we introduce a machine learning (ML) approach to retrieve T_2 from the high-order harmonic generation (HHG) spectrum resulting from the interaction between a strong infrared (IR) laser pulse and solid-state material. The consistency between the experimental and reconstructed HHG spectra validates the efficiency of our scheme. Our ML method offers two key advantages: first, it does not require stringent experimental conditions, and second, the optimization process is fully automated and more reliable than empirical selection of dephasing time values. The ability of our method to reconstruct dephasing time from solid HHG spectra provides a powerful tool for probing the intrinsic properties of materials under extreme conditions. Besides, our method provides another significant advantage, which offers a direct approach to calculating the quantum tunneling time of carriers between different energy bands under light-induced excitation.

Keywords: high-order harmonic generation, dephasing time, machine learning

PACS: 78.47.jb, 42.65.Ky, 43.60.Np

DOI: 10.1088/1674-1056/ade064 **CSTR:** 32038.14.CPB.ade064

1. Introduction

Dephasing time T_2, also defined as coherence time, is a pivotal concept in condensed matter physics,[1–3] representing the timescale over which a quantum system can maintain its phase coherence before interactions with its environment induce phase randomization.[4–6] Understanding and controlling dephasing time is essential for the development of quantum technologies, as it directly impacts the performance and stability of quantum bits (qubits) in quantum computers, where phase coherence is crucial for accurate computation and information processing.[7,8] Additionally, the study of T_2 provides insights into the fundamental interactions between particles and their environments,[9,10] helping to unravel complex behaviors in systems such as superconductors,[11] topological insulators,[12] and quantum dots.[13] Specifically, the dephasing time in plasma resonance can affect the resonance energy linewidth, which is achieved by decoherence of the plasma in the resonance state,[2] and additionally, the dephasing time affects the spectral shape by affecting the phase matching of the longer and shorter trajectories in the solid-state high harmonic generation.[14] Applications of these insights extend to developing novel materials with tailored quantum properties, improving sensors[15] and transistors,[16] and advancing our understanding of quantum phase transitions and other collective phenomena in condensed matter systems.[17]

However, experimental measurement or theoretical calculation of T_2 involves a variety of challenges. The primary difficulty of experimental measurement is isolating the system from environmental noise and disturbances, which can significantly shorten the observed dephasing time. Achieving this isolation requires advanced experimental setups with precise control over temperature, electromagnetic interference, and material impurities. Theoretically, the challenge lies in the intrinsic complexity of the interactions within the system itself, such as electron–phonon interactions,[18,19] spin–orbit coupling,[20,21] defects in materials,[22] or even propagation effects,[23,24] all of which contribute to decoherence in ways that are difficult to predict and model accurately. Overcoming these challenges is crucial for advancing quantum technologies, as precise knowledge and control of dephasing time are vital for the development of stable and efficient quantum devices, including quantum computers and sensors.[25,26]

In this work, we combine the technology of solid

†Corresponding author. E-mail: zhaoxi719@snnu.edu.cn
‡Corresponding author. E-mail: wangjun86@jlu.edu.cn
§Corresponding author. E-mail: song-bin.zhang@snnu.edu.cn

© 2025 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

HHG[27–33] and ML[34,35] to provide an approach to directly reconstruct T_2 from experimental solid HHG spectra. HHG, which is famous for its potential applications in the development of novel light sources[36,37] for industrial and medical purposes, also holds promise for advanced imaging techniques,[38–41] such as high-resolution spectroscopy and microscopy. Solid HHG spectra have already been used to probe the electronic structure and properties of materials with exceptional spatial and temporal resolution, leading to new discoveries in condensed matter physics and material science.[42–44] On the other hand, ML is revolutionizing condensed matter physics and ultrafast science by offering powerful tools to analyze complex data, identify patterns, and accelerate discovery.[45] In condensed matter physics, ML algorithms are employed to predict material properties, optimize experimental conditions, and uncover hidden phases of matter.[46] Techniques such as neural networks and reinforcement learning help in modeling intricate interactions within materials, enabling the discovery of new superconductors, topological insulators, and other exotic states of matter.[47,48] In ultrafast science, ML enhances the interpretation of vast amounts of data generated by advanced spectroscopic techniques.[49] For instance, ML can be used to reconstruct Wigner time delay or the temporal profile of attosecond pulses from experimental photon-ionization electron spectra (i.e., the so-called "streaking trace") and improve the resolution of time-resolved measurements.[50–54] Additionally, ML assists in the design and control of ultrafast experiments by predicting optimal laser parameters and experimental setups, thereby reducing the trial-and-error process.

Solid HHG has been employed as a probing technique to reconstruct band dispersion and momentum-space transition dipole moments.[55–57] Most of these approaches are based on calculating HHG spectra using the solid-state SFA formalism. However, approximations are commonly introduced to avoid direct numerical integration of the SFA expression. For example, in Ref. [55], it is assumed that electronic transitions occur only at a single initial momentum, thereby neglecting the contributions from other quantum trajectories. While such approximations allow for a straightforward relation between the physical quantity of interest and the solid-state HHG spectrum, they inevitably compromise the accuracy of the reconstruction. The ideal approach for retrieving physical parameters from HHG spectra involves a complete numerical evaluation of the solid-state SFA integral, followed by an optimization to infer the quantity from the HHG spectrum. This process naturally aligns with the framework of optimization problems solvable by machine learning techniques, which is why we adopt a machine learning scheme to reconstruct the dephasing time.

As far as we know, previous theoretical calculations of solid HHG have required the empirical selection of a specific T_2 value to compare with experiments, typically ranging from 1 femtosecond (fs) to several hundred fs.[14,23,58–60] Although some experimental works have managed to correlate the dephasing time with a specific characteristic of solid-state HHG under certain laser conditions, thereby measuring the dephasing time,[14] there have been no reports of reconstructing the dephasing time directly from solid-state HHG spectra using ML approaches. The basic idea of our method is straightforward: in the realm of solid HHG, dephasing time strongly influences the quantum tunneling dynamics of the carrier wavepacket between the valence band (VB) and conduction band (CB) and thus the efficiency of HHG spectra yield. Then the T_2 information is encoded in such HHG spectra and we have the opportunity to retrieve T_2 from solid HHG spectra with the ML method.

The advantages of using ML to reconstruct dephasing time are twofold. First, it requires less stringent conditions for the laser field in experiments and does not necessitate complex experimental setups for measurement. Second, the optimization process is fully automated, making it more reliable than the empirical selection of dephasing time values. Our results also demonstrate this reliability.

2. Theoretical methods

The flowchart of our ML algorithm can be found in Fig. 1. The input for the machine learning network is $H_{\rm exp}(T_2,N)$, which can be either an experimental spectrum or a theoretical spectrum calculated using a known T_2 via the SFA or SBE methods. The output data includes the reconstructed T_2 and the HHG spectra $H_{\rm guess}(T_2,N,\beta)$ recalculated using this reconstructed T_2. In the first step of the machine learning network, a random matrix of T_2 is generated by our ML algorithm. Then the $H_{\rm guess}(T_2,N,\beta)$ are calculated using the guessed T_2 and the merit function ${\rm erf}(T_2,\beta)$ are calculated. In the third step, if the merit function converges sufficiently, then we output the optimized T_2. Otherwise, we update the matrix of T_2 and calculate the $H_{\rm guess}(T_2,N,\beta)$ and the merit function until the merit function converges sufficiently.

In the first step, a random matrix of T_2 is generated by our ML algorithm. Then the HHG spectrum, which is set to be our fitness function, and the merit function are calculated in the second step. In the third step, if the merit function converges sufficiently, then we output the optimized T_2. Otherwise, we update the matrix of T_2 and calculate the HHG spectrum and the merit function until the merit function converges sufficiently.

The performance of ML algorithms in physics applications primarily depends on two factors: the efficiency and accuracy of the fitness function calculation and the effectiveness

of the optimization algorithm itself. Let's first look at the fitness function, the interband solid HHG between CB and VB is[14]

$$J_{\rm ter}(\Omega) = \Omega \int_{-\infty}^{+\infty} {\rm d}t_{\rm r}\, {\rm e}^{-{\rm i}\Omega t_{\rm r}} \int_{\overline{\rm BZ}} {\rm d}\boldsymbol{p} d^*(\boldsymbol{p}-\boldsymbol{A}(t_{\rm r})) \times \int_{-\infty}^{t_{\rm r}} {\rm d}t_{\rm i} \boldsymbol{E}(t_{\rm i}) d(\boldsymbol{p}-\boldsymbol{A}(t_{\rm i})) \times {\rm e}^{{\rm i}S(\boldsymbol{p}+\boldsymbol{A}(t_{\rm r}),t_{\rm i},t_{\rm r})-(t_{\rm r}-t_{\rm i})/T_2} + {\rm c.c}, \quad (1)$$

where Ω is the HHG photon energy, $\boldsymbol{p} = \boldsymbol{k} - \boldsymbol{A}(t)$ is the canonical momentum, and $\boldsymbol{k}$ is the crystal momentum. $\boldsymbol{E}(t)$ is the electric field. $d_{mm'}[\boldsymbol{p}+\boldsymbol{A}(t)]$ is the transition dipole momentum (TDM) between the two bands. $S_{mm'}(\boldsymbol{p},t) = \int_{-\infty}^{t} \varepsilon_{mm'}[\boldsymbol{p}+\boldsymbol{A}(t')]{\rm d}t'$ is the classical action. $\varepsilon_{mm'} = E_m - E_{m'}$ is the bandgap and E_m is the band dispersion. T_2 is the dephasing time,[58] which is the unknown function we aim to retrieve. Equation (1) is born from semiconductor Bloch equations (SBE). Keldysh approximation is made in Eq. (1): we assume that the population of CB is approximated to be 1, this approximation will be addressed later.

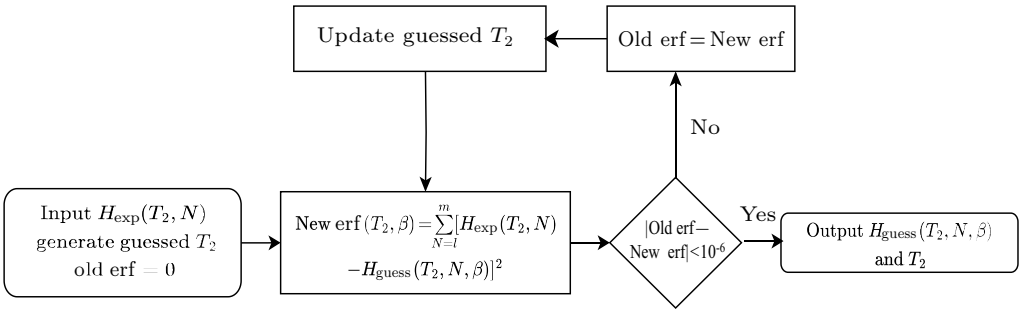


Fig. 1. The flowchart of the T_2 reconstruction.

We perform the 3-loop integration numerically. All the functions and iterative computations in our code are vectorized then only about 3 s is needed to perform a single Eq. (1) calculation with only one Intel CPU core, details of our vectorized calculation can be found in our source code.

For optimization, we assume that the electric field $\boldsymbol{E}(t)$, the transition dipole $d(p)$ and the band gap $\varepsilon(p)$ are known. The dephasing time T_2 is unknown and can be treated as a constant or a crystal momentum dependent function. Besides, an extra unknown parameter β is introduced during the optimization calculation to adjust the relative intensity of HHG spectrum. Moreover, additional parameters can also be introduced to shift the spectra, thus helping the two spectra to overlap better. The merit function, as mentioned in Fig. 1, is defined as

$$\mathrm{erf}(T_2,\beta) = \sum_{N=l}^{m} [H_{\rm exp}(T_2,N) - H_{\rm guess}(T_2,N,\beta)]^2, \quad (2)$$

where $H_{\rm exp}$ is the experimental HHG data, and $H_{\rm guess}$ is the optimized HHG spectrum. The HHG order N sums from l to m. If the merit function before and after the iteration is less than 1×10^{-6}, then the iteration stops and we output the reconstructed T_2.

The optimization algorithm in our method can be switched between genetic algorithm (GA),[62] particle swarm algorithm (PA),[63] and pattern search algorithm (PSA),[64] the comparison of these methods will be discussed in the next section.

Atomic units are used throughout this work, unless otherwise mentioned.

3. Significant of retrieval T_2

When studying solid HHG driven by strong laser–solid interactions, introducing the dephasing time term within the framework of the semiconductor Bloch equations is intended to well align theoretical results with experimental observations. However, a major debate surrounding the dephasing time lies in whether it has genuine physical significance or is merely a mathematical correction.[65,66]

In one of our previous independent articles,[67] we thoroughly discussed the physical effect of this term with our quantum complex analysis model, in which the energy gap $\varepsilon_{mm'}$, the ionization time $t_{\rm i}$, recombination time $t_{\rm r}$ as well as the momentum p are all treated as complex numbers, such an idea has already been implied by investigating the strong laser-induced tunneling time.[68,68] In this framework, the complex energy gap $\varepsilon_{\rm g}$ can be divided into a real part and an imaginary part

$$\varepsilon_{\rm g} = {\rm Re}\left(\varepsilon_{\rm g}\right) + {\rm i}{\rm Im}\left(\varepsilon_{\rm g}\right).$$

From the expression of the exponential term in Eq. (1), we can see that T_2 actually corresponds to the imaginary part of the bandgap. The influence of the term T_2 on the light-induced carrier dynamics can be seen by solving the six-variable complex saddle point equations for ionization time $t_{\rm i}$, recombination time $t_{\rm r}$, and momentum p:

$$\frac{\partial S}{\partial p} = \int_{t_{\rm i}}^{t_{\rm r}} \frac{\partial \varepsilon_{\rm g}}{\partial p} {\rm d}\tau = 0, \quad (3)$$

$$\frac{\partial S}{\partial t_{\rm i}} = \varepsilon_{\rm g} + \frac{\rm i}{T_2}, \quad (4)$$

$$\frac{\partial S}{\partial t_{\mathrm{r}}} = \varepsilon_{\mathrm{g}} + \frac{\mathrm{i}}{T_2} = \Omega. \tag{5}$$

In other words, accurately determining T_2 is essential for precisely solving the saddle point equations to obtain the tunneling time of carriers between different bands. To the best of our knowledge, this is currently the only approach for measuring the laser-induced quantum tunneling time between different energy bands in solids.

Here, we briefly summarize the key points from retrieval work: accurately describing the dephasing time term has at least two important physical implications. First, in our quantum complex analysis model, energy, like time and momentum, is treated as a complex variable. Consequently, the energy gap between the conduction and valence bands can also be extended into the complex plane, with the imaginary part of the energy gap corresponding to the so-called dephasing time. Second, once the dephasing time is accurately determined, it allows us to calculate the laser-induced quantum tunneling time between the conduction and valence bands by solving the complex saddle-point equations, regardless of whether T_2 has a true physical meaning or is merely a purely mathematical correction.

4. Results and discussion

In this section, we provide detailed results of the retrieval calculation. In Subsection 4.1, we use Eq. (1) to generate "experimental data" and perform the retrieval, serving as a test case to demonstrate the validity of our method. In Subsection 4.2, we conduct a similar calculation to that in Subsection 4.1, but with input HHG generated from SBE. In Subsection 4.3, we discuss the error effects from bandgap and dipole. Then, in Subsection 4.4, we present the retrieved T_2 directly from experimental HHG data. Finally, we compare the efficiencies from different algorithms.

The k-dependent T_2 is used to perform the retrieval for the following three reasons. First, in Ref. [69], the authors pointed out that the Coulomb interactions of electrons at different k points scatter differently, so the corresponding values of T_2 are also different at different k points. Second, we found that using k-dependent T_2 in the comparison between theoretical calculations and experimental results yields a closer match to the experimental data.[70] Third, in an independent study published in another paper, we discovered that T_2 plays a role in correcting the photo-induced quantum tunneling process between different energy bands of carriers during strong laser–solid interactions. Electrons at different k-points have different initial momenta, so the corresponding T_2 values at different k-points are also different.

4.1. The T_2 reconstruction from SFA input data

In this subsection, the target crystal is ZnO and a Gaussian laser pulse with a full width at half-maximum (FWHM) of 6 cycles, a peak laser intensity $I = 7.442 \times 10^{11}$ W/cm^2, and a wavelength $\lambda = 3255$ nm is used. Equation (1) is used to generate the input HHG data.

In Fig. 2(a), we present the constant T_2 retrieval. The input T_2 varies from 0.5 fs to about 1100 fs, and the input and output T_2 are almost exactly equal, with errors $\left|T_{2,\mathrm{in}} - T_{2,\mathrm{out}}\right|/T_{2,\mathrm{in}}$ all being less than 1×10^{-6}. Specifically, we apply the particle swarm algorithm to optimize the reconstruction of the constant T_2. Moreover, the reconstruction of each constant T_2 only takes roughly 10 min with four CPU cores in parallel computing.

Figures 2(b) and 2(c) investigate the ability of our code to reconstruct a k-dependent T_2. Such dephasing times have already been employed in some studies that considered Coulomb interactions scattering polarization influencing the excited electron–hole distributions.[70] Two types of input $T_2(k)$ and corresponding HHG (blue and red solid lines in Figs. 2(b) and 2(c)) are tested. The input T_2 in the red line case has a periodic form $T_2(k) = 1.8 + \cos(ak) + 0.2\cos(2ak)$, while the blue line case has non-periodic $T_2(k) = 1 + \frac{1}{2+\mathrm{e}^{100|k|-5}}$. We extend T_2 with sets of unknown parameters x_i, y_i, $i = 1, 2, \ldots$ with different basis functions: $T_2(k) = \sum_{i=0} x_i \cos(i \cdot a \cdot k + y_i)$ for the periodic case, and $T_2(k) = \sum_{i=1}(x_{1i} + \mathrm{e}^{y_{1i}|k|-z_{1i}})/\sum_{i=1}(x_{2i} + \mathrm{e}^{y_{2i}|k|-z_{2i}})$ for the non-periodic case with unknown parameters x_{1i}, y_{1i}, z_{1i}, x_{2i}, y_{2i}, z_{2i}, $i = 1, 2, \ldots$.

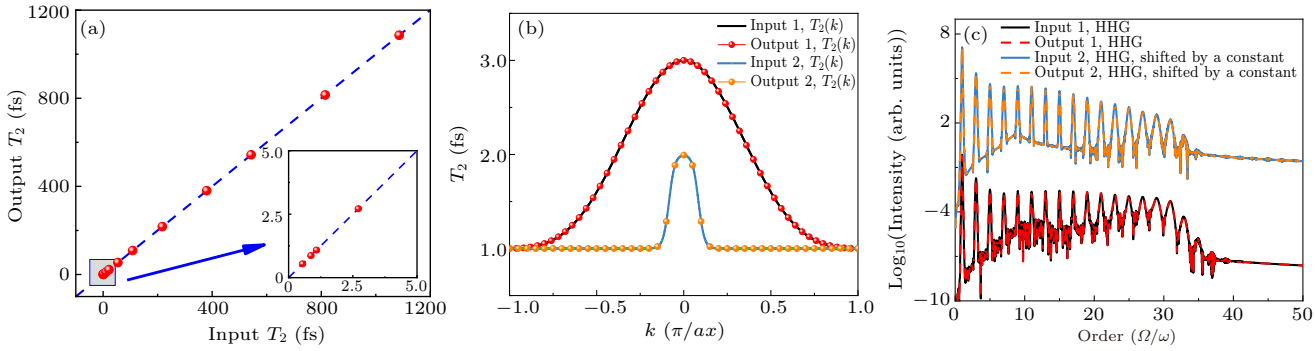


Fig. 2. The comparison of the dephasing time and the spectrum between the input and the output with the input spectrum caculated by SFA. (a) The comparison of constant T_2 between the input and the output. (b) The comparison of constant $T_2(k)$ between the input and the output. (c) The comparison of the spectrum between the input and the output.

These figures show that good agreement between the input and output HHG spectra can be achieved, indicating the validity of our method. GA with an initial population of 200 is used to optimize the process, and convergence is achieved after about one hour of running on a 72-CPU-core server.

4.2. The T_2 reconstruction from SBE input data

As we mentioned in Section 2, we assume the population of the CB is kept at 1 in Eq. (1). This approximation introduces an error in the HHG spectra that needs to be addressed. Here, we use the SBE, a widely used method in studying solid HHG, to generate input HHG as the "experimental data", and use Eq. (1) as our fitness function to perform the retrieval.

The target, laser parameters, input dephasing time, and retrieval processes in this subsection are the same as those in Subsection 4.2. The parameter β is treated as an unknown variable in this section.

We present the constant T_2 retrieval in Fig. 3(a). The input T_2 ranges from 0.5 fs to 100 fs. For T_2 values smaller than 60 fs, the retrieved and input T_2 match well, but for T_2 values larger than 60 fs, the error becomes evident. The reason for the failure of our reconstruction method at larger T_2 values is that HHG becomes less sensitive to changes in T_2 when T_2 is relatively large, making the reconstruction calculation difficult to converge. Although some works claim that the dephasing time can reach the order of 100 fs under certain conditions, most current theoretical calculations use smaller dephasing time values. Our method achieves high accuracy for dephasing times less than 60 fs.

For the k-dependent T_2 retrieval, we use GA to perform the optimization calculations. The calculation converges within one hour using 30 CPU cores. As seen in Fig. 3(b), the input $T_2(k)$ and the output $T_2(k)$ match well. Additionally, the output spectrum is consistent with the input spectrum in the plateau area shown in Fig. 3(c). The success in reconstructing dephasing time from SBE-generated input HHG spectra verifies the feasibility of our method.

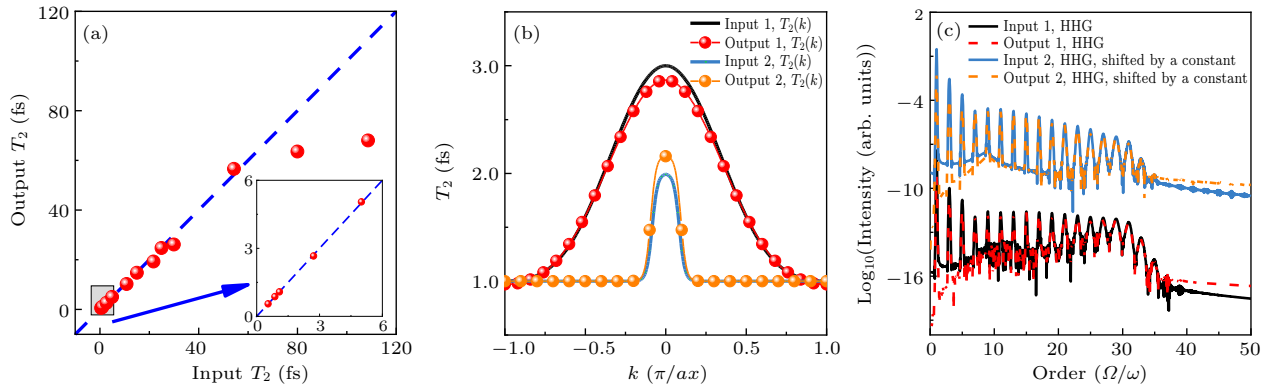


Fig. 3. The comparison of the dephasing time and the spectrum between the input and the output with the input spectrum caculated by SBEs. (a) The comparison of constant T_2 between the input and the output. (b) The comparison of constant $T_2(k)$ between the input and the output. (c) The comparison of the spectrum between the input and the output.

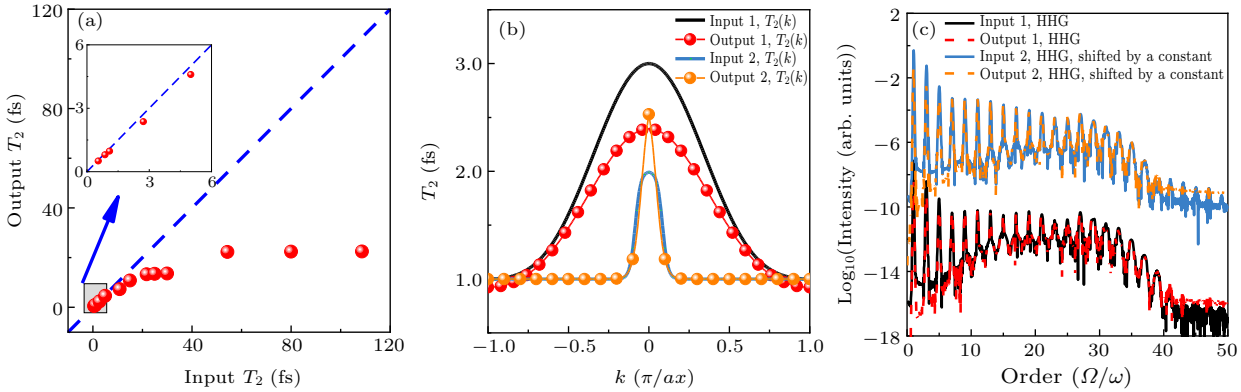


Fig. 4. The comparison of the dephasing time and the spectrum between the input and the output with the input spectrum caculated by SBEs for the peak of the laser intensity $I = 2.251 \times 10^{12}$ W/cm^2. (a) The comparison of constant T_2 between the input and the output. (b) The comparison of constant $T_2(k)$ between the input and the output. (c) The comparison of the spectrum between the input and the output.

However, since our fitness function is based on the Keldysh approximation, which may not work well for cases with high-intensity lasers, we address this issue by applying a higher intensity laser of $I = 2.251 \times 10^{12}$ W/cm^2 to generate input HHG from SBE, keeping the other laser parameters unchanged. GA is used for the optimization. Figure 4 shows the results. The errors for constant T_2 values from 0.5430 fs to 5 fs are less than 13%, but they become significant for larger T_2 values. The reconstructed k-dependent $T_2(k)$ and corresponding HHG spectrum are generally in agreement with the inputs

shown in Figs. 4(b) and 4(c) as long as $T_2(k)$ values are less than 5 fs. Thus, for T_2 values of a few femtoseconds, we can still reconstruct T_2 in high-intensity laser cases in general.

4.3. Error effect

An important criterion for evaluating the success of a numerical method is its ability to reproduce experimental results. Currently, simulations of high-harmonic dynamics in strong laser-solid interactions, based on the SBE framework with dephasing time corrections, can successfully match experimental results. In other words, accurately describing the dipole, energy gap, and T_2 between different energy bands is crucial for determining the performance of the SBE theory. In our method, we assume that the dipole and energy gap are known, while T_2 is unknown. Therefore, errors in the dipole and energy gap will affect our final reconstruction results. Additionally, noise in HHG introduced experimentally due to various factors may also impact the reconstruction. In this section, we discuss the effects of these errors on the reconstruction results.

We first tested a numerical simulation where a random error function $R(\Omega)$ was applied to the HHG spectrum at each photon energy Ω. We multiply an Ω-dependent random function $R(\Omega)$ to the HHG spectrum $J(\Omega)$ at each Ω, such that $J_{\text{new}}(\Omega) = J(\Omega) \cdot R(\Omega)$, where $R(\Omega)$ takes a random value between 0.9 and 1.1 in Fig. 5. Then J_{new} is used as the input HHG data, keeping the laser and target conditions the same as in Subsection 4.1. From Fig. 5(b), we can see that even with a 10% error added to the HHG spectrum, our code can still accurately reconstruct T_2. This test further validates the accuracy of our code.

In our code, we assume the TDM $d(k)$ and the energy bandgap $\varepsilon(k)$ are already known. Before we perform the retrieval calculation, $d(k)$ and $\varepsilon(k)$ are calculated from *ab initio* code packages in advance. Thus, errors from the simulation of $d(k)$ and $\varepsilon(k)$ would be introduced if we apply our code to retrieve T_2 from experimental data. Next, we will discuss the effects of errors from $d(k)$ and $\varepsilon(k)$.

To test the tolerance of our code to errors in the dipole function, we stretched the dipole $d(k)$ by 10% along the k-axis (as shown in Fig. 6(a)). The unstretched $d(k)$ was used to generate the "experimental HHG", while the stretched dipole function was used for the reconstruction calculations. All other laser parameters were kept consistent with those in Fig. 3. From Figs. 6(b) and 6(c), it can be seen that even with a 10% error introduced in the dipole function, our code can still perform the reconstruction calculations reasonably well.

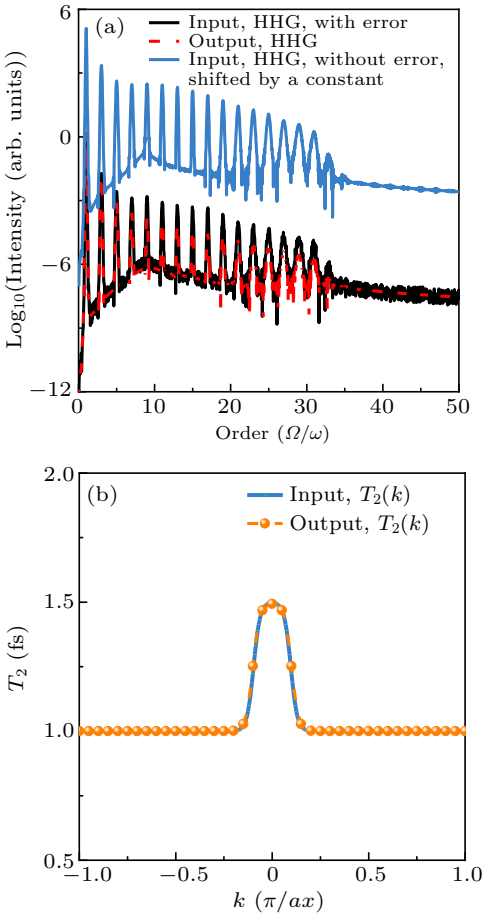


Fig. 5. The comparison of the dephasing time between the input and the output considering the ±10% error in input HHG. (a) The comparison of the spectrum between the input and the output. (b) The comparison of the dephasing time between the input and the output.

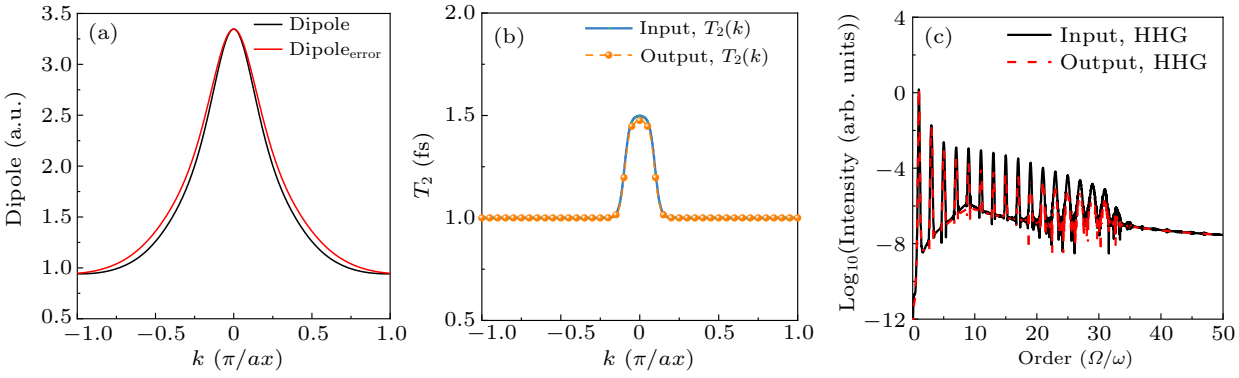


Fig. 6. The comparison of the dephasing time and the spectrum between the input and the output considering the ±10% error in the dipole of the input HHG. (a) The comparison between the dipole with error and the dipole without error. (b) The comparison of constant $T_2(k)$ between the input and the output. (c) The comparison of the spectrum between the input and the output.

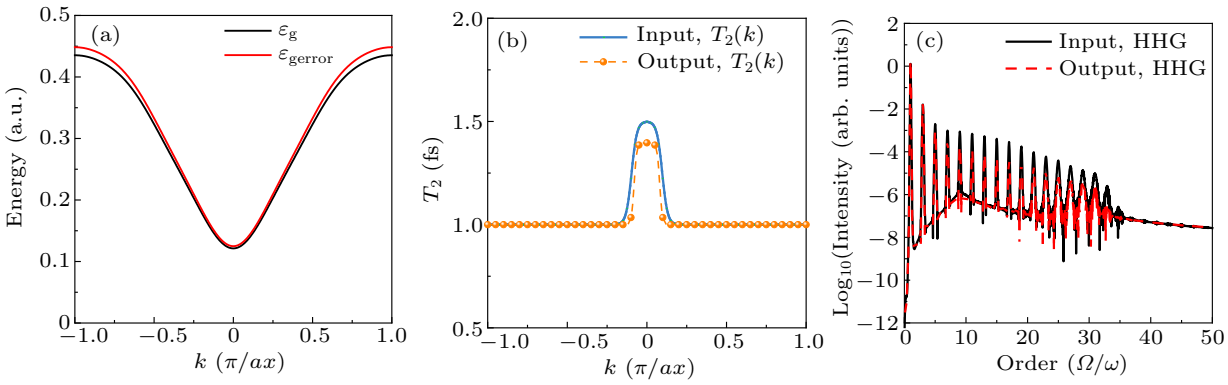


Fig. 7. The comparison of the dephasing time and the spectrum between the input and the output considering the ±5% error in the bandgap of the input HHG. (a) The comparison between the bandgap with error and the bandgap without error. (b) The comparison of constant $T_2(k)$ between the input and the output. (c) The comparison of the spectrum between the input and the output.

Similarly, we also tested the tolerance of our code to errors in the energy gap. We stretched the energy gap function vertically by 5% (as shown in Fig. 7(a)). The unstretched energy gap was used to generate the "experimental HHG", while the stretched energy gap was used for the reconstruction calculations. From Figs. 7(b) and 7(c), it can be seen that under a 5% error in the energy gap, the error in the reconstructed T_2 function is approximately 6% to 7%. Therefore, our code's tolerance for errors in the energy gap is slightly weaker than its tolerance for errors in the dipole function. This is because T_2 and the energy gap jointly determine the light-induced tunneling process of the carrier, making the reconstruction of T_2 more sensitive to the energy gap. Nevertheless, even with a 5% error in the energy gap, our code still performs well.

4.4. The T_2 reconstruction from experimental data

Retrieving T_2 directly from experimental HHG spectra, rather than relying on empirical predictions, is crucial for theoretical investigations of the interaction between strong laser fields and solid materials.

In this example, we apply our code to reconstruct T_2 from experimental HHG spectra of MoS_2, a material with a wide range of applications. The experimental data, dipole moment, and energy bandgap are sourced from Ref. [59]. We follow the suggestions from the references,[69,70] where T_2 is a k-dependent function. $T_2(k)$ is expanded as

$$T_2(k) = \sum_i x_i \cos\left[(i-1)\, a \cdot k\right],$$

where $\{x_i,\ i = 1, 2, 3, \ldots\}$ are the parameters we need to retrieve. The chosen form of $T_2(k)$ satisfies the periodicity of MoS_2 and the axial symmetry with $k = 0$ as the axis of symmetry along the K–K direction in the reciprocal space. Particle swarm algorithm (PA) is applied for the optimization calculations. The retrieved $T_2(k)$ and corresponding HHG spectra are shown in Figs. 8(a) and 8(b). The excellent agreement between the experimental and reconstructed HHG spectra indicates the accuracy of our method.

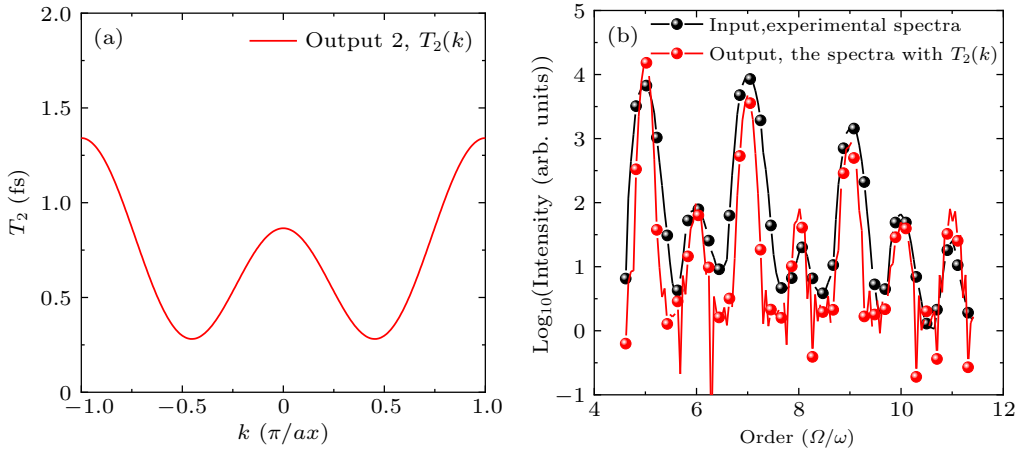


Fig. 8. (a) The output $T_2(k)$ from experimental data. (b) The comparison of experimental and output HHG spectra.

4.5. Comparison of different optimisation algorithms in reconstruction

To execute the reconstruction of T_2 more quickly and efficiently, it is imperative to compare several optimization algorithms including GA, PA, and PSA. GA and PA do not require initial points, while PSA does. We also note that the constraints can accelerate the convergence of the reconstruction.

For constant T_2 reconstruction, even if the initial point is not ideal, the reconstruction using PSA takes less time than that using other algorithms, and the output T_2 can be accurately obtained, as shown in Table 1. If there is no initial point, the reconstruction using PA performs better than using GA. Table 2 documents the comparison of different optimization algorithms for $T_2(k)$ reconstruction. Similarly, the reconstruction using PSA takes the shortest time with a poor initial point, and the output spectra match well. As the number of unknowns rises, constraints are added in the reconstruction using PA and GA. Constraints can be either the range of the unknowns or the linear inequality about the unknowns, which have a significant influence on the reconstruction.

Table 1. The comparison of different optimization algorithms for $T_2 = 5$ fs reconstruction.

	Particle swarm algorithm	Genetic algorithm algorithm	Pattern search algorithm
Initial point	no	no	$T_2 = 0.5$ fs
Constraints	no	no	no
CPUs	2	2	2
Output T_2	5.049 fs	0.3910 fs	5.049 fs
Error	0.98%	92.18%	0.98%
Finish time	2344.29 s	5128.47 s	443.66 s

Table 2. The comparison of different optimization algorithms for $T_2(k) = 1 + \frac{1}{2+e^{100|k|-5}}$ reconstruction.

	Particle swarm algorithm	Genetic algorithm	Pattern search algorithm
Initial point	no	no	$x = [10, 10, 10, 10]$
Constraints	$x_{\min} = [-5, -5, -200, -200]$, $x_{\max} = [-5, -5, -200, -200]$	$x_{\min} = [-5, -5, -200, -200]$, $x_{\max} = [-5, -5, -200, -200]$	no
CPUs	30	30	30
Final erf	32.55	91.38	73.12
Finish time (s)	1215.75	4379.59	1137.90

With the constraints being the range of the unknowns, the reconstruction using PA performs better than that using GA. However, if the range of unknowns is indeterminate or wide, GA would be a better choice.

5. Summary

In summary, we provide a detailed discussion of our ML method for reconstructing the crucial parameter, dephasing time T_2, from HHG spectra. The advantages, validity, efficiency, as well as the limitations of our method are all investigated in this work. We hope that our ML code offers a feasible solution to a longstanding challenge in condensed matter physics and ultrafast physics: reconstructing the dephasing time from experimental data.

Program availability

All the data in this work are available from the authors upon request.

Our ML code module for reconstructing dephasing time is integrated into our developing software package PROBP 2.0, which, along with the user guide, will be released soon. Alternatively, the dephasing time reconstruction module can be obtained by contacting zhaoxi719@snnu.edu.cn. Our software module is available under the Apache License V2.0.

Acknowledgements

This work was supported by the Fundamental Research Funds for the Central Universities (Grant No. GK202207012), QCYRCXM-2022-241, the National Key Research and Development Program of China (Grant No. 2022YFE0134200), the Natural Science Foundation of Jilin Province (Grant No. 20220101016JC), and the National Natural Science Foundation of China (Grant Nos. 12374238, 11934004, and 11974230). C.C.S. receives partial support from the National Natural Science Foundation of China (Grant No. 12274470) and the Natural Science Foundation of Hunan Province for Distinguished Young Scholars (Grant No. 2022JJ10070).

References

[1] Huibers A G, Switkes M, Marcus C M, Campman K and Gossard A C 1997 *Phys. B* **81** 348
[2] Guo Z, Habenicht B F, Liang W Z and Prezhdo O V 2010 *Phys. Rev. B* **81** 760
[3] Ding S and McDowell C A 1999 *J. Phys.: Condens. Matter* **11** L199
[4] Ilzhöfer P, Sohmen M, Durastante G, Politi C, Trautmann A, Natale G, Morpurgo G, Giamarchi T, Chomaz L, Mark M, *et al.* 2021 *Nat. Phys.* **17** 356
[5] Xi Z, Li Y and Fan H 2015 *Sci. Rep.* **5** 10922
[6] Cywiński Ł, Lutchyn R M, Nave C P and Das Sarma S 2008 *Phys. Rev. B* **77** 174509
[7] Kelly S P, Poschinger U, Schmidt-Kaler F, Fisher M and Marino J 2023 *SciPost Phys.* **15** 250
[8] Streltsov A, Adesso G and Plenio M B 2017 *Rev. Mod. Phys.* **89** 041003
[9] Du T Y and Ma C 2022 *Phys. Rev. A* **105** 053125
[10] Orlando G, Ho T S and Chu S I 2020 *J. Opt. Soc. Am. B* **37** 1540
[11] Siddiqi I 2021 *Nat. Rev. Mater.* **6** 875
[12] Assaf B A, Cardinal T, Wei P, Katmis F, Moodera J S and Heiman D 2013 *Appl. Phys. Lett.* **102** 012102
[13] Chow W W, Schneider H and Phillips M 2003 *Phys. Rev. A* **68** 053802
[14] Du T Y 2019 *Phys. Rev. A* **100** 053401
[15] Aslam N, Zhou H, Urbach E K, Turner M J, Walsworth R L, Lukin M D and Park H 2023 *Nat. Rev. Phys.* **5** 157
[16] Picon J D, Bussac M and Zuppiroli L 2007 *Phys. Rev. B* **75** 235106
[17] Li Y C and Lin H Q 2016 *Sci. Rep.* **6** 26365
[18] Fan X, Takagahara T, Cunningham J and Wang H 1998 *Solid State Commun.* **108** 857
[19] Dey P, Paul J, Wang Z, Stevens C, Liu C, Romero A, Shan J, Hilton D and Karaiskaj D 2016 *Phys. Rev. Lett.* **116** 127402
[20] Bermeister A, Keith D and Culcer D 2014 *Appl. Phys. Lett.* **105** 192102
[21] Dugaev V, Sherman E Y and Barnaś J 2011 *Phys. Rev. B* **83** 085306

[22] Hase M and Kitajima M 2010 *J. Phys.: Condens. Matter* **22** 073201
[23] Kilen I, Kolesik M, Hader J, Moloney J V, Huttner U, Hagen M K and Koch S W 2020 *Phys. Rev. Lett.* **125** 083901
[24] Yang J, Sun Q, Ueno K, Shi X, Oshikiri T, Misawa H and Gong Q 2018 *Nat. Commun.* **9** 4858
[25] Albash T and Lidar D A 2015 *Phys. Rev. A* **91** 062320
[26] Montañez-Barrera J, von Spakovsky M R, Damian Ascencio C E and Cano-Andrade S 2022 *Phys. Rev. A* **106** 032426
[27] Du T Y and Bian X B 2017 *Opt. Express* **25** 151
[28] Qiao Y, Chen J, Li Z, Liu Y, Jiang S, Liu W, Yang Y and Chen J 2024 *Opt. Lett.* **49** 3986
[29] Qiao Y, Chen J, Zhou S, Chen J, Jiang S and Yang Y 2024 *Chin. Phys. Lett.* **41** 014205
[30] Li L, Lan P, Zhu X, Huang T, Zhang Q, Lein M and Lu P 2019 *Phys. Rev. Lett.* **122** 193901
[31] Ghimire S, DiChiara A D, Sistrunk E, Agostini P, DiMauro L F and Reis D A 2011 *Nat. Phys.* **7** 138
[32] Uzan A J, Orenstein G, Jiménez-Galán Á, McDonald C, Silva R E, Bruner B D, Klimkin N D, Blanchet V, Arusi-Parpar T, Krüger M, *et al.* 2020 *Nat. Photonics* **14** 183
[33] Qian C, Yu C, Jiang S, Zhang T, Gao J, Shi S, Pi H, Weng H and Lu R 2022 *Phys. Rev. X* **12** 021030
[34] Mahesh B 2020 *Int. J. Sci. Res.* **9** 381
[35] Jordan M I and Mitchell T M 2015 *Science* **349** 255
[36] Luu T T, Garg M, Kruchinin S Y, Moulet A, Hassan M T and Goulielmakis E 2015 *Nature* **521** 498
[37] You Y S, Wu M, Yin Y, Chew A, Ren X, Gholam-Mirzaei S, Browne D A, Chini M, Chang Z, Schafer K J, *et al*. 2017 *Opt. Lett.* **42** 1816
[38] Cavalieri A L, Müller N, Uphues T, Yakovlev V S, Baltuška A, Horvath B, Schmidt B, Blümel L, Holzwarth R, Hendel S, *et al.* 2007 *Nature* **449** 1029
[39] Ramasesha K, Leone S R and Neumark D M 2016 *Annu. Rev. Phys. Chem.* **67** 41
[40] Borsch M, Schmid C P, Weigl L, Schlauderer S, Hofmann N, Lange C, Steiner J, Koch S, Huber R and Kira M 2020 *Science* **370** 1204
[41] Lakhotia H, Kim H, Zhan M, Hu S, Meng S and Goulielmakis E 2020 *Nature* **583** 55
[42] Vampa G, Lu J, You Y S, Baykusheva D R, Wu M, Liu H, Schafer K J, Gaarde M B, Reis D A and Ghimire S 2020 *J. Phys. B: At. Mol. Opt. Phys.* **53** 144003
[43] Park J, Subramani A, Kim S and Ciappina M F 2022 *Adv. Phys.: X* **7** 2003244
[44] Han S, Ortmann L, Kim H, Kim Y W, Oka T, Chacon A, Doran B, Ciappina M, Lewenstein M, Kim S W, *et al*. 2019 *Nat. Commun.* **10** 3272
[45] Zhou T, Wu W, Zhang J, Yu S and Fang L 2023 *Sci. Adv.* **9** eadg4391
[46] Bedolla E, Padierna L C and Castaneda-Priego R 2020 *J. Phys.: Condens. Matter* **33** 053001
[47] Carrasquilla J and Melko R G 2017 *Nat. Phys.* **13** 431
[48] Carrasquilla J 2020 *Adv. Phys.: X* **5** 1797528
[49] Zahavy T, Dikopoltsev A, Moss D, Haham G I, Cohen O, Mannor S and Segev M 2018 *Optica* **5** 666
[50] White J and Chang Z 2019 *Opt. Express* **27** 4799
[51] Brunner C, Duensing A, Schröder C, Mittermair M, Golkov V, Pollanka M, Cremers D and Kienberger R 2022 *Opt. Express* **30** 15669
[52] Zhao X, Wang S J, Yu W W, Wei H, Wei C, Wang B, Chen J and Lin C D 2020 *Phys. Rev. Appl.* **13** 034043
[53] Zhao X, Wei H, Wu Y and Lin C D 2017 *Phys. Rev. A* **95** 043407
[54] Zhao X, Wei H, Wu Y and Lin C D 2017 *Phys. Rev. A* **95** 043407
[55] Zhao Y T, Ma S y, Jiang S C, Yang Y J, Zhao X and Chen J G 2019 *Opt. Express* **27** 34392
[56] Qiao Y, Huo Y, Liang H, Chen J, Liu W, Yang Y and Jiang S 2023 *Phys. Rev. B* **107** 075201
[57] Wu D, Li L, Zhan Y, Huang T, Cui H, Li J, Lan P and Lu P 2022 *Phys. Rev. A* **105** 063101
[58] Vampa G, McDonald C, Orlando G, Klug D, Corkum P and Brabec T 2014 *Phys. Rev. Lett.* **113** 073901
[59] Cao J, Li F, Bai Y, Liu P and Li R 2021 *Opt. Express* **29** 4830
[60] Kolesik M and Moloney J 2023 *Phys. Rev. B* **108** 115433
[61] Vampa G and Brabec T 2017 *J. Phys. B: At. Mol. Opt. Phys.* **50** 083001
[62] Mirjalili S 2019 *Stud. Comput. Intell.* **780** 43
[63] Wang D, Tan D and Liu L 2018 *Soft Comput.* **22** 387
[64] Koessler E and Almomani A 2021 *Optim. Eng.* **22** 1539
[65] Wikmark H, Guo C, Vogelsang J, Smorenburg P W, Coudert-Alteirac H, Lahl J, Peschel J, Rudawski P, Dacasa H, Carlström S, Maclot S, Gaarde M B, Johnsson P, Arnold C L and L'Huillier A 2019 *Proc. Natl. Acad. Sci. USA* **116** 4779
[66] Brown G G, Jiménez-Galán A, Silva R E F and Ivanov M 2024 *Phys. Rev. Res.* **6** 043005
[67] Zhao X, Liu J, Jiang S, Chen J, Wang J and Zhang S 2025 Opticaopen: 28225302.v1[physics.atm–clus]
[68] Yang S, Liu X, Zhang H, Song X, Zuo R, Meier T and Yang W 2024 *Opt. Express* **32** 15862
[69] Langer F, Hohenleutner M, Schmid C P, Pöllmann C, Nagler P, Korn T, Schüller C, Sherwin M, Huttner U, Steiner J, *et al*. 2016 *Nature* **533** 225
[70] Jiang S, Chen J, Wei H, Yu C, Lu R and Lin C 2018 *Phys. Rev. Lett.* **120** 253201