

PyStructureFactor: A Python code for the molecular structure factor in tunneling ionization rates

Shanshan Song^{a,1}, Mingyu Zhu^{a,b,1}, Hongcheng Ni^{a,c,d,*}, Jian Wu^{a,c,d,e,**}

^a*State Key Laboratory of Precision Spectroscopy, East China Normal University, Shanghai 200241, China*

^b*School of Physics and Electronic Science, East China Normal University, Shanghai 200241, China*

^c*NYU-ECNU Joint Institute of Physics, New York University at Shanghai, Shanghai 200062, China*

^d*Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan, Shanxi 030006, China*

^e*CAS Center for Excellence in Ultra-intense Laser Science, Shanghai 201800, China*

Abstract

Tunneling ionization is at the core of strong-field and attosecond science. In this paper, we present PyStructureFactor — a general Python code towards the calculation of the structure factor in the tunneling ionization rate of common molecules under intense laser fields. The numerical implementation is based on the well-developed weak-field asymptotic theory in the integral representation. The information of the electronic structure of the molecules is obtained via the PySCF quantum chemistry package. PyStructureFactor is a general computational framework that can be utilized to compute the molecular structure factor of various types of molecules, including polar and nonpolar diatomic molecules, degenerate molecules, and [multiplet open-shell](#) molecules. Examples are given that are benchmarked against known results with good agreements. The present PyStructureFactor is implemented in an efficient manner and is easily applicable towards larger molecules.

*Corresponding author.

E-mail address: hcni@lps.ecnu.edu.cn

**Corresponding author.

E-mail address: jwu@phy.ecnu.edu.cn

¹These authors contributed equally.

Keywords: Molecular structure factor; Tunneling ionization; Weak-field asymptotic theory; PySCF.

PROGRAM SUMMARY

Program Title: PyStructureFactor

CPC Library link to program files: (to be added by Technical Editor)

Developer’s repository link: <https://github.com/TheStarAlight/PyStructureFactor>

Code Ocean capsule: (to be added by Technical Editor)

Licensing provisions: Apache-2.0

Programming language: Python 3

Supplementary material:

Nature of problem: The structure factor of a molecule in intense laser fields determines its orientation-dependent tunneling ionization rate, which is crucial for the studies of ultrafast molecular dynamics and its control. However, the strong-field community lacks an open-source code to calculate the molecular structure factor, and can only resort to known results of a few molecules.

Solution method: We develop the PyStructureFactor program with the structure factor of arbitrary molecules calculated using the weak-field asymptotic theory in the integral representation. The underlying electronic structure necessary for the calculation is obtained using the PySCF quantum chemistry package.

Restrictions on the Accuracy: The accuracy of the molecular structure factor calculated by PyStructureFactor is restricted by the level of precision of the electronic structure information extracted from the PySCF package.

Running time: The running time depends on the size of the molecule and the basis set of the SCF calculation and other parameters passed to the program. The example in Fig. 2 took 1.2 sec. to finish on an AMD Ryzen 9 7950X CPU on the WSL Ubuntu 22.04 LTS.

1. Introduction

Advances in the laser technology have made intense laser fields widely accessible in typical table-top lab settings. When matter interacts with intense laser fields, a plethora of novel strong-field phenomena emerge that are of fundamental interest, such as high-order harmonic generation (HHG) [1–3], above-threshold ionization [4–6], nonsequential double ionization [7], and laser-induced electron diffraction [8, 9]. HHG is the key to the generation of extreme ultraviolet laser pulses, which has facilitated the improvement of time resolution of ultraprecise measurements down to the attosecond level [10–13].

Common to many of these strong-field phenomena is their first step, known as tunneling ionization [14, 15], which is thus a cornerstone of strong-field and attosecond science. It has attracted widespread interest not only in physics but also in chemistry and related fields. Given the recent development in optochemistry [16, 17], a complementary emerging field to photochemistry, the importance of the tunneling ionization of molecules has been elevated to new heights.

In contrast to photoionization where the laser is manifested as photons in the frequency domain, tunneling ionization is better represented by the highly nonlinear ionization induced by intense optical fields in the time domain $F(t)$, with a tunneling ionization rate [18, 19]

$$W[F(t)] \sim e^{-2\kappa^3/3F(t)} \quad (1)$$

depending exponentially on the field strength $F(t)$, where $\kappa = \sqrt{-2E_0}$ with E_0 the orbital energy of the target. For typical intense infrared laser fields, the photon energy is substantially lower than the ionization potential of the bound electron. Moreover, the timescale of the field oscillation is much longer than the intrinsic attosecond timescale of the bound electron. Therefore, the strong-field ionization process can be approximated by tunneling in a static electric field represented by the instantaneous laser electric field $F(t)$ [20, 21].

In addition to the field factor [Eq. (1)], the tunneling ionization rate of molecules depends on a factor related to the electronic structure of the ionizing orbital, known as the molecular structure factor. It determines the orientation-dependent rate of tunneling ionization that is crucial to the analysis of subsequent strong-field molecular dynamics and photoelectron spectrum [22]. It is the aim of the current work to present the numerical solution to the molecular structure factor.

The molecular structure factor can be obtained either by the molecular Ammosov–Delone–Krainov theory (MO-ADK) [22–26] or via the weak-field asymptotic theory (WFAT) [27–42]. WFAT is applicable in laser fields below the critical field strength that leads to over-barrier ionization. Based on the parabolic adiabatic expansion approach [43], WFAT generalizes isotropic atomic potentials to arbitrary molecular potentials and naturally accounts for the effect of molecular permanent dipole moment (if any), and has been adopted as the theoretical method in some experimental research on molecular strong-field ionization [44–48]. In this work, we employ WFAT towards a numerical implementation of the computation of the molecular structure factor. Under

its framework, the obtained tunneling ionization rate, in its leading order, factorizes into the field factor [Eq. (1)] and the absolute square of the structure factor, the latter of which is an inherent property of the ionizing molecular orbital independent of the external field [30].

In an essence of the perturbation theory, the structure factor of WFAT is ~~taken as the asymptotic expansion for the field strength found through the behavior of the wave function in the asymptotic region~~ in early stages, known as the tail representation [29, 31, 35]. Hence, its reliable evaluation requires an accurate description of the asymptotic tail of the ionizing orbital. To extract the orbital wave function of polyatomic molecules by the Hartree-Fock (HF) method or density functional theory (DFT) with quantum chemistry software packages [31, 34], one normally resorts to the basis-based approaches. However, compared to the correct exponential behavior obtained by the grid-based approach, the standard Gaussian-type basis functions facilitating multicenter integral calculations decay too rapidly in the asymptotic region and thus could result in erratic oscillations. To reproduce the correct local property of the asymptotic tail of the orbital, large basis sets with quadruple or pentuple-zeta quality and variationally optimized exponents are required [30]. This approach significantly increases the computational cost and become infeasible in many cases. The efficiency of such numerical computation needs to be improved urgently.

As a remedy, an alternative integral approach of WFAT was developed with the spirit of scattering theory to overcome the above issue [39–41]. In this approach, the structure factor is expressed as an integral form involving the ionizing orbital and a known analytical function. The integral accumulates at the localized regions where the orbital has large amplitudes, and it is insensitive to the representation of the asymptotic tail far from the nuclei. Hence, the integral approach can be implemented using standard quantum chemistry software packages with Gaussian-type orbitals (GTOs) for arbitrary molecules.

In this work, an open-source program implementing WFAT in the integral representation is developed that can easily and quickly calculate the molecular structure factor of common molecules. In our implementation, named the `PyStructureFactor`, the ionizing molecular orbital is expanded as a linear combination of GTOs at the HF level of theory calculated by the `PySCF` package [49, 50]. The integrals are evaluated numerically using a three-dimensional grid-based method, which exhibits much better numerical stability and is readily applicable to large molecules, compared to the standard analytical integration procedure. Our present implementation of `PyStructureFactor` features

high accuracy and efficient computation towards the calculation of the molecular structure factor.

This work is organized as follows. In Sec. 2, the definition and derivation of the tunneling ionization rate, the molecular structure factor, and other parameters within the WFAT framework are given. In Sec. 3, the details of numerical implementation are presented. In Sec. 4, we detail the usage of the PyStructureFactor program to calculate the molecular structure factor. In Sec. 5, we illustrate the application of the PyStructureFactor program with examples of various types of common molecules. Conclusions are given in Sec. 6. Atomic units are used throughout unless stated otherwise.

2. Theoretical formulation

Tunneling ionization of molecules can be usually modeled within the Born-Oppenheimer and the single-active-electron approximation. Under such a framework, the wave function of the ionizing orbital ψ_0 is the solution to an effective one-electron Schrödinger equation

$$\left[-\frac{1}{2}\nabla^2 + V(\mathbf{r}) \right] \psi_0(\mathbf{r}) = E_0 \psi_0(\mathbf{r}), \quad (2)$$

where the HF potential $V(\mathbf{r})$ describes the interaction between the ionizing electron and the parent ion.

Within WFAT, tunneling ionization is investigated in parabolic coordinates. We assume the electric field vector $\mathbf{F} = F\hat{\mathbf{z}}$, and define the orientation of the molecule using a set of Euler angles (α, β, γ) (within the z - y' - z'' convention [39], which is shown schematically in Fig. 1) representing a rotation $\hat{\mathbf{R}}$ from the laboratory frame (LF) to the molecular frame (MF). The ionization rate does not depend on α and so we set $\alpha = 0$ throughout. The total ionization rate is split into different parabolic channels:

$$\Gamma(\beta, \gamma) = \sum_{\nu} \Gamma_{\nu}(\beta, \gamma), \quad (3)$$

where $\Gamma_{\nu}(\beta, \gamma)$ are partial rates of parabolic quantum number indices

$$\nu = (n_{\xi}, m), \quad n_{\xi} = 0, 1, 2, \dots, \quad m = 0, \pm 1, \pm 2, \dots \quad (4)$$

In the leading-order approximation, the asymptotic expansion of the partial rates for $F \rightarrow 0$ gives

$$\Gamma_{\nu}(\beta, \gamma) = |G_{\nu}(\beta, \gamma)|^2 W_{\nu}(F), \quad (5)$$

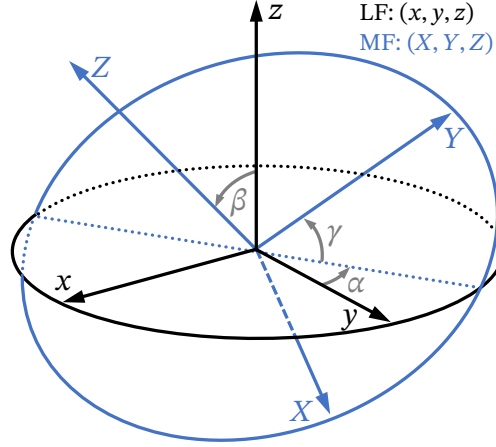


Figure 1: Illustration of the rotation from the LF to the MF using Euler angles of $z\text{-}y'\text{-}z''$ convention. The rotation consists of three steps: (1) The coordinate system (x, y, z) rotates about the z axis by angle α , then (2) rotates about the new y axis by angle β , finally (3) rotates about the rotated z axis by γ , and becomes (X, Y, Z) . The rotations mentioned above are all counter-clockwise.

which can be separated into two factors: The structural part $|G_\nu(\beta, \gamma)|^2$ and the field part $W_\nu(F)$. The field factor, $W_\nu(F)$, has the form [18, 19]

$$W_\nu(F) = \frac{\kappa}{2} \left(\frac{4\kappa^2}{F} \right)^{2Z/\kappa - 2n_\xi - |m| - 1} e^{-2\kappa^3/3F}. \quad (6)$$

The structure factor, namely $G_\nu(\beta, \gamma)$, in the integral representation of WFAT, is given as an integral:

$$G_\nu(\beta, \gamma) = e^{-\kappa\mu_z} \int \Omega_\nu^*(\hat{\mathbf{R}}^{-1}\mathbf{r}) \mathbf{V}_c(\mathbf{r}) \psi_0(\mathbf{r}) d\mathbf{r}, \quad (7)$$

where

$$\boldsymbol{\mu} = \int \psi_0^*(\mathbf{r}) \mathbf{r} \psi_0(\mathbf{r}) d\mathbf{r} \quad (8)$$

denotes the orbital dipole moment in the LF with μ_z its component along the field direction; Ω_ν is a reference function which can be expanded into spherical harmonics:

$$\Omega_\nu(\mathbf{r}) = \sum_{l=|m|}^{\infty} \Omega_{lm}^\nu(\mathbf{r}), \quad (9)$$

where

$$\Omega_{lm}^\nu(\mathbf{r}) = \sum_{l=|m|}^{\infty} R_l^\nu(r) Y_{lm}(\theta, \phi); \quad (10)$$

$V_c(\mathbf{r}) = V(\mathbf{r}) + Z/r$ is the core potential without the Coulomb tail (i.e., $V_c \rightarrow 0$ as $r \rightarrow \infty$) with Z the asymptotic charge of the parent molecular ion. The radial part R_l^ν of the reference function Ω_{lm}^ν in Eq. (10) reads

$$R_l^\nu(r) = \omega_l^\nu (\kappa r)^l e^{-\kappa r} M(l+1-Z/\kappa, 2l+2, 2\kappa r), \quad (11)$$

where $M(a, b, x)$ is the confluent hyper-geometric function [51] and ω_l^ν is the normalization coefficient:

$$\begin{aligned} \omega_l^\nu &= (-1)^{l+(|m|-m)/2+1} 2^{l+3/2} \kappa^{Z/\kappa-(|m|+1)/2-n_\xi} \\ &\times \sqrt{(2l+1)(l+m)!(l-m)!(|m|+n_\xi)!n_\xi!} \frac{l!}{(2l+1)!} \\ &\times \sum_{k=0}^{\min(n_\xi, l-|m|)} \frac{\Gamma(l+1-Z/\kappa+n_\xi-k)}{k!(l-k)!(|m|+k)!(l-|m|-k)!(n_\xi-k)!}. \end{aligned} \quad (12)$$

Under the framework of the HF method, the HF potential $V(\mathbf{r})$ consists of three parts, namely the nuclear Coulomb potential (V_{nuc}), the direct (V_d) and exchange (V_{ex}) parts of inter-electron interactions:

$$V(\mathbf{r}) = V_{\text{nuc}}(\mathbf{r}) + V_d(\mathbf{r}) + V_{\text{ex}}(\mathbf{r}), \quad (13)$$

where

$$V_{\text{nuc}}(\mathbf{r}) = - \sum_{A=1}^{N_{\text{atm}}} \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|}, \quad (14)$$

$$V_d(\mathbf{r}) = \sum_{i=1}^N \int \frac{\psi_i^*(\mathbf{r}') \psi_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}', \quad (15)$$

$$V_{\text{ex}}(\mathbf{r}) \psi_0(\mathbf{r}) = - \sum_{i=1}^N \psi_i(\mathbf{r}) \int \frac{\psi_i^*(\mathbf{r}') \psi_0(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \langle \sigma_i | \sigma_0 \rangle d\mathbf{r}', \quad (16)$$

where N is the number of electrons, N_{atm} is the number of nuclei, $\psi_i(\mathbf{r})$ and σ_i denote the molecular orbital and the spin state of the electron of index i ($\langle \sigma_i | \sigma_j \rangle = 1$ if electrons i and j have the same spin, and $\langle \sigma_i | \sigma_j \rangle = 0$ otherwise), while Z_A and \mathbf{R}_A are the nuclear charge and position of atom of index A .

3. Numerical implementation

For the practical numerical implementation of the computation of the molecular structure factor within the framework of WFAT, we first refer to the evaluation of the integral in Eq. (7). Directly evaluating the integrals for each molecular orientation (β, γ) is inefficient. Nevertheless, we may utilize the spherical harmonic expansion of the reference function Ω_ν , and express the rotated spherical harmonic functions as a linear combination of spherical harmonic functions using the Wigner- d matrix. Substituting Eqs. (9) and (10) into Eq. (7) yields [39]

$$G_\nu(\beta, \gamma) = e^{-\kappa\mu_z} \sum_{l=|m|}^{\infty} \sum_{m'=-l}^l I_{lm'}^\nu d_{mm'}^l(\beta) e^{-im'\gamma}, \quad (17)$$

where $d_{mm'}^l(\beta)$ is the Wigner- d matrix, and the integral $I_{lm'}^\nu$ is expressed as

$$I_{lm'}^\nu = \int \Omega_{lm'}^{\nu*}(\mathbf{r}) V_c(\mathbf{r}) \psi_0(\mathbf{r}) d\mathbf{r}. \quad (18)$$

In this way the computational overhead is significantly reduced — with the values of the integrals $I_{lm'}^\nu$, we may obtain the structure factor of arbitrary molecular orientation $G_\nu(\beta, \gamma)$ with little computational cost.

The next topic is concerned with the calculation of the HF potential operator acting on the ionizing orbital, i.e., $V(\mathbf{r})\psi_0(\mathbf{r})$. In the practical application, the molecular orbitals are usually obtained by the self-consistent-field (SCF) approach, which is commonly implemented in standard quantum chemistry packages. The molecular orbitals are usually expanded into linear combinations of Gaussian-type orbitals (GTOs):

$$\psi_i(\mathbf{r}) = \sum_{\alpha} C_{i\alpha} \chi_{\alpha}(\mathbf{r} - \mathbf{R}_{\alpha}), \quad (19)$$

where $\chi_{\alpha}(\mathbf{r} - \mathbf{R}_{\alpha})$ denotes a GTO of index α centering at \mathbf{R}_{α} , and the coefficients $C_{i\alpha}$ are obtained through the SCF approach. In this way we obtained $V_d(\mathbf{r})$ and $V_{\text{ex}}(\mathbf{r})\psi_0(\mathbf{r})$ in Eqs. (15) and (16) expressed in coefficients $C_{i\alpha}$ for practical calculations. Defining the integral

$$J_{\alpha\beta}(\mathbf{r}) = \int \frac{\chi_{\alpha}(\mathbf{r}' - \mathbf{R}_{\alpha}) \chi_{\beta}(\mathbf{r}' - \mathbf{R}_{\beta})}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (20)$$

which can be calculated analytically for GTOs, Eqs. (15) and (16) can be written as

$$V_d(\mathbf{r}) = \sum_{i=1}^N \sum_{\alpha,\beta} C_{i\alpha} C_{i\beta} J_{\alpha\beta}(\mathbf{r}), \quad (21)$$

$$V_{\text{ex}}(\mathbf{r})\psi_0(\mathbf{r}) = - \sum_{i=1}^N \langle \sigma_i | \sigma_0 \rangle \sum_{\alpha,\beta,\gamma} C_{i\alpha} C_{i\beta} C_{0\gamma} J_{\beta\gamma}(\mathbf{r}) \chi_{\alpha}(\mathbf{r} - \mathbf{R}_{\alpha}). \quad (22)$$

The last topic is related to the convergence of the algorithm. The integral in Eq. (18) is numerically evaluated on a grid using the Becke fuzzy cell integration scheme [52] implemented in the PySCF package, which uses higher density of grid points near the nuclei to accelerate convergence of the calculation. Apart from adopting the Becke grids in the numerical integration, we also choose the origin such that the dipole of the parent ion

$$\mathbf{D} = \left[\sum_{A=1}^{N_{\text{atm}}} Z_A \mathbf{R}_A - \sum_{i=1}^N \int \psi_i^*(\mathbf{r}) \mathbf{r} \psi_i(\mathbf{r}) d\mathbf{r} \right] - \boldsymbol{\mu} \quad (23)$$

vanishes, which reduces the grid size required for the integral Eq. (18) to converge.

4. Program usage

The usage of the program is simple and straightforward. Firstly, import necessary packages, including `PyStructureFactor`, `PySCF` and `NumPy` in Python. Next, initialize the molecule object by calling the method `pyscf.M`, and provide necessary information of the molecule. Finally, get the structure factor of the molecule by invoking the method `PyStructureFactor.get_structure_factor`, whose return value is a NumPy array containing the structure factor at different molecular orientations (β, γ).

To better illustrate the usage of the program, we present a description of the input parameters of the two main methods, including `pyscf.M`, which defines the molecular and orbital information, and `PyStructureFactor.get_structure_factor`, which outputs the corresponding structure factor, in Tables 1 and 2, respectively. Shown in Fig. 2 is a minimal example of the program which calculates the molecular structure factor $G_{00}(\beta)$ corresponding to the highest occupied molecular orbital (HOMO) of the hydrogen molecule. For

Parameter	Description	Default
atom	Information of the molecular structure. Refer to Fig. 2 for an example.	—
unit	The unit used for the position coordinates of the atoms in the molecule. 'B' or 'AU' indicates Bohr and otherwise Angström.	'Angström'
basis	Basis set used for calculation. Refer to pyscf.gto.basis for more information. An accurate basis set is suggested for optimal convergence.	'sto-3g'
charge	Charge of the molecule or molecular ion.	0
spin	Number of spin (i.e., $2S$) of the molecule.	0

Table 1: Essential input parameters of the `pyscf.M` method in the `PyStructureFactor` program.

the diatomic molecules, the γ Euler angle does not come into play, and β denotes the angle between the electric field vector and the molecular axis in the LF. More examples are available in Sec. 5.

Now we present convergence and reliability tests of the program in Figs. 3 and 4 with N_2 and CO molecules, respectively. To this end, we obtain their structure factors as a function of β and compare the results to reference data extracted from the literature [34] which is obtained using the tail representation of WFAT. In both figures, the pc-4 basis is adopted in the calculations, and the cut-off limit of the index l in the summation of Eq. (17), i.e., the l_{\max} , is set to the default value of 10. The “grid_level” labels denote the input parameter “atom_grid_level” of the program. Here, grid levels 1, 3, 5, and 7 correspond to radial and angular sizes (40, 194), (75, 302), (105, 770), and (135, 1202) for each spherical grid around the nuclei, respectively. It is clear from the figures that for “atom_grid_level” larger than 3, the obtained results are hardly distinguishable from each other even after enlargement, which reveals a stable and fast convergence of the present program. In addition, the converged result is close to the reference data. This illustrates that the calculation of the molecular structure factor is able to achieve accurate results under default parameters. Minor discrepancy from the reference data results from the difference in the tail and integral presentations of WFAT and the choice of basis.

```

from PyStructureFactor import get_structure_factor
import numpy as np
import pyscf
n_beta = 90
n_gamma = 1
molH2 = pyscf.M(atom="H□0,0,0.37;□H□0,0,-0.37", basis="pc-1",
    spin=0)
beta_grid = np.linspace(0, np.pi, n_beta)
G_grid = get_structure_factor(
    mol = molH2, orbital_index = 0, channel = (0,0),
    lmax = 10, hf_method = "RHF",
    atom_grid_level = 3,
    orient_grid_size = (n_beta, n_gamma))

```

Figure 2: A minimal example of the program calculating the structure factor $G_{00}(\beta)$ of the HOMO of the hydrogen molecule (H_2).

5. Illustrative examples

In this section we take a number of common molecules as examples, and present their structure factors obtained from PyStructureFactor. The examples cover molecules with various properties, as shown in Figs. 5 to 9: diatomic molecules (H_2 , CO, and O_2), linear polyatomic molecules (C_2H_2) and planar molecules (C_6H_6); [multiplet open-shell](#) molecules (O_2 [and \$C_6H_6\$](#)); molecules with degenerate HOMOs (double HOMO degeneracy for O_2 , C_2H_2 and C_6H_6).

For all examples presented, we set the input parameter l_{\max} to the default value of 10 and set the "atom_grid_level" parameter to the default value of 3. For linear molecules, the γ dependence of the structure factor drops out and it only depends on β . For nonlinear molecules, on the other hand, the structure factor is presented on the (β, γ) grid, see Fig. 9 for an example of the C_6H_6 molecule.

It is worthwhile noting that the molecular structure factor varies not only from molecule to molecule, but also changes from orbital to orbital for a particular molecule. Novel laser technologies have enabled the exploration of ionization from different molecular orbitals, and studies show that subsequent molecular dynamics critically depends on the ionizing orbital. Not surprisingly, the structure factor also depends on the orbital from which ionization occurs, as shown in Figs. 5 and 6.

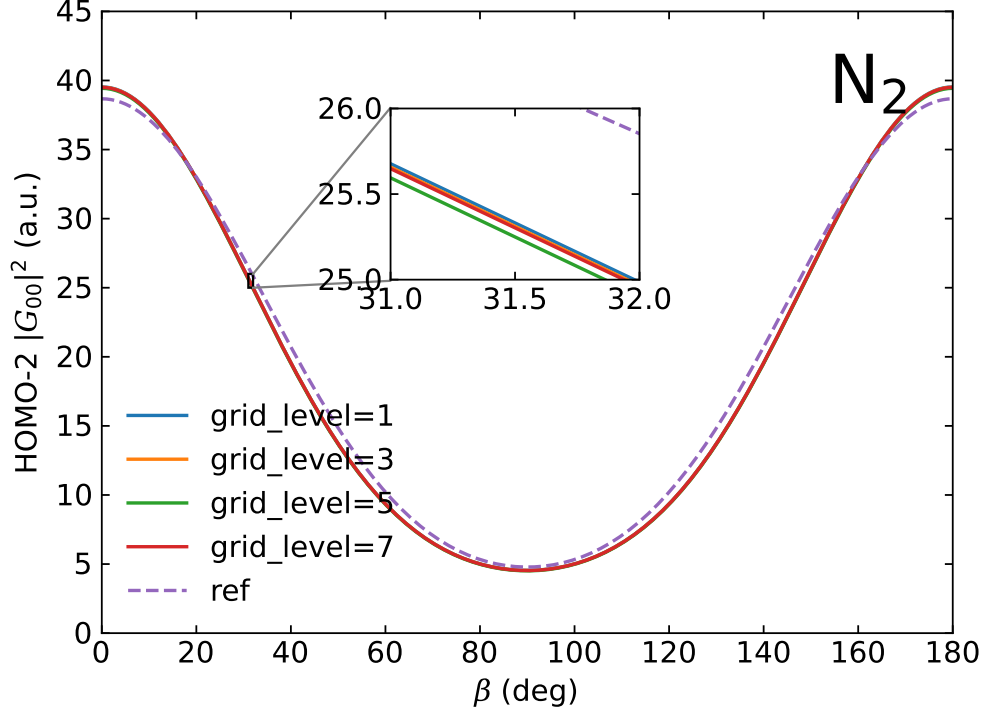


Figure 3: Squared structure factor $|G_{00}|^2$ of the HOMO-2 of the nitrogen molecule (N_2). The reference data are extracted from Ref. [34].

Furthermore, for molecules with degenerate orbitals, the structure factor depends on the orientation of the orbital. The HOMOs of O_2 , for example, are π orbitals that are two degenerate orbitals perpendicular to each other. Each of them has a nodal plane, and we use the nodal plane to identify the two degenerate orbitals. “HOMO- yz ” corresponds to the orbital that has a nodal plane on the y - z plane while “HOMO- xz ” represents that with a nodal plane on the x - z plane. Clearly, the structure factor depends on the orientation of the orbital and thus varies for different degenerate HOMO orbitals, as shown in Figs. 7 and 8. We note that tunneling ionization could occur simultaneously from multiple orbitals, and the final tunneling ionization rate might be a summation over various contributions.

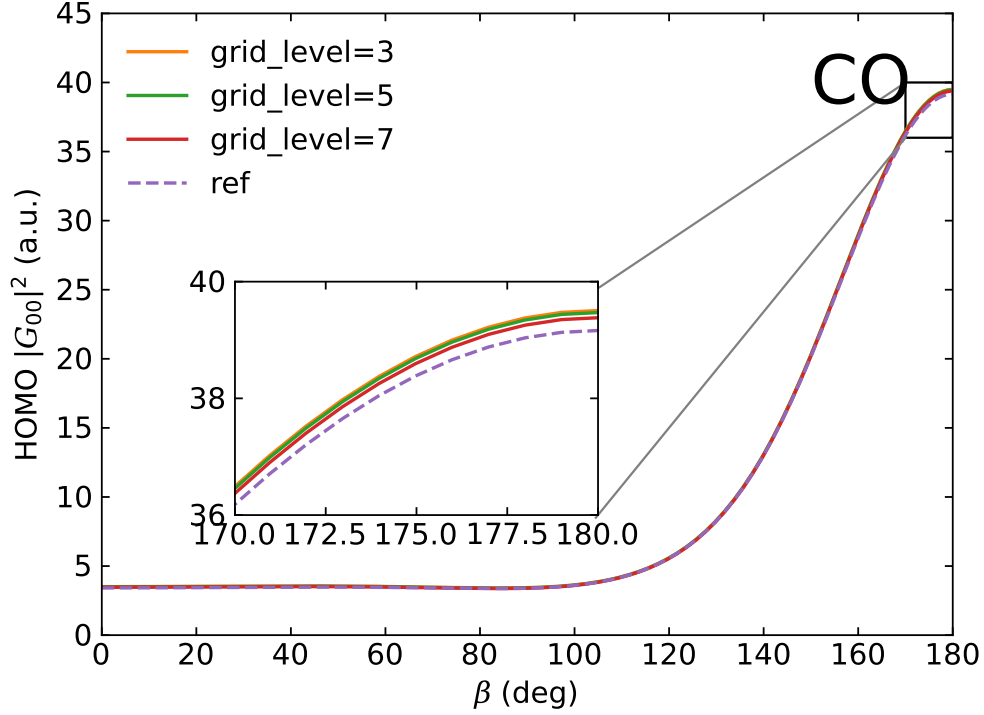


Figure 4: Squared structure factor $|G_{00}|^2$ of the HOMO of the carbon monoxide molecule (CO). The reference data are extracted from Ref. [34].

6. Conclusions

The tunneling ionization rate of molecules depends on the molecular orientation. Such dependence is determined by the molecular structure factor. In this work, we have presented the PyStructureFactor program, which computes the molecular structure factor within the WFAT theoretical framework based on the PySCF quantum chemistry package. PyStructureFactor is implemented in a manner that is both computationally efficient and easily accessible to the end users. It provides a straightforward approach to obtain the molecular structure factor of virtually all kinds of molecules, and therefore lays a solid foundation for the study of strong-field molecular dynamics.

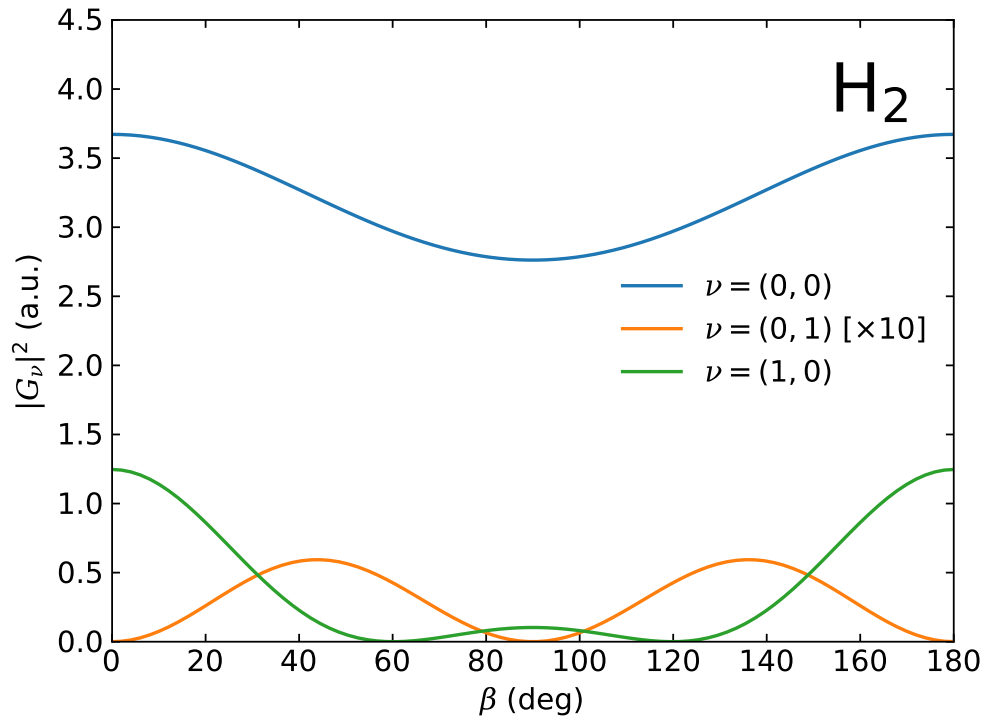


Figure 5: Squared structure factor $|G_{\nu}|^2$ of the HOMO of the hydrogen molecule (H_2). The molecular axis is aligned along the z -axis in the MF and the internuclear distance is 0.74 \AA . The basis used for calculation is pc-4.

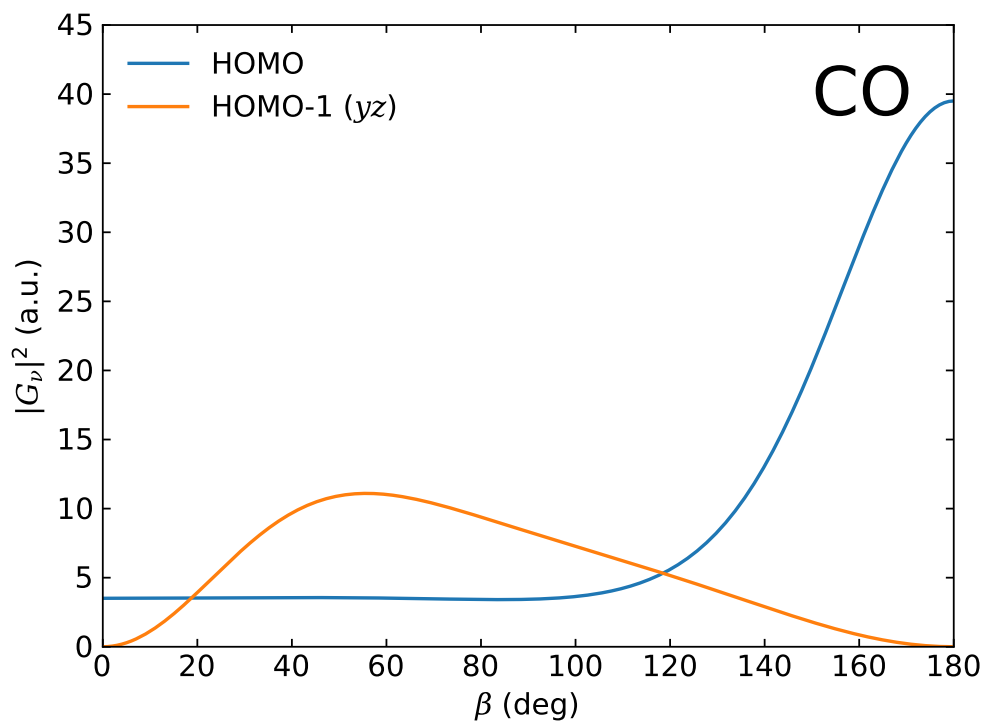


Figure 6: Squared structure factor $|G_{00}|^2$ of the HOMO and HOMO-1 of the carbon monoxide molecule (CO, with the nodal plane on the y-z plane). The molecular axis is aligned along the z-axis in the MF with the +z axis pointing from the C to the O nucleus, and the internuclear distance is 1.13 Å. The basis used for calculation is pc-4.

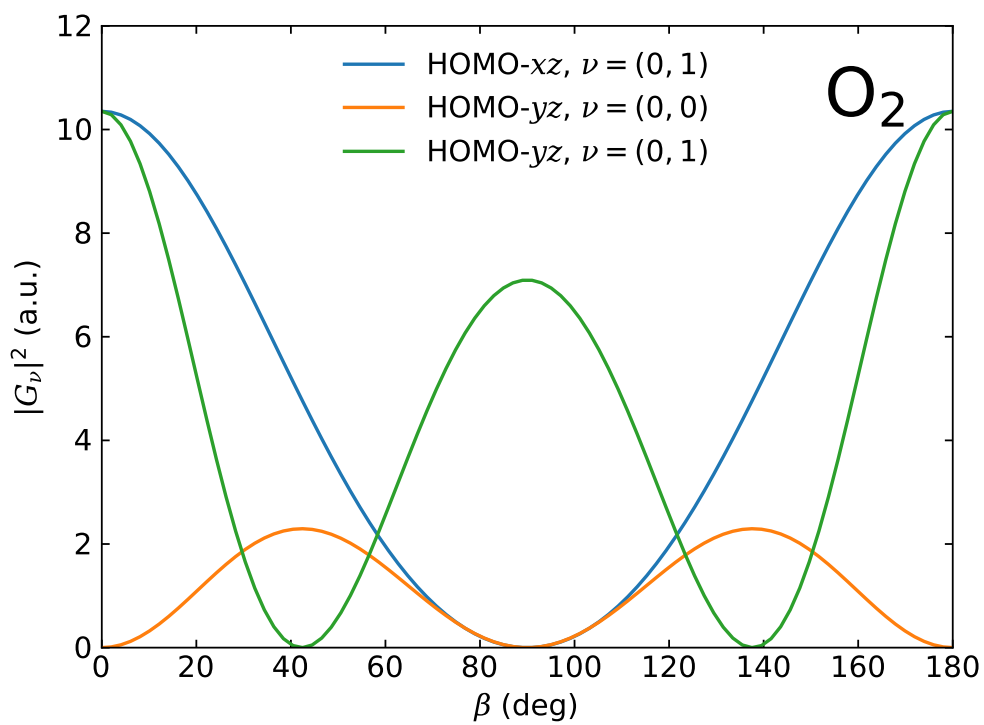


Figure 7: Squared structure factor $|G_v|^2$ of the two degenerate HOMOs of the oxygen molecule (O_2 , denoted by their nodal planes xz and yz). The molecular axis is aligned along the z -axis in the MF and the internuclear distance is 1.21 Å. The basis used for calculation is pc-4. The oxygen molecule at the ground state has two unpaired electrons and thus has a total spin $S = 1$, which requires parameters `spin=2` and `hf_method="UHF"` to be passed to the program.

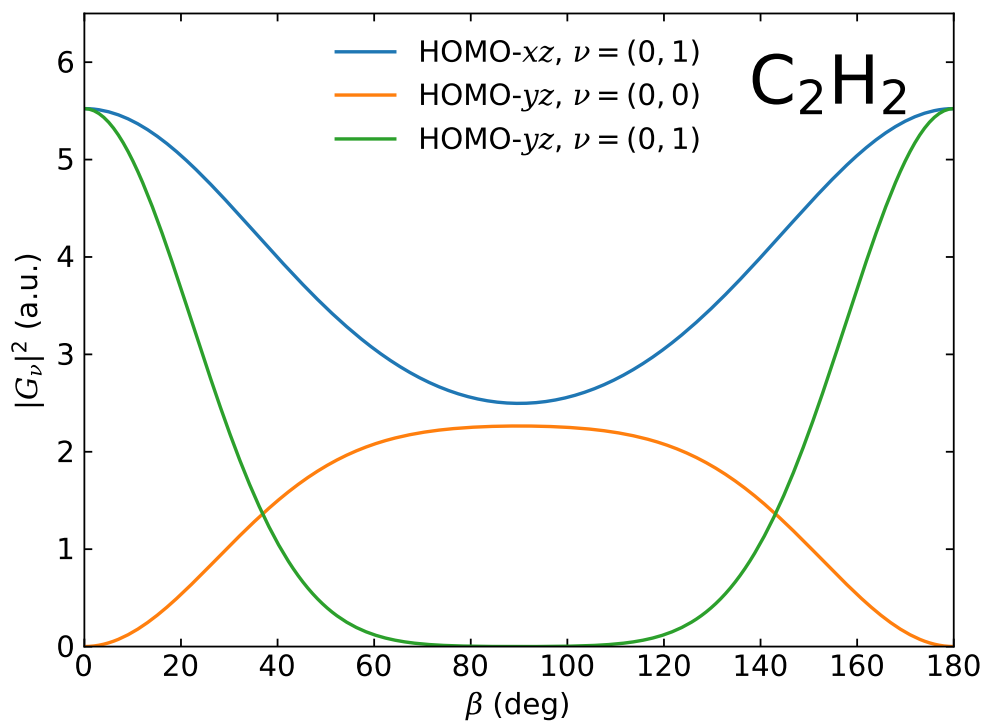


Figure 8: Squared structure factor $|G_v|^2$ of the two degenerate HOMOs of the acetylene molecule (C_2H_2). The molecular axis is aligned along the z -axis in the MF, the internuclear distance between the C nuclei is 1.20 \AA , and it is 1.06 \AA between the C nucleus and the neighboring H nucleus. The basis used for calculation is pc-2.

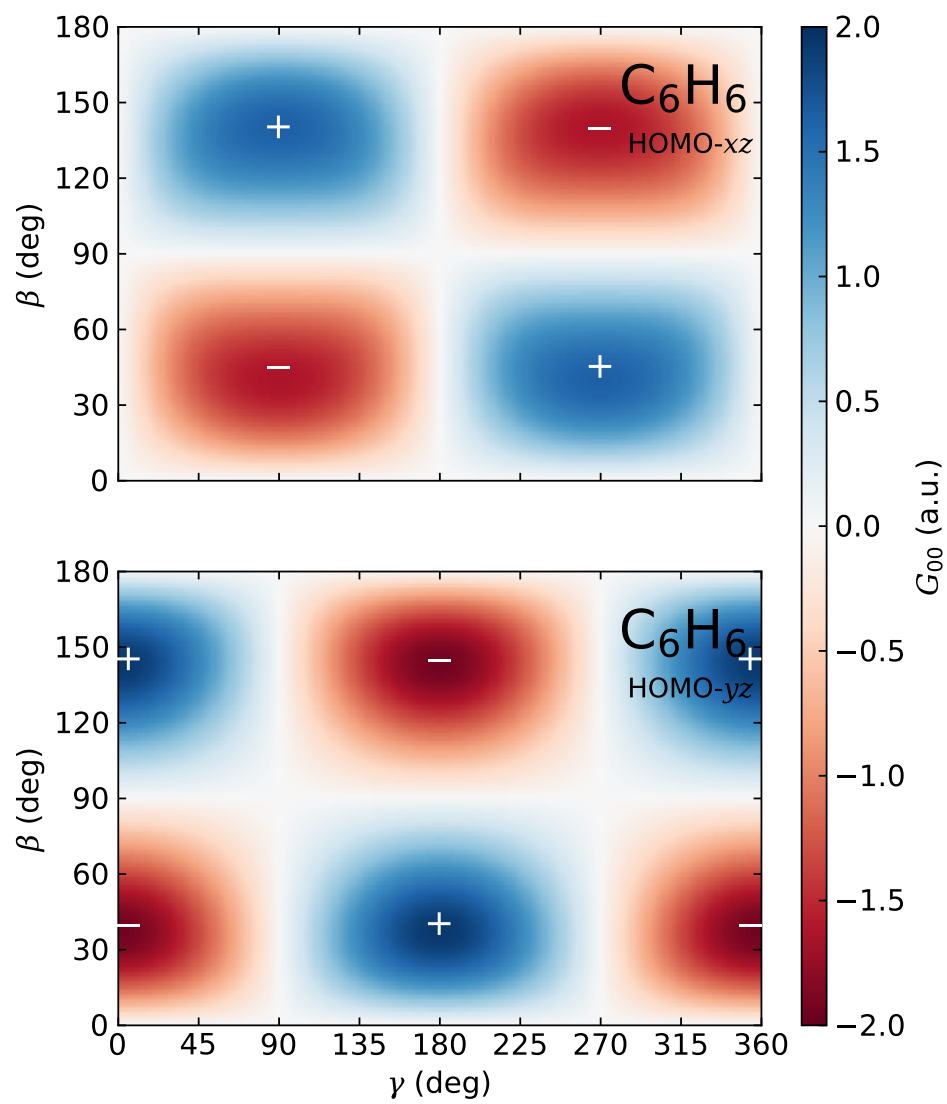


Figure 9: Structure factor G_{00} of the two degenerate HOMOs of the benzene molecule (C_6H_6). The molecule is placed in the x - z plane in the MF, with the z -axis pointing from one C nucleus to its neighboring H nucleus. The basis used for calculation is pc-1.

Acknowledgments

We would like to thank Qiming Sun, Hao Huang, and Shengzhe Pan for helpful discussions. This work is supported by the National Natural Science Foundation of China (Grant No. 92150105) and the Science and Technology Commission of Shanghai Municipality (Grant No. 21ZR1420100). Numerical computations were in part performed on the East China Normal University Multifunctional Platform for Innovation (001).

References

- [1] J. L. Krause, K. J. Schafer, K. C. Kulander, High-order harmonic generation from atoms and ions in the high intensity regime, *Phys. Rev. Lett.* 68 (1992) 3535–3538. [doi:10.1103/PhysRevLett.68.3535](https://doi.org/10.1103/PhysRevLett.68.3535).
- [2] P. B. Corkum, Plasma perspective on strong field multiphoton ionization, *Phys. Rev. Lett.* 71 (1993) 1994–1997. [doi:10.1103/PhysRevLett.71.1994](https://doi.org/10.1103/PhysRevLett.71.1994).
- [3] T. Popmintchev, M.-C. Chen, P. Arpin, M. M. Murnane, H. C. Kapteyn, The attosecond nonlinear optics of bright coherent X-ray generation, *Nat. Photon.* 4 (2010) 822–832. [doi:10.1038/nphoton.2010.256](https://doi.org/10.1038/nphoton.2010.256).
- [4] P. Agostini, F. Fabre, G. Mainfray, G. Petite, N. K. Rahman, Free-Free Transitions Following Six-Photon Ionization of Xenon Atoms, *Phys. Rev. Lett.* 42 (1979) 1127–1130. [doi:10.1103/PhysRevLett.42.1127](https://doi.org/10.1103/PhysRevLett.42.1127).
- [5] W. Becker, F. Grasbon, R. Kopold, D. Milošević, G. Paulus, H. Walther, Above-Threshold Ionization: From Classical Features to Quantum Effects, *Adv. At. Mol. Opt. Phys.* 48 (2002) 35–98. [doi:10.1016/S1049-250X\(02\)80006-4](https://doi.org/10.1016/S1049-250X(02)80006-4).
- [6] D. B. Milošević, G. G. Paulus, D. Bauer, W. Becker, Above-threshold ionization by few-cycle pulses, *J. Phys. B* 39 (2006) R203. [doi:10.1088/0953-4075/39/14/R01](https://doi.org/10.1088/0953-4075/39/14/R01).
- [7] B. Walker, B. Sheehy, L. F. DiMauro, P. Agostini, K. J. Schafer, K. C. Kulander, Precision Measurement of Strong Field Double Ionization of Helium, *Phys. Rev. Lett.* 73 (1994) 1227–1230. [doi:10.1103/PhysRevLett.73.1227](https://doi.org/10.1103/PhysRevLett.73.1227).

- [8] C. I. Blaga, J. Xu, A. D. DiChiara, E. Sistrunk, K. Zhang, P. Agostini, T. A. Miller, L. F. DiMauro, C. D. Lin, Imaging ultrafast molecular dynamics with laser-induced electron diffraction, *Nature* 483 (2012) 194–197. doi:[10.1038/nature10820](https://doi.org/10.1038/nature10820).
- [9] B. Wolter, M. G. Pullen, A.-T. Le, M. Baudisch, K. Doblhoff-Dier, A. Sentsleben, M. Hemmer, C. D. Schröter, J. Ullrich, T. Pfeifer, R. Moshhammer, S. Gräfe, O. Vendrell, C. D. Lin, J. Biegert, Ultrafast electron diffraction imaging of bond breaking in di-ionized acetylene, *Science* 354 (2016) 308–312. doi:[10.1126/science.aah3429](https://doi.org/10.1126/science.aah3429).
- [10] F. Krausz, M. Ivanov, Attosecond physics, *Rev. Mod. Phys.* 81 (2009) 163–234. doi:[10.1103/RevModPhys.81.163](https://doi.org/10.1103/RevModPhys.81.163).
- [11] J. M. Dahlström, A. L’Huillier, A. Maquet, Introduction to attosecond delays in photoionization, *J. Phys. B* 45 (2012) 183001. doi:[10.1088/0953-4075/45/18/183001](https://doi.org/10.1088/0953-4075/45/18/183001).
- [12] R. Pazourek, S. Nagele, J. Burgdörfer, Attosecond chronoscopy of photoemission, *Rev. Mod. Phys.* 87 (2015) 765–802. doi:[10.1103/RevModPhys.87.765](https://doi.org/10.1103/RevModPhys.87.765).
- [13] A. S. Kheifets, Wigner time delay in atomic photoionization, *J. Phys. B* 56 (2023) 022001. doi:[10.1088/1361-6455/acb188](https://doi.org/10.1088/1361-6455/acb188).
- [14] L. V. Keldysh, Ionization in the Field of a Strong Electromagnetic Wave, *Sov. Phys. JETP* 20 (1965) 1307.
- [15] S. L. Chin, From Multiphoton to Tunnel Ionization, Vol. Volume 16 of *Advances in Multi-Photon Processes and Spectroscopy*, WORLD SCIENTIFIC, 2004. doi:[10.1142/9789812796585_0003](https://doi.org/10.1142/9789812796585_0003).
- [16] H. Li, X. Gong, H. Ni, P. Lu, X. Luo, J. Wen, Y. Yang, X. Qian, Z. Sun, J. Wu, Light-Induced Ultrafast Molecular Dynamics: From Photochemistry to Optochemistry, *J. Phys. Chem. Lett.* 13 (2022) 5881–5893. doi:[10.1021/acs.jpcllett.2c01119](https://doi.org/10.1021/acs.jpcllett.2c01119).
- [17] J. Ma, L. Xu, H. Ni, C. Lu, W. Zhang, P. Lu, J. Wen, F. He, O. Faucher, J. Wu, Transient Valence Charge Localization in Strong-Field Dissociative Ionization of HCl Molecules, *Phys. Rev. Lett.* 127 (2021) 183201. doi:[10.1103/PhysRevLett.127.183201](https://doi.org/10.1103/PhysRevLett.127.183201).

- [18] M. V. Ammosov, N. B. Delone, V. P. Krainov, Tunnel ionization of complex atoms and of atomic ions in an alternating electromagnetic field, *Sov. Phys. JETP* 64 (1986) 1191.
- [19] N. B. Delone, V. P. Krainov, Tunneling and barrier-suppression ionization of atoms and ions in a laser radiation field, *Phys. Usp.* 41 (1998) 469. doi:[10.1070/PU1998v041n05ABEH000393](https://doi.org/10.1070/PU1998v041n05ABEH000393).
- [20] O. I. Tolstikhin, T. Morishita, S. Watanabe, Adiabatic theory of ionization of atoms by intense laser pulses: One-dimensional zero-range-potential model, *Phys. Rev. A* 81 (2010) 033415. doi:[10.1103/PhysRevA.81.033415](https://doi.org/10.1103/PhysRevA.81.033415).
- [21] O. I. Tolstikhin, T. Morishita, Adiabatic theory of ionization by intense laser pulses: Finite-range potentials, *Phys. Rev. A* 86 (2012) 043417. doi:[10.1103/PhysRevA.86.043417](https://doi.org/10.1103/PhysRevA.86.043417).
- [22] D. Pavičić, K. F. Lee, D. M. Rayner, P. B. Corkum, D. M. Villeneuve, Direct Measurement of the Angular Dependence of Ionization for N₂, O₂, and CO₂ in Intense Laser Fields, *Phys. Rev. Lett.* 98 (2007) 243001. doi:[10.1103/PhysRevLett.98.243001](https://doi.org/10.1103/PhysRevLett.98.243001).
- [23] X. M. Tong, Z. X. Zhao, C. D. Lin, Theory of molecular tunneling ionization, *Phys. Rev. A* 66 (2002) 033402. doi:[10.1103/PhysRevA.66.033402](https://doi.org/10.1103/PhysRevA.66.033402).
- [24] L. Holmegaard, J. L. Hansen, L. Kalkhøj, S. Louise Kragh, H. Stapelfeldt, F. Filsinger, J. Küpper, G. Meijer, D. Dimitrovski, M. Abu-samha, C. P. J. Martiny, L. Bojer Madsen, Photoelectron angular distributions from strong-field ionization of oriented molecules, *Nat. Phys.* 6 (2010) 428–432. doi:[10.1038/nphys1666](https://doi.org/10.1038/nphys1666).
- [25] R. Johansen, K. G. Bay, L. Christensen, J. Thøgersen, D. Dimitrovski, L. B. Madsen, H. Stapelfeldt, Alignment-dependent strong-field ionization yields of carbonyl sulfide molecules induced by mid-infrared laser pulses, *J. Phys. B: At. Mol. Opt. Phys.* 49 (2016) 205601. doi:[10.1088/0953-4075/49/20/205601](https://doi.org/10.1088/0953-4075/49/20/205601).
- [26] S.-F. Zhao, J.-K. Li, G.-L. Wang, P.-C. Li, X.-X. Zhou, Accurate Structure Parameters for Tunneling Ionization Rates of Gas-Phase Lin-

- ear Molecules, Commun. Theo. Phys. 67 (2017) 289. [doi:10.1088/0253-6102/67/3/289](https://doi.org/10.1088/0253-6102/67/3/289).
- [27] O. I. Tolstikhin, T. Morishita, L. B. Madsen, Theory of tunneling ionization of molecules: Weak-field asymptotics including dipole effects, Phys. Rev. A 84 (2011) 053423. [doi:10.1103/PhysRevA.84.053423](https://doi.org/10.1103/PhysRevA.84.053423).
 - [28] L. B. Madsen, O. I. Tolstikhin, T. Morishita, Application of the weak-field asymptotic theory to the analysis of tunneling ionization of linear molecules, Phys. Rev. A 85 (2012) 053404. [doi:10.1103/PhysRevA.85.053404](https://doi.org/10.1103/PhysRevA.85.053404).
 - [29] L. B. Madsen, F. Jensen, O. I. Tolstikhin, T. Morishita, Structure factors for tunneling ionization rates of molecules, Phys. Rev. A 87 (2013) 013406. [doi:10.1103/PhysRevA.87.013406](https://doi.org/10.1103/PhysRevA.87.013406).
 - [30] V. H. Trinh, O. I. Tolstikhin, L. B. Madsen, T. Morishita, First-order correction terms in the weak-field asymptotic theory of tunneling ionization, Phys. Rev. A 87 (2013) 043426. [doi:10.1103/PhysRevA.87.043426](https://doi.org/10.1103/PhysRevA.87.043426).
 - [31] L. B. Madsen, F. Jensen, O. I. Tolstikhin, T. Morishita, Application of the weak-field asymptotic theory to tunneling ionization of H₂O, Phys. Rev. A 89 (2014) 033412. [doi:10.1103/PhysRevA.89.033412](https://doi.org/10.1103/PhysRevA.89.033412).
 - [32] O. I. Tolstikhin, L. B. Madsen, T. Morishita, Weak-field asymptotic theory of tunneling ionization in many-electron atomic and molecular systems, Phys. Rev. A 89 (2014) 013421. [doi:10.1103/PhysRevA.89.013421](https://doi.org/10.1103/PhysRevA.89.013421).
 - [33] I. Y. Tolstikhina, T. Morishita, O. I. Tolstikhin, Application of the many-electron weak-field asymptotic theory of tunneling ionization to atoms, Phys. Rev. A 90 (2014) 053413. [doi:10.1103/PhysRevA.90.053413](https://doi.org/10.1103/PhysRevA.90.053413).
 - [34] R. Saito, O. I. Tolstikhin, L. B. Madsen, T. Morishita, Structure factors for tunneling ionization rates of diatomic molecules, At. Data Nucl. Data Tables 103-104 (2015) 4–49. [doi:10.1016/j.adt.2015.02.001](https://doi.org/10.1016/j.adt.2015.02.001).
 - [35] V. H. Trinh, V. N. T. Pham, O. I. Tolstikhin, T. Morishita, Weak-field asymptotic theory of tunneling ionization including the first-order correction terms: Application to molecules, Phys. Rev. A 91 (2015) 063410. [doi:10.1103/PhysRevA.91.063410](https://doi.org/10.1103/PhysRevA.91.063410).

- [36] A. I. Dnestryan, O. I. Tolstikhin, Integral-equation approach to the weak-field asymptotic theory of tunneling ionization, *Phys. Rev. A* 93 (2016) 033412. [doi:10.1103/PhysRevA.93.033412](https://doi.org/10.1103/PhysRevA.93.033412).
- [37] J. Svensmark, O. I. Tolstikhin, L. B. Madsen, Theory of dissociative tunneling ionization, *Phys. Rev. A* 93 (2016) 053426. [doi:10.1103/PhysRevA.93.053426](https://doi.org/10.1103/PhysRevA.93.053426).
- [38] V. H. Trinh, O. I. Tolstikhin, T. Morishita, First-order correction terms in the weak-field asymptotic theory of tunneling ionization in many-electron systems, *J. Phys. B: At. Mol. Opt. Phys.* 49 (2016) 195603. [doi:10.1088/0953-4075/49/19/195603](https://doi.org/10.1088/0953-4075/49/19/195603).
- [39] L. B. Madsen, F. Jensen, A. I. Dnestryan, O. I. Tolstikhin, Structure factors for tunneling ionization rates of molecules: General Hartree-Fock-based integral representation, *Phys. Rev. A* 96 (2017) 013423. [doi:10.1103/PhysRevA.96.013423](https://doi.org/10.1103/PhysRevA.96.013423).
- [40] A. I. Dnestryan, O. I. Tolstikhin, L. B. Madsen, F. Jensen, Structure factors for tunneling ionization rates of molecules: General grid-based methodology and convergence studies, *J. Chem. Phys.* 149 (2018) 164107. [doi:10.1063/1.5046902](https://doi.org/10.1063/1.5046902).
- [41] P. K. Samygin, T. Morishita, O. I. Tolstikhin, Weak-field asymptotic theory of tunneling ionization from nearly degenerate states, *Phys. Rev. A* 98 (2018) 033401. [doi:10.1103/PhysRevA.98.033401](https://doi.org/10.1103/PhysRevA.98.033401).
- [42] H. Matsui, O. I. Tolstikhin, T. Morishita, Weak-field asymptotic theory of tunneling ionization of the hydrogen molecule including core polarization, spectator nucleus, and internuclear motion effects, *Phys. Rev. A* 103 (2021) 033102. [doi:10.1103/PhysRevA.103.033102](https://doi.org/10.1103/PhysRevA.103.033102).
- [43] P. A. Batishchev, O. I. Tolstikhin, T. Morishita, Atomic Siegert states in an electric field: Transverse momentum distribution of the ionized electrons, *Phys. Rev. A* 82 (2010) 023416. [doi:10.1103/PhysRevA.82.023416](https://doi.org/10.1103/PhysRevA.82.023416).
- [44] H. Ohmura, N. Saito, T. Morishita, Molecular tunneling ionization of the carbonyl sulfide molecule by double-frequency phase-controlled laser fields, *Phys. Rev. A* 89 (2014) 013405. [doi:10.1103/PhysRevA.89.013405](https://doi.org/10.1103/PhysRevA.89.013405).

- [45] P. M. Kraus, O. I. Tolstikhin, D. Baykusheva, A. Rupenyan, J. Schneider, C. Z. Bisgaard, T. Morishita, F. Jensen, L. B. Madsen, H. J. Wörner, Observation of laser-induced electronic structure in oriented polyatomic molecules, *Nat. Commun.* 6 (2015) 7039. doi:[10.1038/ncomms8039](https://doi.org/10.1038/ncomms8039).
- [46] S. G. Walt, N. Bhargava Ram, A. von Conta, O. I. Tolstikhin, L. B. Madsen, F. Jensen, H. J. Wörner, Role of Multi-Electron Effects in the Asymmetry of Strong-Field Ionization and Fragmentation of Polar Molecules: The Methyl Halide Series, *J. Phys. Chem. A* 119 (2015) 11772–11782. doi:[10.1021/acs.jpca.5b07331](https://doi.org/10.1021/acs.jpca.5b07331).
- [47] T. Endo, A. Matsuda, M. Fushitani, T. Yasuike, O. I. Tolstikhin, T. Morishita, A. Hishikawa, Imaging Electronic Excitation of NO by Ultrafast Laser Tunneling Ionization, *Phys. Rev. Lett.* 116 (2016) 163002. doi:[10.1103/PhysRevLett.116.163002](https://doi.org/10.1103/PhysRevLett.116.163002).
- [48] H. Fujise, M. Uemura, H. Hasegawa, D. Ikeya, A. Matsuda, T. Morishita, L. B. Madsen, F. Jensen, O. I. Tolstikhin, A. Hishikawa, Helicity-dependent dissociative tunneling ionization of CF₄ in multicycle circularly polarized intense laser fields, *Phys. Chem. Chem. Phys.* 24 (2022) 8962–8969. doi:[10.1039/D1CP05858D](https://doi.org/10.1039/D1CP05858D).
- [49] Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, G. K. Chan, PySCF: the Pythonbased simulations of chemistry framework, *WIREs Comput. Mol. Sci.* 8 (2018) e1340. doi:[10.1002/wcms.1340](https://doi.org/10.1002/wcms.1340).
- [50] Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, J. J. Eriksen, Y. Gao, S. Guo, J. Hermann, M. R. Hermes, K. Koh, P. Koval, S. Lehtola, Z. Li, J. Liu, N. Mardirossian, J. D. McClain, M. Motta, B. Mussard, H. Q. Pham, A. Pulkin, W. Purwanto, P. J. Robinson, E. Ronca, E. R. Sayfutyarova, M. Scheurer, H. F. Schurkus, J. E. T. Smith, C. Sun, S.-N. Sun, S. Upadhyay, L. K. Wagner, X. Wang, A. White, J. D. Whitfield, M. J. Williamson, S. Wouters, J. Yang, J. M. Yu, T. Zhu, T. C. Berkelbach, S. Sharma, A. Y. Sokolov, G. K.-L. Chan, Recent developments in the PySCF program package, *J. Chem. Phys.* 153 (2020) 024109. doi:[10.1063/5.0006074](https://doi.org/10.1063/5.0006074).
- [51] F. Olver, D. Lozier, R. Boisvert, C. Clark, *The NIST Handbook of Mathematical Functions*, Cambridge University Press, 2010.

- [52] A. D. Becke, A multicenter numerical integration scheme for polyatomic molecules, *J. Chem. Phys.* 88 (1988) 2547–2553. [doi:10.1063/1.454033](https://doi.org/10.1063/1.454033).

Parameter	Description	Default
<code>mol</code>	The PySCF molecule object. Initialized by invoking <code>pyscf.Mol</code> or <code>pyscf.gto.M</code> .	—
<code>orbital_index</code>	Index of the ionizing orbital relative to HOMO, e.g., HOMO \rightarrow 0, LUMO \rightarrow +1, HOMO - 1 \rightarrow -1,	0
<code>channel</code>	Parabolic channel $\nu = (n_\xi, m)$. The program would calculate the structure factor of channel (n_ξ, m) , i.e., $G_{n_\xi, m}$.	(0, 0)
<code>lmax</code>	The maximum cut-off limit of the angular quantum number (larger l would be neglected) used in the summation in Eq. (17).	10
<code>hf_method</code>	Indicates whether 'RHF' or 'UHF' should be used in the molecular HF calculation. Note: 'UHF' must be used for multiplet open-shell molecules.	'RHF'
<code>atom_grid_level</code>	Level of fineness of the grid used in integration (see also <code>pyscf.dft.Grid</code>), which controls the number of radial and angular grids around each atom in the evaluation of the integration , ranging from 0 to 9.	3
<code>orient_grid_size</code>	Size of the output (β, γ) grid. The grid is uniform, with β ranging from 0 to π and γ ranging from 0 to 2π . Setting the γ grid count to 1 indicates that γ would be zero throughout the calculation.	(90, 1)
<code>move_dip_zero</code>	Indicates whether to shift the molecular coordinates such that its dipole moment [Eq. (23)] vanishes (i.e., $\mathbf{D} = \mathbf{0}$).	True
<code>rmax</code>	Indicates the cut-off limit of the radial grid points, points of radii larger than <code>rmax</code> would be neglected in the calculation.	40
Returns	A NumPy array containing the structure factor $G_{n_\xi, m}$ of the given channel (n_ξ, m) on the (β, γ) orientation grid, whose shape is initialized from given by the <code>orient_grid_size</code> .	

Table 2: Input parameters and return value of the `PyStructureFactor.get_structure_factor` method.