

iSPECTRON: a simulation interface for linear and nonlinear spectra with ab-initio quantum chemistry software

Francesco Segatta¹, Artur Nenov¹, Daniel R. Nascimento², Niranjan Govind^{2,*}, Shaul Mukamel^{3,*}, and Marco Garavelli^{1,*}

¹Dipartimento di Chimica Industriale “Toso Montanari”, Università degli Studi di Bologna, Viale del Risorgimento 4, I-40136 Bologna, Italy

²Physical and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, Washington 99352, United States

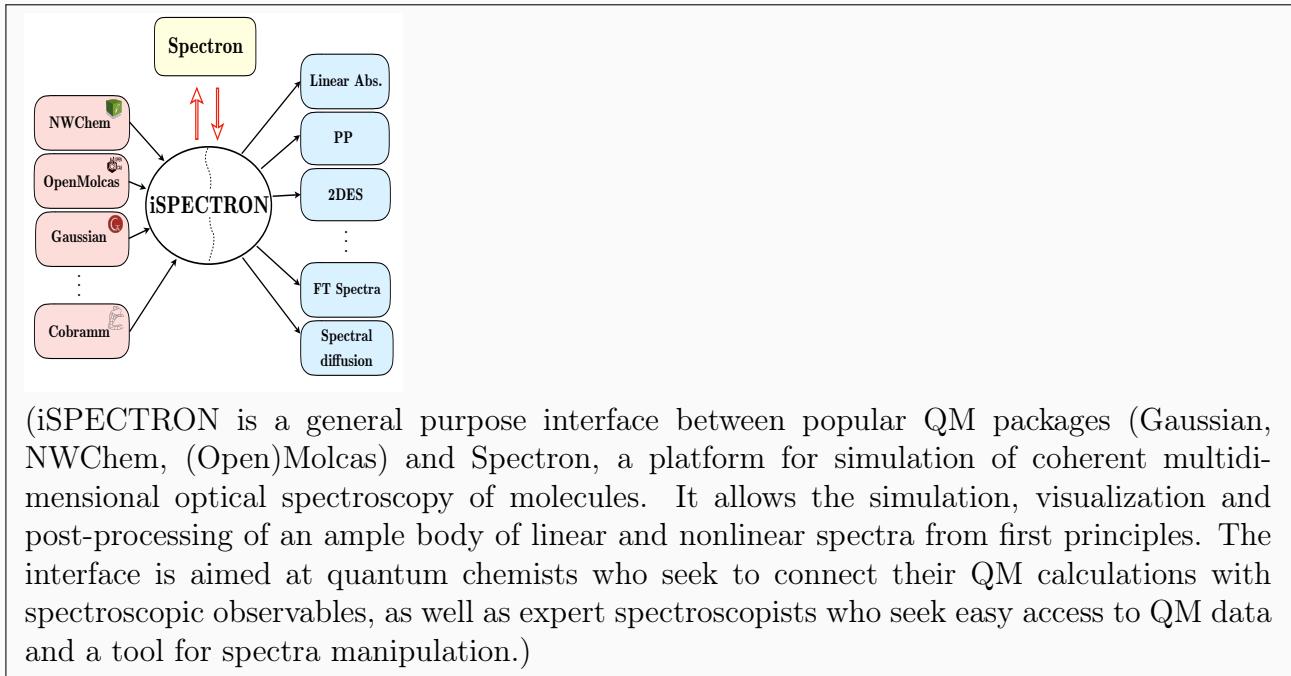
³Department of Chemistry and Department of Physics and Astronomy, University of California, Irvine, 92697, USA

September 23, 2020

Abstract

We introduce iSPECTRON, an open source (under the Educational Community License version 2.0) program that parses data from common quantum chemistry software (NWChem, OpenMolcas, Gaussian, Cobramm, etc.), produces the input files for the simulation of linear and nonlinear spectroscopy of molecules with the Spectron code, and analyzes the spectra with a broad range of tools. Vibronic spectra are expressed in term of the electronic eigenstates, obtained through quantum chemistry computations, and vibrational/bath effects are incorporated in the framework of the displaced harmonic oscillator model, where all required quantities are computed at the Franck-Condon point. The code capabilities are illustrated by simulating linear absorption, transient absorption and two dimensional electronic spectra of the pyrene molecule. Two levels of electronic structure theory, TDDFT (with NWChem) and RASSCF/RASPT2 (with OpenMolcas), are compared where possible.

Keywords: linear spectroscopy, two-dimensional nonlinear optical spectroscopy, CASSCF/CASPT2, TDDFT, line shape function, displaced harmonic oscillator, cumulant expansion of Gaussian fluctuations ■



INTRODUCTION

Spectroscopy is a valuable experimental tool for investigating electronic and vibrational properties and dynamics of molecular systems. The past few decades have witnessed the birth of multi-dimensional multi-pulse techniques, which allow one to study elementary molecular events such as energy and charge transfer processes, formation and evolution of vibrational and electronic coherences, conformational and solvent dynamics and even to study the evolution of a system simultaneously in time and space (4D spectroscopy)[88, 17, 79, 60, 10, 28, 81, 63]. In combination with ultra-short laser pulses with high phase stability, these novel techniques have equipped researchers with the tools necessary to unravel gas- and condensed-phase dynamics in the sub-femtosecond (fs, 10^{-15} s) and even in the attosecond (as, i.e. 10^{-18} s)[52, 87] regimes with an unprecedented level of detail. Nonetheless, connecting the optical response of the system to the underlying quantum-mechanical structure and vibronic dynamics remains an open computational challenge. Theoretical simulations constitute an invaluable tool for unravelling the complex underlying dynamics. The simulation of molecular spectroscopy is based on two ingredients: i) an accurate description of the electronic and vibrational structure and dynamics of the molecular system, and ii) an effective approach to account for the light-matter interaction and response.

Step i) is usually tackled by employing the tools of quantum chemistry, i.e., by solving the Schrödinger equation at some level of approximation, which yields the manifold of electronic states of the system and the transition dipole between them. Broadly, quantum mechanical calculations fall into two families - wavefunction (WF) based and density functional theory (DFT) based - both offering advantages and limitations due to the underlying approximations and computational cost. The description of molecular vibrations, environment fluctuations, as well as the mechanism and time scale of energy dissipation (e.g. internal conversion, energy or charge transfer) are key additional ingredients to properly access the physical phenomena that govern the recorded line-shapes. This can be achieved at various levels of sophistication ranging from numerically exact schemes solving the time-dependent Schrödinger equation (quantum dynamics) to model effective Hamiltonians which parametrize the couplings between the electronic states, as well as between the electronic

and vibrational degrees of freedom.

Step ii) is usually treated by quantum master equations which employ the non-perturbative or perturbative (i.e. weak-field) treatment of the light-matter interactions. A comprehensive perturbative theoretical framework for linear and nonlinear optical spectroscopies was first introduced a few decades ago.[53] The key quantity, the response function, incorporates all the relevant molecular information for the response to external light-induced perturbations. The response function requires, as a minimal input, the electronic levels of the system under study and the transition dipoles between them, but it also incorporates the vibronic dynamics at various degrees of sophistication.

Over the years, specialized programs that address steps i) and ii) have been developed by different groups and made available to the user community. For the former, various proprietary and open source packages for QM calculations exist - Gaussian[26], Molcas[27, 5], NWChem[84, 4], GAMESS[7] to name a few - but none of them encompass the entire array of existing and emerging electronic structure methods. General purpose codes for simulating spectroscopy are less common, reflecting the complexity associated with system specific input parameters required to simulate the nonlinear response. Among the publicly available codes we note FCclasses[74] and EXAT[33] for simulating the linear response in the eigenstate and excitonic picture, respectively, as well as Spectron, a platform developed by the Mukamel group specifically to describe coherent multidimensional optical spectroscopy of molecules, chromophore aggregates and semiconductor nanostructures from the IR to the X-ray.[1, 90]¹. There is a clear gap between Steps (i) and (ii) that needs to be bridged for the general user. In the present paper we introduce iSPECTRON, an open source (under the Educational Community License version 2.0) code conceived as a stepping stone towards a general purpose interface that would fill this gap. As schematically shown in Figure 1), it has been designed with a twofold goal:

- (a) Automatize the parsing of the output data from a range of popular quantum chemistry software (NWChem, OpenMolcas, Gaussian, Cobramm, etc.) to produce the necessary input files to simulate linear and nonlinear spectroscopy with the Spectron code;

¹Spectron can be obtained for free from <https://mukamel.ps.uci.edu/software.html>

- (b) Read the nonlinear response data produced by the Spectron code and produce, display and analyze a broad arsenal of possible spectra (transient absorption, 2DES, power spectra, etc.).

The interface is aimed at quantum chemists who seek to connect their QM calculations with spectroscopic observables, as well as at expert spectroscopists who seek easy access to QM data and a tool for spectra manipulation. In this work we present the capabilities of version 1.0 of the interface, which facilitates the first principles simulation of linear and nonlinear spectra within the electronic eigenstate expansion of the response function, relying on model Hamiltonians to describe the coupling to a bath². In the methodology section (Sec. 2) we provide a synopsis of the basic equations at the heart of the Spectron code, with the focus on the link to QM calculations. The workflow and capabilities of the interface are presented in a nutshell in Sec. 3 and employed in Sec. 4 to track the photo-induced vibrational dynamics and ultrafast internal conversion of pyrene, a polycyclic aromatic hydrocarbon (PAH) consisting of four fused benzene rings. We compare linear spectra at the time-dependent (TD) DFT (DFT-based) and RASSCF/RASPT2 (WF-based) levels of theory, employing data obtained from the NWChem and OpenMolcas software packages, respectively. Finally we present nonlinear spectra at the RASSCF/RASPT2 level of theory and analyze them with the post-processing tools of the interface.

METHODOLOGY

Theory

In this section we summarize the theoretical background used by the interface and by the Spectron code to simulate a broad variety of spectroscopic signals.

²The bath comprises the intra-molecular and the environment (e.g. a protein cage, a solvent, a nearby molecule) degrees of freedom (DOF)

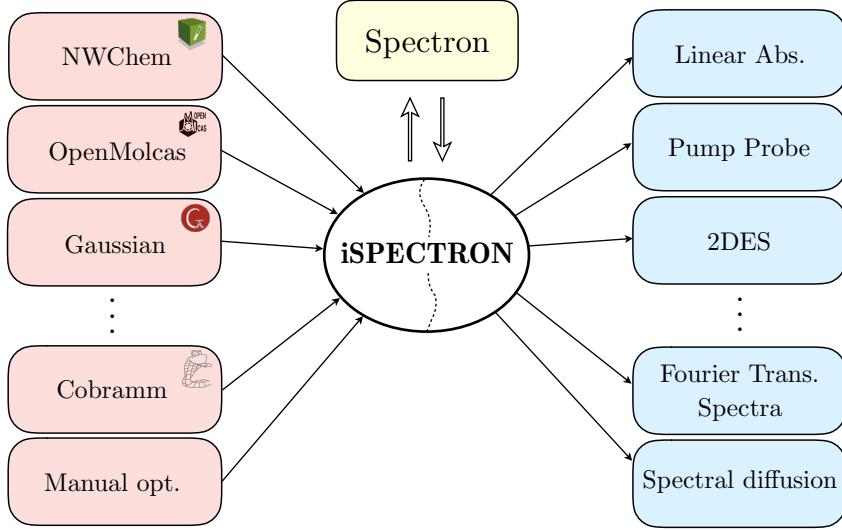


Figure 1: Schematic representation of the iSPECTRON links with quantum chemistry software, its double-sided flow of data with the Spectron software, and the range of possible spectroscopic techniques and properties that can be produced with it.

Electronic structure methods

The essential quantities for the simulation of spectroscopy are the manifold of energies of each state and the transition dipole moments between these states. Additionally, calculations of normal modes and frequencies at the ground state (GS) minimum, and evaluation of vertical gradients for the states of interest may be required. All of these properties may be obtained at the QM level by employing either WF-based or DFT-based approaches. Here, our focus is on multi-configurational WF techniques labelled as multi-configurational self-consistent field (MCSCF) methods, and the techniques emerging through the linear-response time-dependent (TD) extension of DFT categorized as TDDFT methods.

Multi-configuration methods

A many-electron wavefunction (WF) can be generally written as[82]:

$$|\Psi^{el}\rangle = c_0 |\psi_0\rangle + \sum_{r,a} c_a^r |\psi_a^r\rangle + \sum_{rs,ab} c_{ab}^{rs} |\psi_{ab}^{rs}\rangle + \dots \quad (1)$$

where $|\psi_0\rangle$ is the closed shell determinant built on top of the molecular spin orbitals (MOs), while $|\psi_a^r\rangle$, $|\psi_{ab}^{rs}\rangle$, and so forth, are the singly, doubly, ..., excited determinants, respectively. The coefficients in front of the various determinants are known as configuration interaction (CI) coefficients. This straight-forward expansion of the WF can treat closed and open shell electronic states on an equal footing, thus allowing for a description of photochemical processes such as bond breaking, dissociation and passage through a conical intersection. They also facilitate the calculation of transition probabilities between pairs of states. The factorial increase of the full CI basis with the system size limits the practical application of full CI methods to small systems and calls for additional approximations.

The complete active space self consistent field (CASSCF) method[71, 72] is a multi-configurational WF-based method in which the WF (eq. 1) is expanded over a restricted number of user-selected (active) molecular orbitals and electrons which constitute the *active space* (AS). In contrast to CI techniques, CASSCF WF is obtained variationally by simultaneously optimizing the CI and MO coefficients. Provided that the AS is properly assembled, this procedure can provide a qualitatively correct description of the multi-configurational nature of the electronic states along the potential energy surface. Quantitative results are obtained by utilizing second-order perturbation theory (CASPT2[3]) on top of the CASSCF WF which describes the missing dynamic correlation, i.e. the relaxation of the electrons outside the AS.

The CASSCF /CASPT2 protocol has been extensively used over the past 30 years and is widely considered the gold standard for a number of applications in photochemistry[70]. The major bottlenecks of the protocol remains, however, the CI computational cost which currently restricts the AS size of the conventional CASSCF/CASPT2 scheme to maximally 16 electrons and 16 orbitals. Recent years have witnessed a renaissance in the exploration of ways to cope with the CI cost. Several notable techniques such as the density matrix renormalization group (DMRG)[16, 6], the full configuration interaction quantum Monte Carlo (FCIQMC) approach[50], the variational two-electron reduced-density-matrix-driven (v2RDM)-CASSCF method[24] and the Heat-bath Configuration Interaction (HCI) active space solver[80], allow to handle AS with many tens of orbitals. Other approaches have focused on mitigating the factorial increase problem of CI by removing “configurational

deadwood” from the list of configuration state functions by introducing a flexible approach to the construction of the AS. Restricted active space (RAS)SCF/RASPT2[61, 49] and their generalized version GASSCF/GASPT2[46, 47], as well as the closely related occupation restricted multiple active space ORMAS/MRMP2 method[32, 18], also allow the construction of AS with few tens of orbitals. A simple but elegant projection technique, in combination with RASSCF theory has made accessible single and multiple core-valence transition energies and dipole moments for simulation of X-ray spectroscopy.[45]

The CASSCF/CASPT2 and RASSCF/RASPT2 methods are available in packages such as Gaussian and OpenMolcas, whereas OpenMolcas comes equipped with some of the recent CI solvers including DMRG and FCIQMC, as well as the more generalized GASSCF framework.

With the wavefunction at hand, it is possible to obtain (virtually) all properties necessary for the spectroscopy simulation: state energies, permanent and transition dipole moments between them, frequencies, gradients, etc. The CASSCF/CASPT2 protocol has been utilized in the past few years to simulate transient absorption, two-dimensional IR, Raman and electronic spectra of organic compounds (e.g. pyrene, uracil and thio-uracil, azobenzene) from first principles.

Density Functional and Time-Dependent Density Functional Theory

Density functional theory (DFT) is the most widely used theory for electronic structure calculations. Since it offers a good compromise between accuracy and computational cost, it is broadly applicable over the chemical and materials space across the periodic table. Formally, DFT is founded on the Hohenberg-Kohn theorems and is an exact theory with the electron density $\rho(\mathbf{r})$ serving as the fundamental variable instead of the wavefunction.[31] The computational efficiency of DFT arises from the exact mapping of the many-electron problem onto a set of single particle Schrödinger-like equations, known as the Kohn-Sham equations[34], where the complexities of the many-electron interactions are subsumed into the exchange-correlation functional, which is unknown.[64, 21] Approximate expressions, which can be divided into different families[65, 36], have been developed over the years. A detailed discussion of the pros and cons of DFT is beyond the scope of this paper and we

refer the interested reader to more comprehensive reviews on the subject.[64, 21, 20, 68]

Time-dependent density functional theory (TDDFT) is the extension of DFT to excited states and is also widely applicable over a broad range of systems.[15] Like DFT, TDDFT is also formally exact and founded on the Runge-Gross theorem[73], with the time-dependent electronic density, $\rho(\mathbf{r}, t)$, as the fundamental variable. However, in practice, it relies on the same exchange-correlation approximations as in DFT, thereby limiting its applicability to single excitations.[51, 48] TDDFT is typically formulated either in the frequency domain (or linear response) via the Casida equations[14] or in real-time.[43] In this paper, we utilize the former. Two other key limitations of traditional TDDFT include the failure to describe double excitations[22] and describing the correct topology near conical intersections.[39] Despite these limitations, TDDFT has been successfully utilized to study both valence[37] as well as core-level[44, 89] molecular and materials properties.[29, 85, 54] TDDFT developments remain an active area of research. We refer the reader to more comprehensive reviews.[62, 83, 51, 48]

Because of their broad applicability and computational efficacy, DFT and TDDFT have been implemented in all established quantum chemistry packages. Here, we utilize the implementation in NWChem[84, 4], including newly developed capabilities to calculate excited-state transition dipole moments using the TDDFT pseudo-wavefunction approach.

Including Environment effects

The study of molecular systems often requires to consider their native condensed-phase environment. Generally, homogeneous environment (i.e. a solvent) is described implicitly within an analytical framework known as the polarizable continuum model (PCM)[41]. PCM captures the polarization effect of the solvent shell on the electronic structure, thereby profiting from simplicity (in terms of input preparation) and generality (can be coupled to any quantum-chemical method). Various flavours of PCM that can be used in combination with both WF and DFT methods are found in every QM package. Instead, heterogeneous environment (e.g. the binding pocket of a protein) is usually treated with hybrid schemes where a high level QM description is applied to a spatially confined photo-responsive center (high layer), while a low level theory is applied to the typically much larger environ-

ment (low layer). If molecular mechanics (MM) is employed as the low level method the method is known as QM/MM. Schemes relying on conventional ‘electrostatic embedding’ such as ONIOM[19] or COBRAMM[86] provide flexible layer compartmentalization such as QM(high-level)/QM(low-level)/MM or QM/MM(movable)/MM(frozen). ONIOM is implemented within the Gaussian package and can profit from all the QM and semi-empirical methods available therein but it can be also coupled to an external program for the calculations for one or more layers. COBRAMM is a QM/MM dedicated interface which couples various QM software (e.g. Gaussian, OpenMolcas) with the MM suite Amber and relies on the Gaussian optimizer to perform geometry optimizations.

In recent years, next generation ‘polarizable embedding’ schemes have been established, relying on inclusion of atom multipole moments in addition to the charges or on polarizable point dipoles and treating the mutual polarization of the QM and MM parts self-consistently.[8, 42]

Spectroscopy simulations

The term “spectroscopy” comprises the conspicuous body of techniques that use radiation to obtain information on the structure and dynamical properties of matter. In the dipole (long wavelength) approximation, where the field does not vary across the molecule, the system is described by the Hamiltonian

$$\hat{H} = \hat{H}_{mol} - \hat{\boldsymbol{\mu}} \cdot \mathbf{E}(t) \quad (2)$$

where the first term represent the molecular Hamiltonian, which can be handled via quantum chemistry approaches; the second term represents the (semi-classical) interaction between the external (classical) electromagnetic field and the molecular system, through the dipole operator $\hat{\boldsymbol{\mu}}$.

A density matrix (rather than a wavefunction) representation of the system is required when dealing with open quantum systems and time-dependent perturbation theory is applied with respect to the field-matter interaction term: different orders of the perturbative expansion give rise to specific signals. We focus here on first- (linear) and third-order (nonlinear) signals.

The interaction between the external electromagnetic field and the electrons produces a

time-dependent polarization. This is the source of a new field, which eventually creates the measured signal of spectroscopic experiments.

The linear and nonlinear response functions

The linear (first-order) polarization is connected to the incident radiation field through the linear relation [1]

$$\mathbf{P}^{(1)}(\mathbf{r}, t) = \int_0^\infty dt_1 R^{(1)}(t_1) E(\mathbf{r}, t - t_1) \quad (3)$$

$R^{(1)}(t_1)$ is the (first-order) system *response function*, and contains the information of the first-order perturbed system. It may be recast as:

$$R^{(1)}(t_1) = \left(\frac{i}{\hbar} \right) \text{Tr} [\hat{\mu} \mathbb{G}(t_1) \mu^\times \hat{\rho}_0] \quad (4)$$

where $\hat{\rho}_0$ is the equilibrium density matrix (i.e. prior to any field interaction), $\mathbb{G}(t_1)$ is a superoperator that describes the density matrix evolution in absence of the external field (defined as $\mathbb{G}(t) \bullet = e^{-\frac{i}{\hbar} \hat{H}_{mol} t} \bullet e^{\frac{i}{\hbar} \hat{H}_{mol} t}$), and μ^\times is the dipole superoperator (defined as $\mu^\times \bullet = [\hat{\mu}, \bullet]$) that accounts for the interaction between the system and the external field.

Similarly, the third-order polarization reads [1]

$$\begin{aligned} \mathbf{P}^{(3)}(\mathbf{r}, t) = & \int_0^\infty dt_3 \int_0^\infty dt_2 \int_0^\infty dt_1 R^{(3)}(t_3, t_2, t_1) \times \\ & E(\mathbf{r}, t - t_3) E(\mathbf{r}, t - t_3 - t_2) E(\mathbf{r}, t - t_3 - t_2 - t_1) \end{aligned} \quad (5)$$

where the corresponding response function, $R^{(3)}(t_3, t_2, t_1)$, is given by

$$R^{(3)}(t_3, t_2, t_1) = \left(\frac{i}{\hbar} \right)^3 \text{Tr} [\hat{\mu} \mathbb{G}(t_3) \mu^\times \mathbb{G}(t_2) \mu^\times \mathbb{G}(t_1) \mu^\times \hat{\rho}_0] \quad (6)$$

Eqs. 4 and 6, can be interpreted as follows:

1. Start from the the equilibrium density matrix ρ_0 ;
2. Interact once with the laser via the dipole superoperator μ^\times ;
3. Evolve the perturbed density matrix by the field-free molecular Hamiltonian H_{mol} for a time t_1 ;
4. Repeat steps 2. and 3. for times t_2 and t_3 ;

5. Compute, with the perturbed and time-dependent density matrix $\rho^1(t_1)$ ($\rho^3(t_1, t_2, t_3)$), the first (third) order response function $P^{(1)}(t_1) = \text{Tr} [\mu \rho^{(1)}]$ ($P^{(3)}(t_1, t_2, t_3) = \text{Tr} [\mu \rho^{(3)}]$);

The polarization must be averaged over all possible system orientations to represent a randomly oriented sample.³

A general solution to eqs. 4 and 6 can be derived[53] which applies to arbitrary level schemes (see Appendix). In the case of third-order polarization, the number of terms that appear in the general expression is very large due to the various Liouville space pathways⁴ and the permutation over the incident fields. Therefore, the response is often recast for approximate models of two or three distinct manifolds (i.e. bands, see section “Sum-over-states expansion” for definition of the manifolds) in combination with temporally well separated pulses, resonant with the energy gaps between the manifolds. Fortunately, in resonant conditions, only a few of the terms survive when a specific (called phase-matched) combination of the incoming fields wave vectors $\mathbf{k}_i, i = 1, 2, 3$ is chosen, as highly oscillatory terms (where molecular and field frequencies) add up rather than subtract can be safely neglected, an approximation known as the rotating wave approximation (RWA). Experimentally, different types of linear and nonlinear signals that radiate from the sample can be distinguished by placing the detector in different directions, identified by a specific combination of the incoming fields wave vectors $\pm \mathbf{k}_1 \pm \mathbf{k}_2 \pm \mathbf{k}_3$.

Signals in a given direction can be further decomposed into ground state bleaching (GSB), stimulated emission (SE) and excited state absorption (ESA)⁵. The nomenclature reflects the physical processes induced by the light-matter interactions.[1] The response function of eq. 6 in a given phase-matching direction can be then expressed as the sum of the individual (GSB, SE, ESA) contributions. Notably, ESA contributions have an opposite sign to GSB and SE contributions, thus their spectral overlap could potentially lead to

³Note that in some specific cases the systems in the sample may be globally oriented in a specific direction: in that case such orientational average should not be performed.

⁴The Liouville pathways are the terms that appear when the action of the super-operator μ^\times is written out explicitly: being a commutator, it can act either on the left or on the right, producing multiple contributions known as Liouville space pathways.

⁵Note that, in a real experiment, GSB, SE and ESA associated with the same phase-matching condition cannot be detected separately.

complex interference patterns and occasionally to signal cancellation. For this reason, the calculation of the individual components offered by Spectron is essential.

In linear spectroscopy the time-dependent polarization (eq. 3) is Fourier transformed with respect to t

$$\mathbf{P}^{(1)}(\Omega) = \int_0^\infty dt \mathbf{P}^{(1)}(t) \exp\{i\Omega t\} \quad (7)$$

to provide an intensity vs. Ω spectrum - the LA spectrum. In third-order spectroscopy (rephasing, non-rephasing and quasi absorptive signals) any pair of time variables can be used to generate a 2D plot. The time-dependent polarization (eq. 5) is usually Fourier transformed with respect to variables t_1 and t_3

$$\mathbf{P}^{(3)}(\Omega_1, t_2, \Omega_3) = \int_0^\infty dt_3 \int_0^\infty dt_1 \mathbf{P}^{(3)}(t_3, t_2, t_1) \exp\{i\Omega_3 t_3 + i\Omega_1 t_1\} \quad (8)$$

to provide a two-dimensional map (intensity vs. Ω_1 vs. Ω_3) - the 2D spectrum - for each (waiting) time t_2 . The double quantum coherence signal is Fourier transformed with respect either to intervals t_1 and t_2 or t_2 and t_3 . Finally, the transient absorption (pump probe) spectra (intensity vs. Ω_3 at different t_2 times) can be readily obtained from the two-dimensional maps by integrating over Ω_1 .

$$\mathbf{P}^{(3)}(t_2, \Omega_3) = \int_0^\infty d\Omega_1 \mathbf{P}^{(3)}(\Omega_1, t_2, \Omega_3) \quad (9)$$

Sum-over-states (SOS) expansion: connecting the electronic structure of the system to a spectroscopic signal

In the SOS approach, the response function is expanded in the eigenstate basis. These states are usually divided in three sets: on the bottom of the energy ladder there is the unique ground state (denoted by g). Above it, there is the single-excitation manifold (with states denoted by the index e). Finally, one finds the double-excitation manifold (index f). The manifolds are connected by the dipole operator μ , and we assume that only transitions between $g \rightarrow e$ (characterized by energy ω_{eg} and transition dipole moment μ_{eg}) and $e \rightarrow f$ (ω_{fe} and μ_{fe}) can occur (Figure 2(d)). Linear techniques only interrogate states of the first-excitation manifold, while in third-order techniques the second excitation manifold should be considered as well.

We note that the terms “single-excitation” and “double-excitation” are historically related to the study of excitonic systems and referred to exciton states having a single site or two sites simultaneously excited.[1, 90] In the framework of the current level scheme (Figure ??), the f -manifold can include states generated by single-excitation from the ground state which are coupled to states from the e -manifold by a non-vanishing transition dipole moment. Furthermore, as typical in long conjugated systems, doubly excited states can lie among the lowest excited states of the system and, while not coupled to the ground state (i.e they are dark) they participate in internal conversion. A more convenient way to think of the e and f manifolds is in terms of being resonant with the energy of the pump (e manifold, ω_{eg} energies) and probe (f manifold, ω_{fe} energies) pulses.[76, 57] Thus, the assignment of states to either manifold is connected to the experimental conditions.

The f states are not directly involved in the photo-induced dynamics (during t_2). However, as each e state has its own characteristic manifold of higher lying f states the ESA associated with their population upon interaction with the probe pulse can serve as indirect signatures of the underlying dynamics in the photo-active states. This is particularly valuable for resolving states intermediately populated during the non-adiabatic decay which are dark with respect to the ground state (i.e. dark or phantom states). Thus, states from the f manifold could be considered “spectator states” in the terminology of transient IR spectroscopy.[76, 5]

Three non vanishing (under the RWA) independent third-order signals can be obtained for the three-level scheme (Figure 2(d)) from different combinations of the wave vectors of the incident pulses \mathbf{k}_1 , \mathbf{k}_2 and \mathbf{k}_3 . The signals are referred to as rephasing ($\mathbf{K}_I = -\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3$), non-rephasing ($\mathbf{K}_{II} = +\mathbf{k}_1 - \mathbf{k}_2 + \mathbf{k}_3$) and double-quantum coherence ($\mathbf{K}_{III} = +\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3$)(see Liouville space ladder diagrams in Fig. 2). The various signals encompass a different complementary information about the system. For example, following the above interpretation in the SE contribution to the non-rephasing response function $R_{SE}^{(\mathbf{K}_{II})}(t_1, t_2, t_3)$ (Fig. 2) (middle), after the first interaction with the incident pulse, the molecular system density matrix evolves in a coherence state (ρ_{ge}) during the t_1 interval, i.e. the *ket* and *bra* states of the density matrix propagate under the influence of the ground sate (g) and excited state (e) Hamiltonians, respectively. The interaction with the second pulse creates the density matrix element $\rho_{e'e}$, which can be either a population state ($e = e'$) or a coherence

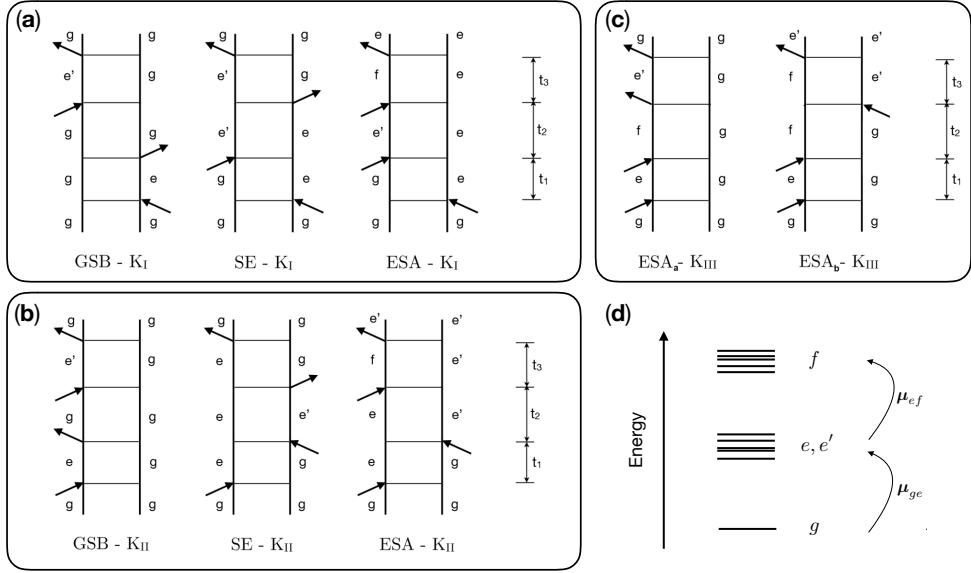


Figure 2: Ladder diagrams for the signal generated in the direction \mathbf{K}_I (a), \mathbf{K}_{II} (b) and \mathbf{K}_{III} (c); GSB: ground-state bleaching; SE: excited-state emission; ESA: excited-state absorption. (d) Energy level diagram and allowed transitions between the states in the various manifolds.

state ($e \neq e'$), which evolves during the t_2 interval. Finally, after interaction with the third pulse the density matrix evolves in a coherence state $\rho_{e'g}$.

By adding the simulated/recorded rephasing and non-rephasing contributions one obtains (quasi-)absorptive spectra which provide the highest spectral resolution thanks to favorable cancellation of dispersive spectral components. The simulation of the various third-order signals is an integral part of the Spectron code.

In the SOS framework the first- and third-order response functions (eqs. 4 and 6) are expanded in terms of the system eigenstates. $R^{(1)}(t_1)$ thus becomes:

$$R^{(1)}(t_1) = \left(\frac{i}{\hbar} \right) \theta(t_1) \sum_e |\boldsymbol{\mu}_{eg}|^2 \exp\{-i(\omega_{eg} - i\gamma_{eg})t_1\} + c.c. \quad (10)$$

We have added a phenomenological dephasing rate γ_{eg} to the transitions. This accounts for the decay of the coherence and the associated signal broadening in an open system due to processes such as pure dephasing, internal conversion, etc., whose detailed and explicit description is omitted at this lowest⁶ level of theory.

⁶“Lowest” in terms of the description of the signal line-shapes.

The linear absorption spectrum represents the real (absorptive) part of the Fourier transformed polarization $\text{Re}\{\mathbf{P}^{(1)}(\Omega)\}$ (eq. 3).

$$\mathbf{P}^{(1)}(\Omega) = \sum_e \frac{|\boldsymbol{\mu}_{eg}|^2}{\Omega - \omega_{eg} + i\gamma_{eg}} \quad (11)$$

It is given as a sum of transitions centered at ω_{eg} , having Lorenzian line-shapes whose widths are governed by the dephasing rates γ_{eg} .

Similarly, the SE diagram of the non-rephasing third-order signal becomes:

$$R_{SE}^{(\mathbf{K}_{II})}(t_1, t_2, t_3) = \left(\frac{i}{\hbar}\right)^3 \theta(t_1)\theta(t_2)\theta(t_3) \sum_{ee'} |\boldsymbol{\mu}_{eg}|^2 |\boldsymbol{\mu}_{e'g}|^2 \times \exp\{-i(\omega_{ge} - i\gamma_{ge})t_1 - i(\omega_{e'e} - i\gamma_{e'e})t_2 - i(\omega_{e'g} - i\gamma_{e'g})t_3\} \quad (12)$$

whereas the spectrum resulting from the Fourier transform of the third-order polarization (eq. 5)

$$\mathbf{P}_{SE}^{(\mathbf{K}_{II})}(\Omega_1, t_2, \Omega_3) = \sum_e \frac{|\boldsymbol{\mu}_{eg}|^2}{\Omega_1 - \omega_{eg} + i\gamma_{eg}} \cdot \frac{|\boldsymbol{\mu}_{e'g}|^2}{\Omega_3 - \omega_{e'g} + i\gamma_{eg}} \exp\{-i(\omega_{e'e} - i\gamma_{e'e})t_2\} \quad (13)$$

has a more complex line shape with positive and negative regions (phase twists). Similar expressions can be derived for the remaining third-order diagrams[53, 1]. Within the framework of eqs. 11 and 13, the necessary QM inputs to the spectroscopy simulation with Spectron are limited to the manifold of states energies (from which one builds the required energy differences ω_{eg} and ω_{fe}) and the transition dipole moments ($\boldsymbol{\mu}_{eg}$ and $\boldsymbol{\mu}_{fe}$) between them. The dephasing rate γ may be fitted to describe the observed line broadening.

The line shapes: incorporating the coupling to a harmonic bath

Modelling the line shapes requires to describe the field-free dynamics of the system coupled to a bath. This can be realized at various levels of sophistication: a) the exact numerical solution of the Liouville equation in the reduced system space; b) semi-classical and mixed-quantum classical trajectory-based approaches which approximate or neglect altogether some quantum mechanical aspects (ranging from accurate path-integral methods, that can be made arbitrarily accurate, to crude phenomenological separations between quantum and classical degrees of freedom); c) methods based on parametrized Hamiltonians. The latter

rely on the (strong) approximation that the potential energy surfaces of the electronic states of interest can be expressed by analytical expressions (e.g harmonic potentials) which can be parametrized by quantum mechanical calculations (energies, gradients, normal modes etc.). The appeal of this approach is boosted by the availability of analytical expressions for the signal line shapes. Therefore, facilitating line shape description based on the coupling between the electronic structure to a model Hamiltonian describing the intra-molecular and the environment vibrational degrees of freedom (DOF) is naturally the first strategy made available through the iSPECTRON interface that goes beyond purely phenomenological line shapes.

Here we outline the displaced harmonic oscillator (DHO) model Hamiltonian used to describe the coupling to intra-molecular DOF and the resulting line shape functions. The model assumes that the potential energy surfaces of the electronic states are quadratic in the coordinates, with different equilibrium positions (i.e. electronic state a is *displaced* by \tilde{d}_{ak} along the k -th mass-weighted normal mode coordinate \tilde{q}_k). It is also assumed that the different electronic states are described by the same normal modes and frequencies (which are thus computed only for a single electronic state, generally being the electronic GS)⁷. The displacements \tilde{d}_{ak} of the various ES wells with respect to the GS can be computed with the vertical gradient approach, implemented in iSPECTRON, i.e. by projecting energy gradients for the states of interest onto the GS normal modes. Starting with the expression for the energy gap $\Delta E_{ab}(\tilde{q}_k)$ between states a and b in the framework of the DHO model

$$\Delta E_{ab}(\tilde{\mathbf{q}}) = \varepsilon_a - \varepsilon_b + \sum_k \frac{1}{2} \omega_k^2 (\tilde{d}_{ak}^2 - \tilde{d}_{bk}^2) - \omega_k^2 (\tilde{d}_{ak} - \tilde{d}_{bk}) \tilde{q}_k \quad (14)$$

$\varepsilon_a - \varepsilon_b$ is the adiabatic excitation energy between states a and b . The gradient of the energy gap $\nabla_{\tilde{\mathbf{q}}} \Delta E_{ab}(\tilde{\mathbf{q}})$ reads

$$\nabla_{\tilde{\mathbf{q}}} \Delta E_{ab}(\tilde{\mathbf{q}}) = [\nabla_{\mathbf{Q}} E_a(\mathbf{Q}) - \nabla_{\mathbf{Q}} E_b(\mathbf{Q})] \nabla_{\tilde{\mathbf{q}}} \mathbf{Q} = \mathbb{P}^\dagger \mathbb{M}^{-\frac{1}{2}} (\mathbf{f}_a - \mathbf{f}_b) \quad (15)$$

where \mathbb{P} is a matrix whose columns are the normal modes $\tilde{\mathbf{q}}$ expressed in normalized mass-weighted Cartesian coordinates \mathbf{Q} , \mathbb{M} is a diagonal matrix with the nuclear masses and

⁷Formulations for the linear response beyond this approximation, which incorporate so called Dushinsky rotation matrices, exist.[23]

$\mathbf{f}_{a/b} = \nabla_{\mathbf{Q}} E_{a/b}(\mathbf{Q})$ are the Cartesian energy gradient in electronic state a and b . Taking the derivative with respect to $\tilde{\mathbf{q}}$ in eq. 14 and inserting the expression derived in eq. 15 gives

$$\tilde{\mathbf{d}}_a = -\mathbb{W}^{-2}\mathbb{P}^\dagger\mathbb{M}^{-\frac{1}{2}}(\mathbf{f}_a - \mathbf{f}_b) + \tilde{\mathbf{d}}_b \quad (16)$$

where $\tilde{\mathbf{d}}_a$ and $\tilde{\mathbf{d}}_b$ are k -dimensional arrays of the displacements along all modes on states a and b .

The state-specific displacements along each normal mode k , i.e. \tilde{d}_{ak} , are directly related to spectroscopic observables like the dimensionless Huang-Rhys factors S_{ak} or the reorganization energy λ_{ak} through the following relations:

$$S_{ak} = \frac{\tilde{d}_{ak}^2 \omega_k}{2\hbar} \quad \text{and} \quad \lambda_{ak} = \frac{\tilde{d}_{ak}^2 \omega_k^2}{2\hbar}. \quad (17)$$

and are processed by the interface script to formulate the spectral density functions $J_{ab}(\omega)$ which represent the frequency dependent strength of (linear) system-bath coupling[53, 1]

$$J_{ab}(\omega) = \pi \sum_k \tilde{d}_{ak} \tilde{d}_{bk} \omega_k^3 [\delta(\omega - \omega_k)] \quad (18)$$

The spectral density is the central quantity in the formulation of the line shape functions $g_{ab}(t)$ [53, 1]

$$g_{ab}(t) = -\frac{1}{2\pi} \int_0^{+\infty} d\omega \frac{J_{ab}(\omega)}{\omega^2} \left[\coth\left(\frac{\beta\omega}{2}\right)(1 - \cos(\omega t)) + i(\sin(\omega t) - \omega t) \right] \quad (19)$$

that replaces the phenomenological broadening in the linear and nonlinear response. The precise form of the response function is derived via the cumulant expansion of Gaussian fluctuations (CGF)[53]. For linear techniques, the first-order response function becomes

$$R^{(1)}(t_1) = \left(\frac{i}{\hbar}\right) \theta(t_1) \sum_e |\boldsymbol{\mu}_{eg}|^2 \exp\{-i(\omega_{ge} - i\tau_{ge}^{-1}) - 2g_{ee}(t_1)\} + c.c. \quad (20)$$

τ_{eg}^{-1} here indicates the rate of coherence decay due to finite lifetime (τ_{eg}) associated with internal conversion, energy or charge transfer (thus, representing part of the phenomenological broadening γ_{ab}) which causes a homogeneous broadening of the spectral lines.

For the third-order response function we again consider the same diagram shown in eq. 12, which now becomes

$$\begin{aligned} R_{SE}^{(K_{II})}(t_1, t_2, t_3) &= \left(\frac{i}{\hbar}\right)^3 \theta(t_1)\theta(t_2)\theta(t_3) \sum_{ee'} |\boldsymbol{\mu}_{eg}|^2 |\boldsymbol{\mu}_{e'g}|^2 \times \\ &\times \exp\{-i(\omega_{ge} - i\tau_{ge}^{-1})t_1 - i(\omega_{e'e} - i\tau_{e'e}^{-1})t_2 - i(\omega_{e'g} - i\tau_{e'g}^{-1}) + \varphi_{ee'}^{SE}(t_1, t_2, t_3)\} \end{aligned} \quad (21)$$

and $\varphi_{ee'}^{SE}(t_1, t_2, t_3)$ is a multipoint line-shape function given by a specific combinations of $g_{ab}(t)$ evaluated at different times[53, 1]

$$\begin{aligned}\varphi_{ee'}^{SE}(t_1, t_2, t_3) = & -g_{e'e'}(-t_2) - g_{ee}(t_1 + t_2 + t_3) - g_{ee'}(t_1) \\ & + g_{ee'}(-t_2 - t_3) + g_{ee'}(t_1 + t_2) - g_{ee'}(-t_3)\end{aligned}\quad (22)$$

Similar expressions for the multipoint line-shape functions which enter in the remaining third-order diagrams are readily implemented in Spectron.[1]

The coupling to the environment (e.g. a protein cage, a solvent, a nearby molecule, etc.) may also be included in the form of spectral densities, and treated within the CGF framework presented above. While it can be computed from first principles based on molecular dynamics simulations[77, 38], the motion of the environment can be also accounted for via model functions, as it is usually characterized by smooth spectral densities with significant contributions mainly in the low frequency regime. For example, the overdamped brownian oscillator (OBO) spectral density[13, 40] available in Spectron has the form

$$J_{ab}(\omega) = 2\lambda_b \frac{\omega\Lambda_b}{\omega^2 + \Lambda_b^2} \quad (23)$$

and its parameters (the cutoff frequency Λ_b and coupling strength λ_b) can be fitted to reproduce the experimental line shapes. In general the spectral densities describing the coupling to the environment are state-specific, e.g. the energies of ionic state are more susceptible, i.e. strongly coupled, to the environment fluctuations.[77]. Nevertheless, to a first approximation a single OBO spectral density is commonly used for all states.

The decomposition of the spectral density into intra-chromophore and environment contributions is a reasonable approximation as energy fluctuations caused by inter- and intra-molecular vibrations (i.e. by short- and long- range interactions) can be considered uncorrelated.

Beyond the adiabatic approximation: incorporating internal conversion, charge and energy transfer

The CGF approach is exact for fluctuations with Gaussian statistics and in the absence of electronic relaxation. Approximate expressions may be introduced to incorporate internal-conversion or energy transfer between different states. In the limit of decoupled evolution

of the populations and coherence terms in the density matrix SE and ESA diagrams are separated in population (i.e. $e = e'$ during t_2) and coherence (i.e. $e \neq e'$ during t_2) terms. Population relaxation during t_2 is described by the Pauli master equation[1]:

$$\dot{\rho}_{ee}(t) = - \sum_{e'} K_{e'e',ee} \rho_{e'e'}(t) \quad (24)$$

with \mathbf{K} the rate matrix with elements $K_{e'e',ee}$ depicting the population transfer rate from state e into state e' . The solution of this equation is formally recast by the population Green's function $\rho_{e'e'}(t) = \sum_e \mathbf{G}_{e'e',ee}(t) \rho_{ee}(0)$ and the elements of the matrix \mathbf{G} act as time-dependent weighting factors in the population pathways describing SE and ESA. For example, the population SE diagram of the k_{II} signal becomes

$$R_{SE,pop}^{(\mathbf{K}_{II})}(t_1, t_2, t_3) = \left(\frac{i}{\hbar} \right)^3 \theta(t_1) \theta(t_2) \theta(t_3) \sum_{ee'} |\boldsymbol{\mu}_{eg}|^2 |\boldsymbol{\mu}_{e'g}|^2 \mathbf{G}_{e'e',ee}(t_2) \times \times \exp \left\{ -i (\omega_{ge} - i\tau_{ge}^{-1}) t_1 - i (\omega_{e'g} - i\tau_{e'g}^{-1}) + \varphi_{ee'}^{SE}(t_1, t_2, t_3) \right\} \quad (25)$$

whereas the coherence SE diagram retains its form from eq. 21.

The rates τ_{ab}^{-1} entering the rate matrix $K_{e'e',ee}$ can be obtained from ab-initio computations that treat the coupling between different states (such as quantum-dynamics, semi-classical approaches equipped with surface hopping, Redfield/Förster transport methods in the context of excitons, etc.)[12] or may be fitted from available experimental information.[55] Finally, we note that Spectron offers variations of the multipoint line shape function $\varphi(t_1, t_2, t_3)$ depending on the relative time scales of transport, bath fluctuations and coherence dephasing in the system[1]. The different formulations are all expressed in terms of line shape functions $g_{ab}(t)$.

Beyond impulsive excitations: incorporating pulse envelopes

So far we have shown how to compute the response function for linear and nonlinear (third-order) techniques at various levels of sophistication. The response function coincides with the signal itself in the limit of infinitely short laser pulses (known as the “impulsive limit”). Within this limit, the pulses are solely characterized by their wave-vector and mutual time delays.

To include the realistic pulse envelopes, the response function has to be convoluted with the fields, as prescribed by eqs. 3 and 5. The effect of the fields, which is entirely captured by the multiple integrals of these equations, can be accounted for in an approximate way by: (a) weighting the data along the frequency axes with the profile of the field (in the frequency domain) (b) convoluting the data along the time (t_2) to describe the effect of the finite time resolution that a realistic pulse (with finite duration) necessarily introduces. Exact equations that include the finite pulse-width effect when assuming temporally separated pulses and in absence of population transfer, are reported in Ref. [1].

Additionally, each pulse may have its own polarization: the polarization configuration of the pulse sequence weights differently the various contributions and affect the final spectra. The weights are given by the product of the field polarization $\hat{\nu}_i$ ($i \in x, y, z$) and the transition dipole moment μ_{ab} , i.e. $\hat{\nu}_i \cdot \mu_{ab}$. A posteriori linear combinations of signals obtained with different pulse polarization configurations can help highlight weak features like cross-peaks which are often masked by the more intense but less interesting diagonal features in the spectra.[59, 58]

Interface Workflow(s)

iSPECTRON is a command-line interface written in Python 3 designed for Unix platforms⁸. It is composed of two parts: part (a) is designed to digest QC outputs and some parameters specified by the user to produce the necessary input files for the Spectron code; part (b) reads through the Spectron output data and processes and manipulates the simulated spectra. The discussion that follows is thus developed in two sub-sections, one for each part of the interface.

Interface - part a

The interface reads in inputs and outputs of the QM codes employed for the QM calculation to retrieve the following information: transition energies, transition dipole moments and vertical gradients for all states of interest (i.e. states of the $e-$ manifold in case of linear techniques, states of the $e-$ and $f-$ manifolds for nonlinear techniques). Alternatively, this

⁸The code requires the NUMPY and SCIPY packages.

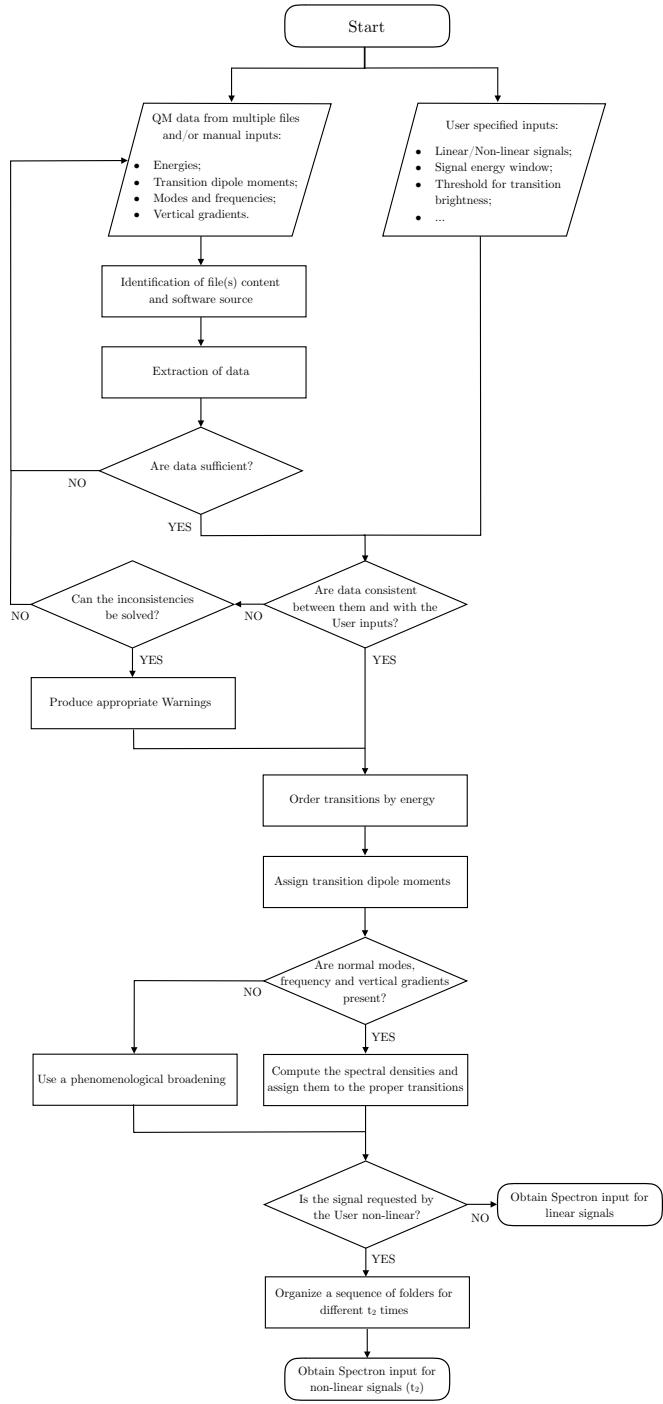


Figure 3: Workflow of the iSPECTRON program functioning for the collection and re-elaboration of quantum chemistry data

information can be provided manually by the user (via communication files with well-defined format, see Supporting Information). The QM inputs and outputs can constitute a homo-

geneous set (i.e. all obtained with the same QM code) or coming from different sources (as, e.g., QM/MM calculations with COBRAMM). The interface interprets the provided inputs/outputs and tags them according to the software of origin and the type of data inside the file. In addition to the system-specific data, the user needs to specify several parameters for the ‘virtual experiment’ such as the energy window, the type of signal of interest (linear or transient absorption, $\mathbf{K}_I - \mathbf{K}_{III}$, quasi-absorptive) and the simulation time t_2 . Optional parameters relevant to the system under scrutiny (e.g. phenomenological broadening or OBO parameters), the utilized pulses (e.g. polarization) and advanced simulation details (e.g., the selection of specific diagrams) can be also provided.

After all files have been collected and identified by the interface, the resulting input is analyzed for consistency. If there are no conflicts the code re-elaborates the data by ordering the manifold of energies and transition dipole moments and, if requested, computes spectral densities. Otherwise, if the inconsistencies are beyond remedy (e.g. no energies or transition dipoles are found or no transitions are present in the selected energy window) the code exits with an appropriate error message.

The code then prepares various intermediate files to be processed by Spectron. The minimal required information to run a basic Spectron simulation (in the SOS representation utilizing phenomenological broadening constants γ_{ab}) is as follows:

- The Spectron main input file, containing information about the chosen spectroscopic technique, the energy window, the number of states, etc.;
- The (single column) file with the list of transition energies relative to the GS, including the GS energy itself (zero by definition), in increasing order.
- The (five column) file with the transition dipoles; the first two indices represent initial and final states a and b , while the last three numbers are the x , y and z components of the transition dipole μ_{ab} .

For simulations which go beyond the phenomenological broadening of the spectral line-shapes, making use of the CGF framework of Spectron, a few additional files are also created:

- The (single column) file containing, in consecutive order, spectral densities $J_{ab}(\omega)$ associated with different states;

- The coupling file, that defines the coupling between transitions and spectral densities (a sequence of lower triangular $N_{states} \times N_{states}$ matrices, one for each spectral density, in which the entry ab is the coupling strength to the spectral density $J_{ab}(\omega)$);
- The transport rates file (square $N_{states} \times N_{states}$ matrix) which contains the rate matrix \mathbf{K} (24) with the inverse of each state lifetime on the diagonal and the energy transfer rates between the states off-diagonal;

Finally, two additional files, named INFO.txt and VIB_INFO.txt, are also produced: the former contains the command line utilized by the user to launch the interface; the latter, summarizes the states reorganization energy λ (eq. 17). The workflow of the first part of the interface is reported in Scheme 3.

Interface - part b

Once Spectron has finished simulating the linear or nonlinear signal, its output can be processed by the second leg of interface code.

In the case of linear techniques, the output provided by Spectron is a single two-column file (energy vs. intensity). For nonlinear techniques, instead, the output is organized in a directory tree structure:

- **top level**, named according to the simulated signal (e.g. KI). If a folder with the same name is already present (as in case of repeating a simulation with different parameters) an integer is attached to the tag (e.g. KI_1, KI_2);
- **second level** (optional), with branches according to the diagrams, i.e. GSB, SE or ESA;
- **third level**, with branches associated with the waiting time t_2 (in case of rephasing, non-rephasing and quasi-absorptive signals) or t_1 or t_3 (in case of double quantum coherence), e.g. $t2_0, t2_4 \dots t2_100$, in fs.

This allows to visualize and manipulate selectively one or multiple snapshots of individual diagrams constituting the overall signal.

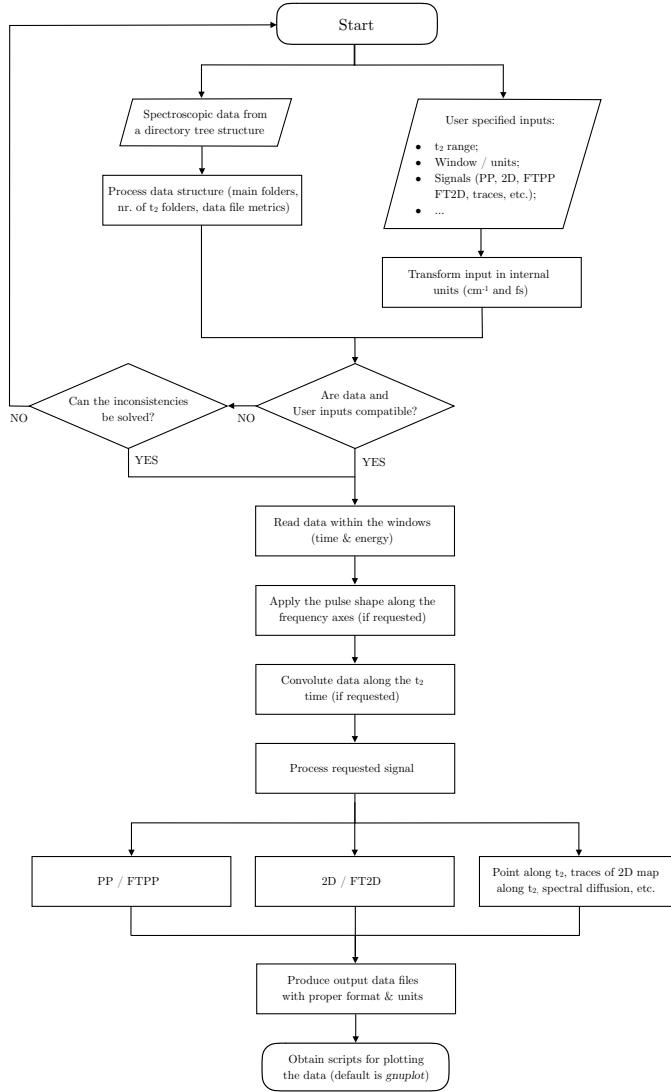


Figure 4: Workflow of the iSPECTRON program functioning for the production and analysis of a variety of spectra.

In the post-processing the user can choose to further manipulate the 2D spectra by convoluting them with realistic pulses, by integrating them to obtain their marginal, the TA spectra (eq. 9), or by Fourier transforming them along the t_2 time; one can request 1D and 2D spectral cuts or time traces and can adjust a number of additional optional parameters (e.g. the spectral window). Specifically, in the post-processing one has access to:

- One-dimensional transient absorption spectra (1DPP), that resolve the signal along a single frequency axis (the detection frequency Ω_3). Single t_2 times snapshots and/or

two-dimensional time/frequency heat-maps may be provided. The t_2 axis may be Fourier transformed, obtaining a frequency/frequency map that resolves the frequencies underneath the signal beating along t_2 , that may be different at different detection frequencies. This new frequency axis is labeled as Ω_2 ;

- Two-dimensional (2D) spectra, which resolve the signal along two frequency axes (excitation - Ω_1 - and detection - Ω_3 - frequency) are given as one or multiple t_2 snapshots (according to the user specifications). Maps at different Ω_2 frequencies may also be obtained by Fourier transform of the data along t_2 ;
- Cuts of 2D maps along specific excitation or detection frequencies may be performed. The code was implemented to perform cuts along any specified couple of points in the Ω_1/Ω_3 plane (such as diagonal or anti-diagonal cuts);
- The (t_2) time trace of the value of a selected Ω_1/Ω_3 2D map point (or the integrated value of a 2D map squared region) may be also extracted from the nonlinear data to visualize the beatings of the signals in different regions of the 2D map;
- Two approaches to extract the spectral diffusion of peaks of interest in a 2D map as a function of the waiting time t_2 are also implemented: i) evaluation of the slope of the line that connects the peak maxima along Ω_1 at every fixed Ω_3 value; ii) evaluation of the ratio between the widths of the diagonal and anti-diagonal cuts that goes through the center of the peak. In both cases, one expects to see how the observable approaches a value close to zero with t_2 , a signature of the erasure of the system memory of the initial excitation energy, caused by the system-bath interaction.

Finally, output files with all requested features are created, together with *gnuplot* input files, with all the information necessary to plot these data.

The workflow of the second part of the iSPECTRON interface code is reported in Scheme 4. Note that this second part of the code is independent from the first one: indeed, the Spectron code can be used with other levels of theory implemented therein (excitons, quasi-particle approach, stochastic models for the system-bath coupling, etc.) and the third-order response, if organized in the folder structure presented above, may be processed by

iSPECTRON.

EXAMPLES

In this section we demonstrate how the linear and nonlinear spectra generated with the iSPECTRON interface can be utilized to monitor the photo-induced vibrational dynamics and ultrafast internal conversion of pyrene. Pyrene is a widely used PAH owing to its interesting photophysical properties with a remarkably long excited state lifetime and high fluorescence yields. It presents well separated absorption bands with clear Franck-Condon progressions in the UV (at 320 nm and at 280 nm), ultrafast internal conversion to the lowest (dark) excited state and few characteristic ESA signatures, making it an excellent model case for assessing novel spectroscopic techniques.[35, 11, 66, 10, 9] The availability of high resolution experimental LA, PP and 2D maps naturally makes pyrene the system of choice to assess the capabilities of iSPECTRON.

First we present linear absorption of the *La* band - the lowest bright electronic excited state of pyrene absorbing in the near-UV - simulated at two levels of theory - RASSCF/RASPT2 (through OpenMolcas) and TDDFT (through NWChem) - and at two temperatures - at room temperature (i.e. 300 K) and at cryogenic temperature (i.e. 5 K). Subsequently, we report nonlinear spectra simulations at the RASSCF/RASPT2 level of theory. In all simulations coupling to the intra-molecular and environment vibrational DOF is described within the CGF framework (Section) in its most general realization capable of treating bath fluctuations with arbitrary time scales and population transfer[2]. The required spectral densities describing the coupling to the intra-molecular DOF are modeled based on the undamped DHO Hamiltonian utilizing the vertical gradient approach (eq. 16). The spectral density describing the coupling to the environment DOF is obtained with the OBO model (eq. 23) fitted to reproduce the width of the experimental spectra. The ultrafast $La \rightarrow Lb$ internal conversion is modeled on the basis of the rate eq. 24. Thereby, a lifetime τ_{eg} of 100 fs is assigned to the *La* state, in agreement with experimental findings[9]. Energies, transition dipole moments, normal modes, frequencies, and gradients have been

computed from first principles in vacuum⁹. D_{2h} symmetry was enforced for the QC computations behind the nonlinear spectra. Additional information on the quantum chemical calculations, simulation parameters as well as the iSpectron commands used to generate the plots in Figures 5 and 6 are reported in the Supporting Information.

Linear spectra

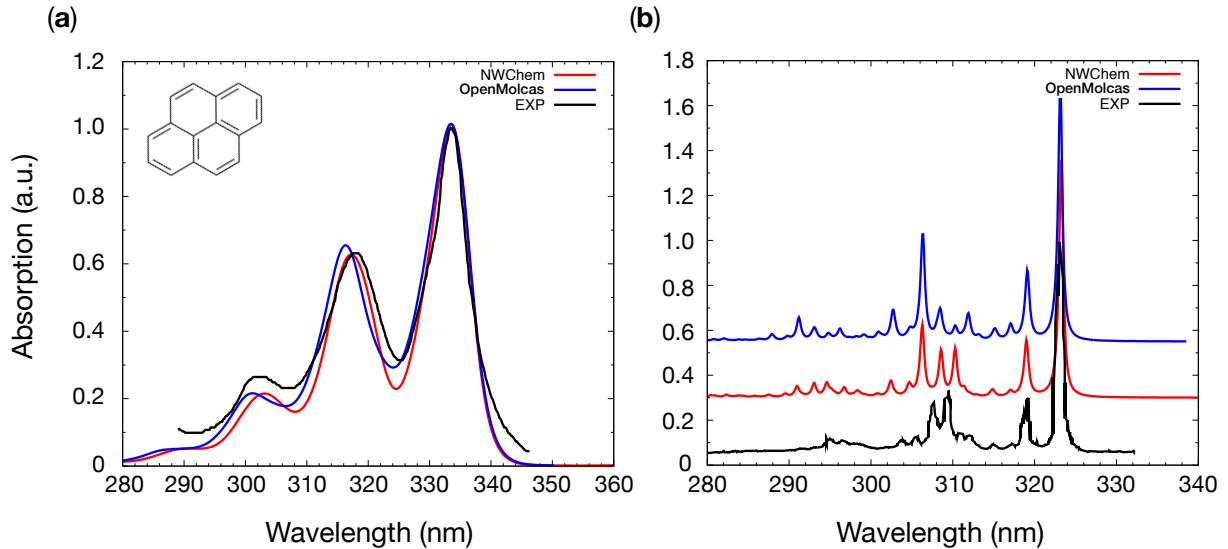


Figure 5: Comparison of the experimental (black line) and theoretical (TDDFT -red line- and RASSCF/RASPT2 -blue line-) linear absorption spectrum of pyrene recorded in (a) aqueous solution at room temperature[75] and in (b) a Ne-matrix at 5 K.[30] The theoretical and experimental spectra are normalized and centered at the fundamental (0–0) transition. Simulated spectra at 5 K have been vertically shifted to facilitate the comparison between them and with the experiment. The structure of the pyrene molecule is also reported (top left of panel (a)).

The first bright state of pyrene, dominated by a transition between the highest occupied orbital (HOMO) into the lowest unoccupied orbital (LUMO) and labeled *La* following Platt’s

⁹The solvent has little effect on the line shapes, as different solvents: water, ethanol, polycrystalline powder, acetonitrile and benzene provide similar Raman spectra[78] which more over are in excellent agreement with gas-phase simulations.[25, 56]

notation[67], shows a well resolved vibronic progression in the linear spectrum, with peaks at 334 nm, 318 nm and 302 nm with the fundamental accumulating the highest oscillator strength. Figure 5 documents the excellent agreement between theory and experiment for both aqueous solution room temperature data set[75], and Ne-matrix spectrum at 5 K[30]. The fine structure of the spectrum arises due to coupling to intra-molecular symmetry-conserving vibrational modes with frequencies $\sim 400 \text{ cm}^{-1}$ (ring breathing), $\sim 1100 \text{ cm}^{-1}$ and $\sim 1250 \text{ cm}^{-1}$ (in-plane bendings), $\sim 1450 \text{ cm}^{-1}$ and $\sim 1700 \text{ cm}^{-1}$ (carbon-carbon stretching, see Table 2 in the Supporting Information). Similar results have been reported by Freidzon et al. for LA simulations in water[25]. The 400 cm^{-1} mode, clearly seen at 5K (black spectrum in fig. 5, right) and accurately reproduced by both at RASSCF/RASPT2 and TDDFT levels of theory, is responsible for the shoulder in the fundamental of the LA band at room temperature. Regarding the high-frequency region, RASSCF/RASPT2 predicts the 1700 cm^{-1} stretching mode to dominate the line shape, whereas TDDFT predicts equal contributions from the 1250 cm^{-1} , 1450 cm^{-1} and 1700 cm^{-1} modes.

The excitation energy for the *La* state has been shifted by 2550 cm^{-1} and 280 cm^{-1} at room temperature, and by 3540 cm^{-1} and 1250 cm^{-1} at 5 K, for RASSCF/RASPT2 and TDDFT, respectively. Note that for both levels of theory, the difference between the applied shifts at room and cryogenic temperature is about 1000 cm^{-1} : this should therefore be ascribed to a solvatochromic shift induced by the aqueous environment of the 300 K experiment.

Nonlinear spectra

Figure 6 summarizes some of the possible spectra obtained form general 2D data computed by Spectron. In Figure 6 (a) we report the quasi-absorptive (i.e. summing the \mathbf{K}_I and \mathbf{K}_{II} signals 2DES heat maps at different t_2 times (0, 50, 100 and 150 fs). In particular, we note the dynamic multi-peak pattern of the (overlapping) GSB and SE contributions (shown in red) arising due to the vibronic structure of the absorption band. Furthermore, we can recognize ESA signatures of the *La* state, a weak one at 400 nm and an intense one at 610 nm. The 610 nm ESA has been reported experimentally.[69, 9]. It is established that the *La*

state decays through a non-adiabatic process mediated by a conical intersection to a lower lying dark excited state, denoted *Lb* in Platt's notation, on a 100 fs time scale. In the spectra this population transfer is reflected in: a) decrease of the intensity of the diagonal bleach signal due to the decay of the SE; b) gradual decay of the spectral fingerprints of the *La* state; c) the delayed build up of the *Lb* signatures at 370 nm and 480 nm, both of which have been reported experimentally.[69, 35, 11, 9] These observations are quantified by the time traces (reported in Figure 6 (b)) generated by iSPECTRON at wavelengths coinciding with the maximum of each spectral signature, highlighted with a green (SE), red (*La* ESA) and violet (*Lb* ESA) arrow, respectively, in the maps.

Figure 6 (c) shows two snapshots from the spectral dynamics of the GSB peak, highlighting in particular the elongated shape at early times (50 fs) and the circular shape at longer times (3 ps). This circularization of the 2D peaks is explained in terms of a memory erasure over time of the system initial excitation energy due to its interaction with the bath and of the finite life-time of system vibrations. This can be quantified by measuring the so called *flattening* parameter, which is the difference between the diagonal and the anti-diagonal peak widths, divided by the former. The magnitude of the flattening parameter is larger when the line shape significantly deviates from a circle, while it is zero when the two widths equate. The evolution of the flattening parameter is reported in Figure 6 (d), evidencing its progressive decay. The beating pattern is related to the long-living 400 cm^{-1} intra-molecular mode activated in the *La* state in the GSB process.

Broadband (300-650 nm) PP heat maps for the first 500 fs with both infinite resolution ("impulsive limit") and finite resolution (employing pulses with realistic duration of ~ 18 fs) are reported in Figure 6 (e), together with a Fourier transform of the residual ESA in the 550-650 nm window (Figure 6 (d)) after subtracting the decaying part of the signal (due to the $La \rightarrow Lb$ population transfer). The observed intensity beating, giving rise to the peaks in the Fourier transform spectra, is to be ascribed to the *La* ESA and as such it carries the information of the coherent wave-packet dynamics on the *La* potential energy surface initiated by the pump pulse. The effect of introducing finite time-resolution is to wash out the high frequency features in the spectral dynamics of the ESA thereby uncovering low frequency modes, accentuating the necessity of ultra short pulses (preferably with a sub-10

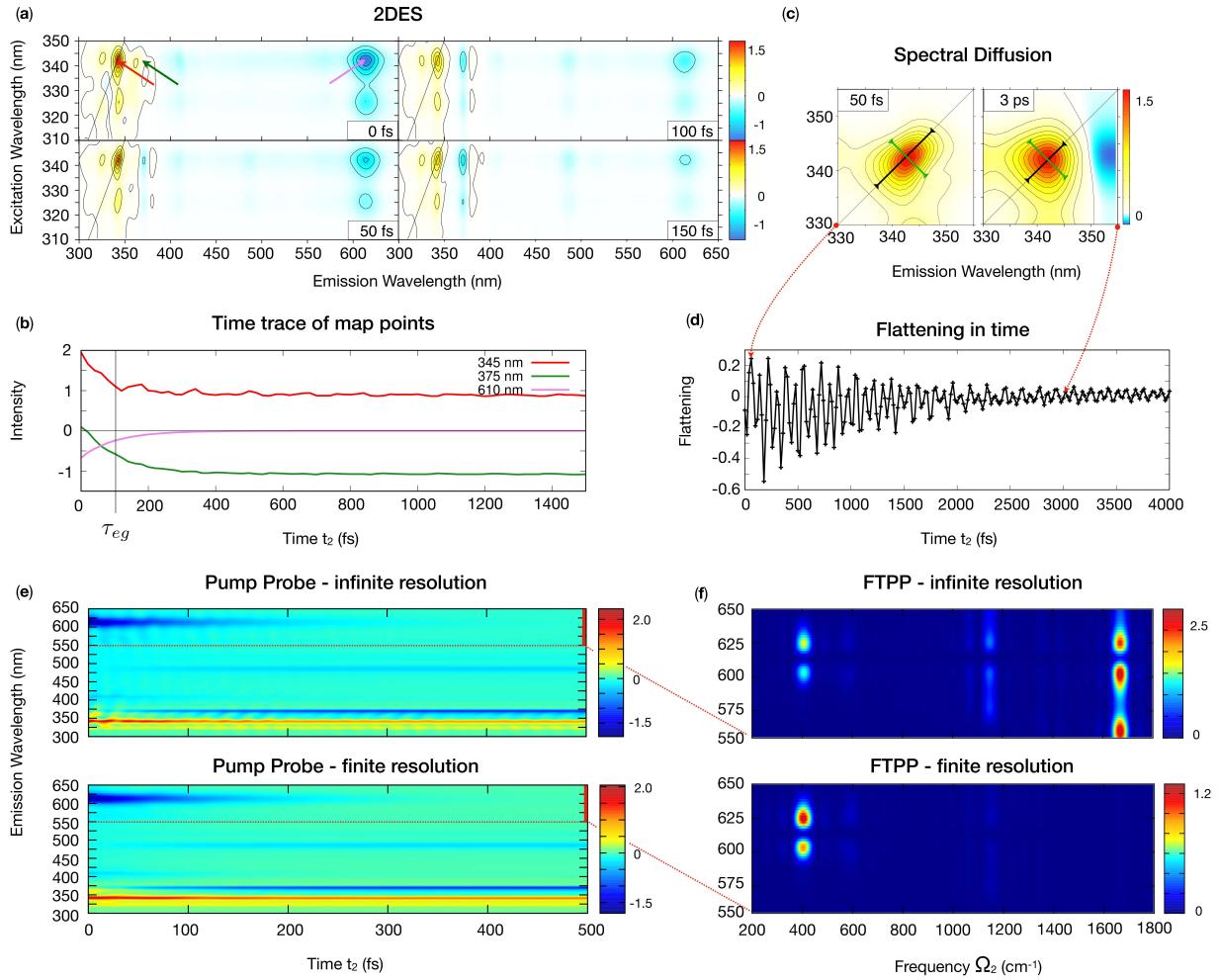


Figure 6: Various spectra and signal analysis obtained through iSPECTRON by processing 2DES simulation data produced by Spectron based on quantum chemistry computations performed at the CASPT2 level with OpenMolcas). (a) 2D maps at different t_2 times; positive peaks represent GSB and SE contributions, while negative peaks represent ESA contributions. The evolution of the various peaks due to the internal conversion between La and Lb states is made apparent by the time traces of selected points of the maps (indicated with arrows in (a)). (c) and (d) show the spectral diffusion and the evolution of the flattening parameter along t_2 , for the main positive peak in the maps. (e) and (f) report PP heat maps (simulated from 0 to 500 fs) and the Fourier transform (along t_2) of the region between 550-650 nm

fs duration) in order to thoroughly unravel the vibrational dynamics of photo-excited organic molecules[66].

SUMMARY AND FUTURE DEVELOPMENTS

The simulation of linear and nonlinear spectra is important to support the interpretation of experimental data and identify the microscopic origin of macroscopic spectral features. iSPECTRON was developed to facilitate a) the calculation of spectra via the Spectron code via the automatised extraction of QM properties form a variety of quantum chemistry software and b) the re-elaboration of these data to produce and analyze different spectra (1DPP, 2DES, etc) from the Spectron third-order response output.

We have demonstrated the capabilities of the developed code by simulating linear and nonlinear spectra for the pyrene molecule at different levels of theory, and with different QC software.

The future development of iSPECTRON aims at expanding the readout capabilities to more routines of the presently interfaced codes (i.e. GASSCF/GASPT2, DMRG, etc.) and towards the integration further QM codes. A second future development direction involves the calculation of the OBO parameters from first principles by integrating the calculation and evaluation of energy auto-correlation functions on top of MM simulations of the environmental DOF[77]. Finally, a third development direction aims at expanding the range of systems that can be treated through the interface. In this regard Spectron offers many more appealing features. One such features is the calculation of spectral densities from stochastic data (e.g. energy fluctuations along a given trajectory). Coupling molecular dynamics codes (such as COBRAMM for example) to this feature through iSPECTRON would allow one to describe vibrational dynamics beyond the DHO and CGF approximations. Another Spectron feature of interest, simulations in the exciton picture, would open the door to the spectroscopy simulation of systems which are currently beyond the capabilities of QM software.

APPENDIX

The first-order system response function (eq. 4) can be recast as

$$R^{(1)}(t_1) = \left(\frac{i}{\hbar}\right) \Theta(t_1) Tr \left[\hat{\boldsymbol{\mu}} \mathbb{G}(t_1) \left(\hat{\boldsymbol{\mu}} \hat{\rho}_0 \right) + \hat{\boldsymbol{\mu}} \mathbb{G}(t_1) \left(\hat{\rho}_0 \hat{\boldsymbol{\mu}} \right) \right] \quad (26)$$

Considering a generalized multi-level Hamiltonian with electronic states labeled a, b, c, d the superoperator $\mathbb{G}(t)$ assumes the form

$$\mathbb{G}(t_1) \bullet = \sum_{a,b} \exp\{-i(\omega_{ab} - i\gamma_{ab})\} |a\rangle \langle a| \bullet |b\rangle \langle b| \quad (27)$$

where the phenomenological dephasing rate γ_{ab} accounts for the decay of the coherence in an open quantum system. Inserting eq. 27 in the expression for the linear response (eq. 26) gives[53]

$$R^{(1)}(t_1) = \left(\frac{i}{\hbar}\right) \Theta(t_1) \sum_{a,b} |\boldsymbol{\mu}_{ab}|^2 \exp\{-i(\omega_{ab} - i\gamma_{ab}) t_1\} - c.c. \quad (28)$$

Similarly, the expression for the third-order system response function (eq. 6) can be recast as

$$\begin{aligned} R^{(3)}(t_1, t_2, t_3) = & \left(\frac{i}{\hbar}\right)^3 \Theta(t_1) \Theta(t_2) \Theta(t_3) Tr \left[\hat{\boldsymbol{\mu}} \mathbb{G}(t_3) \left(\hat{\boldsymbol{\mu}} \mathbb{G}(t_2) \left(\hat{\boldsymbol{\mu}} \mathbb{G}(t_1) \left(\hat{\boldsymbol{\mu}} \hat{\rho}_0 \right) \right) \right) \right. \\ & + \hat{\boldsymbol{\mu}} \mathbb{G}(t_3) \left(\hat{\boldsymbol{\mu}} \mathbb{G}(t_2) \left(\hat{\boldsymbol{\mu}} \mathbb{G}(t_1) \left(\hat{\rho}_0 \hat{\boldsymbol{\mu}} \right) \right) \right) + \hat{\boldsymbol{\mu}} \mathbb{G}(t_3) \left(\hat{\boldsymbol{\mu}} \mathbb{G}(t_2) \left(\mathbb{G}(t_1) \left(\hat{\boldsymbol{\mu}} \hat{\rho}_0 \right) \hat{\boldsymbol{\mu}} \right) \right) \\ & + \hat{\boldsymbol{\mu}} \mathbb{G}(t_3) \left(\hat{\boldsymbol{\mu}} \mathbb{G}(t_2) \left(\mathbb{G}(t_1) \left(\hat{\rho}_0 \hat{\boldsymbol{\mu}} \right) \hat{\boldsymbol{\mu}} \right) \right) + \hat{\boldsymbol{\mu}} \mathbb{G}(t_3) \left(\mathbb{G}(t_2) \left(\hat{\boldsymbol{\mu}} \mathbb{G}(t_1) \left(\hat{\boldsymbol{\mu}} \hat{\rho}_0 \right) \right) \hat{\boldsymbol{\mu}} \right) \quad (29) \\ & + \hat{\boldsymbol{\mu}} \mathbb{G}(t_3) \left(\mathbb{G}(t_2) \left(\hat{\boldsymbol{\mu}} \mathbb{G}(t_1) \left(\hat{\rho}_0 \hat{\boldsymbol{\mu}} \right) \right) \hat{\boldsymbol{\mu}} \right) + \hat{\boldsymbol{\mu}} \mathbb{G}(t_3) \left(\mathbb{G}(t_2) \left(\mathbb{G}(t_1) \left(\hat{\boldsymbol{\mu}} \hat{\rho}_0 \right) \hat{\boldsymbol{\mu}} \right) \hat{\boldsymbol{\mu}} \right) \\ & \left. + \hat{\boldsymbol{\mu}} \mathbb{G}(t_3) \left(\mathbb{G}(t_2) \left(\mathbb{G}(t_1) \left(\hat{\rho}_0 \hat{\boldsymbol{\mu}} \right) \hat{\boldsymbol{\mu}} \right) \hat{\boldsymbol{\mu}} \right) \right] \end{aligned}$$

Inserting the definition of the superoperator (eq. 27) in the expression of the nonlinear response (eq. 29) gives[53]

$$\begin{aligned}
R^{(3)}(t_1, t_2, t_3) = & \left(\frac{i}{\hbar} \right)^3 \Theta(t_1) \Theta(t_2) \Theta(t_3) \left[\right. \\
& \sum_{a,b,c,d} \boldsymbol{\mu}_{ab} \boldsymbol{\mu}_{bc} \boldsymbol{\mu}_{cd} \boldsymbol{\mu}_{da} \exp \{ -i (\omega_{dc} - i\gamma_{dc}) t_3 - i (\omega_{db} - i\gamma_{db}) t_2 - i (\omega_{da} - i\gamma_{da}) t_1 \} \\
& \sum_{a,b,c,d} \boldsymbol{\mu}_{ab} \boldsymbol{\mu}_{bc} \boldsymbol{\mu}_{cd} \boldsymbol{\mu}_{da} \exp \{ -i (\omega_{dc} - i\gamma_{dc}) t_3 - i (\omega_{db} - i\gamma_{db}) t_2 - i (\omega_{ab} - i\gamma_{ab}) t_1 \} \\
& \sum_{a,b,c,d} \boldsymbol{\mu}_{ab} \boldsymbol{\mu}_{bc} \boldsymbol{\mu}_{cd} \boldsymbol{\mu}_{da} \exp \{ -i (\omega_{dc} - i\gamma_{dc}) t_3 - i (\omega_{ac} - i\gamma_{ac}) t_2 - i (\omega_{ab} - i\gamma_{ab}) t_1 \} \\
& \sum_{a,b,c,d} \boldsymbol{\mu}_{ab} \boldsymbol{\mu}_{bc} \boldsymbol{\mu}_{cd} \boldsymbol{\mu}_{da} \exp \{ -i (\omega_{ba} - i\gamma_{ba}) t_3 - i (\omega_{ca} - i\gamma_{ca}) t_2 - i (\omega_{da} - i\gamma_{da}) t_1 \} - c.c. \left. \right] \\
& \quad (30)
\end{aligned}$$

Eqs. 28 and 30 are general and apply to any multi-level quantum system.

ACKNOWLEDGEMENTS

F.S., A.N., D.R.N., N.G., S.M, M.G. acknowledge support from the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division under Award Nos. DE-SC0019484, KC-030103172684. The Spectron code was developed with support from the National Science Foundation (Grant CHE-1953045). This research benefited from computational resources provided by EMSL, a DOE Office of Science User Facility sponsored by the Office of Biological and Environmental Research and located at PNNL. PNNL is operated by Battelle Memorial Institute for the United States Department of Energy under DOE Contract No. DE-AC05-76RL1830.

References

1. D. Abramavicius, B. Palmieri, D. V. Voronine, F. Šanda, and S. Mukamel. Coherent multidimensional optical spectroscopy of excitons in molecular aggregates; quasiparticle

- versus supermolecule perspectives. *Chemical Reviews*, 109(6):2350–2408, 2009. PMID: 19432416.
2. D. Abramavicius, L. Valkunas, and S. Mukamel. Transport and correlated fluctuations in the nonlinear optical response of excitons. *Europhysics Letters (EPL)*, 80(1):17005, Sept. 2007.
 3. K. Andersson, P. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski. Second-order perturbation theory with a casscf reference function. *J. Phys. Chem.*, 94(14):5483–5488, 1990.
 4. E. Aprà, E. J. Bylaska, W. A. De Jong, N. Govind, K. Kowalski, T. P. Straatsma, M. Valiev, H. van Dam, Y. Alexeev, J. Anchell, et al. Nwchem: Past, present, and future. *The Journal of chemical physics*, 152(18):184102, 2020.
 5. F. Aquilante, J. Autschbach, A. Baiardi, S. Battaglia, V. A. Borin, L. F. Chibotaru, I. Conti, L. D. Vico, M. Delcey, I. F. Galván, N. Ferré, L. Freitag, M. Garavelli, X. Gong, S. Knecht, E. D. Larsson, R. Lindh, M. Lundberg, P. Å. Malmqvist, A. Nenov, J. Norell, M. Odelius, M. Olivucci, T. B. Pedersen, L. Pedraza-González, Q. M. Phung, K. Pierloot, M. Reiher, I. Schapiro, J. Segarra-Martí, F. Segatta, L. Seijo, S. Sen, D.-C. Sergentu, C. J. Stein, L. Ungur, M. Vacher, A. Valentini, and V. Veryazov. Modern quantum chemistry with [open]molcas. *The Journal of Chemical Physics*, 152(21):214117, June 2020.
 6. A. Baiardi and M. Reiher. The density matrix renormalization group in chemistry and molecular physics: Recent developments and new challenges. *The Journal of Chemical Physics*, 152(4):040903, Jan. 2020.
 7. G. M. J. Barca, C. Bertoni, L. Carrington, D. Datta, N. De Silva, J. E. Deustua, D. G. Fedorov, J. R. Gour, A. O. Gunina, E. Guidez, T. Harville, S. Irle, J. Ivanic, K. Kowalski, S. S. Leang, H. Li, W. Li, J. J. Lutz, I. Magoulas, J. Mato, V. Mironov, H. Nakata, B. Q. Pham, P. Piecuch, D. Poole, S. R. Pruitt, A. P. Rendell, L. B. Roskop, K. Ruedenberg, T. Sattasathuchana, M. W. Schmidt, J. Shen, L. Slipchenko, M. Sosonkina, V. Sundriyal,

- A. Tiwari, J. L. Galvez Vallejo, B. Westheimer, M. Wloch, P. Xu, F. Zahariev, and M. S. Gordon. Recent developments in the general atomic and molecular electronic structure system. *The Journal of Chemical Physics*, 152(15):154102, Apr. 2020.
8. M. Bondanza, M. Nottoli, L. Cupellini, F. Lipparini, and B. Mennucci. Polarizable embedding QM/MM: the future gold standard for complex (bio)systems? *Physical Chemistry Chemical Physics*, 22(26):14433–14448, 2020.
9. R. Borrego-Varillas, L. Ganzer, G. Cerullo, and C. Manzoni. Ultraviolet transient absorption spectrometer with sub-20-fs time resolution. *Applied Sciences*, 8(6):989, June 2018.
10. R. Borrego-Varillas, A. Nenov, L. Ganzer, A. Oriana, C. Manzoni, A. Tolomelli, I. Rivolta, S. Mukamel, M. Garavelli, and G. Cerullo. Two-dimensional UV spectroscopy: a new insight into the structure and dynamics of biomolecules. *Chemical Science*, 10(43):9907–9921, 2019.
11. R. Borrego-Varillas, A. Oriana, L. Ganzer, A. Trifonov, I. Buchvarov, C. Manzoni, and G. Cerullo. Two-dimensional electronic spectroscopy in the ultraviolet by a birefringent delay line. *Optics Express*, 24(25):28491, Nov. 2016.
12. R. Borrego-Varillas, D. C. Teles-Ferreira, A. Nenov, I. Conti, L. Ganzer, C. Manzoni, M. Garavelli, A. M. de Paula, and G. Cerullo. Observation of the sub-100 femtosecond population of a dark state in a thiobase mediating intersystem crossing. *Journal of the American Chemical Society*, 140(47):16087–16093, Oct. 2018.
13. W. B. Bosma, Y. J. Yan, and S. Mukamel. Impulsive pump-probe and photon-echo spectroscopies of dye molecules in condensed phases. *Physical Review A*, 42(11):6920–6923, Dec. 1990.
14. M. E. Casida and D. Chong. Recent advances in density functional methods. *Computational Chemistry: Reviews of Current Trends*, 1995.
15. M. E. Casida and M. Huix-Rotllant. Progress in time-dependent density-functional theory. *Annual review of physical chemistry*, 63:287–323, 2012.

16. G. K.-L. Chan and S. Sharma. The density matrix renormalization group in quantum chemistry. *Annual Review of Physical Chemistry*, 62(1):465–481, May 2011.
17. P. C. Chen. An introduction to coherent multidimensional spectroscopy. *Applied Spectroscopy*, 70(12):1937–1951, Dec. 2016.
18. Y.-K. Choe, Y. Nakao, and K. Hirao. Multireference møller–plesset method with a complete active space configuration interaction reference function. *The Journal of Chemical Physics*, 115(2):621–629, July 2001.
19. L. W. Chung, W. M. C. Sameera, R. Ramozzi, A. J. Page, M. Hatanaka, G. P. Petrova, T. V. Harris, X. Li, Z. Ke, F. Liu, H.-B. Li, L. Ding, and K. Morokuma. The ONIOM method and its applications. *Chemical Reviews*, 115(12):5678–5796, Apr. 2015.
20. A. J. Cohen, P. Mori-Sánchez, and W. Yang. Challenges for density functional theory. *Chemical reviews*, 112(1):289–320, 2012.
21. R. M. Dreizler and E. K. Gross. *Density functional theory: an approach to the quantum many-body problem*. Springer Science & Business Media, 2012.
22. P. Elliott, S. Goldson, C. Canahui, and N. T. Maitra. Perspectives on double-excitations in tddft. *Chemical Physics*, 391(1):110–119, 2011.
23. F. J. A. Ferrer, J. Cerezo, J. Soto, R. Improta, and F. Santoro. First-principle computation of absorption and fluorescence spectra in solution accounting for vibronic structure, temperature effects and solvent inhomogenous broadening. *Computational and Theoretical Chemistry*, 1040-1041:328–337, July 2014.
24. J. Fosso-Tande, T.-S. Nguyen, G. Gidofalvi, and A. E. DePrince. Large-scale variational two-electron reduced-density-matrix-driven complete active space self-consistent field methods. *Journal of Chemical Theory and Computation*, 12(5):2260–2271, Apr. 2016.
25. A. Y. Freidzon, R. R. Valiev, and A. A. Berezhnoy. Ab initio simulation of pyrene spectra in water matrices. *RSC Adv.*, 4(79):42054–42065, 2014.

26. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakerzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox. Gaussian[®]16 Revision C.01, 2016. Gaussian Inc. Wallingford CT.
27. I. F. Galván, M. Vacher, A. Alavi, C. Angeli, F. Aquilante, J. Autschbach, J. J. Bao, S. I. Bokarev, N. A. Bogdanov, R. K. Carlson, L. F. Chibotaru, J. Creutzberg, N. Dattani, M. G. Delcey, S. S. Dong, A. Dreuw, L. Freitag, L. M. Frutos, L. Gagliardi, F. Gendron, A. Giussani, L. González, G. Grell, M. Guo, C. E. Hoyer, M. Johansson, S. Keller, S. Knecht, G. Kovačević, E. Källman, G. L. Manni, M. Lundberg, Y. Ma, S. Mai, J. P. Malhado, P. Å. Malmqvist, P. Marquetand, S. A. Mewes, J. Norell, M. Olivucci, M. Oppel, Q. M. Phung, K. Pierloot, F. Plasser, M. Reiher, A. M. Sand, I. Schapiro, P. Sharma, C. J. Stein, L. K. Sørensen, D. G. Truhlar, M. Ugandi, L. Ungur, A. Valentini, S. Vancoillie, V. Veryazov, O. Weser, T. A. Wesołowski, P.-O. Widmark, S. Wouters, A. Zech, J. P. Zobel, and R. Lindh. OpenMolcas: From source code to insight. *Journal of Chemical Theory and Computation*, 15(11):5925–5964, Sept. 2019.
28. A. Gelzinis, R. Augulis, V. Butkus, B. Robert, and L. Valkunas. Two-dimensional spectroscopy for non-specialists. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1860(4):271–285, Apr. 2019.
29. N. Govind, K. Lopata, R. Rousseau, A. Andersen, and K. Kowalski. Visible light absorption of n-doped TiO_2 rutile using (lr/rt)-tddft and active space eomccsd calculations.

The Journal of Physical Chemistry Letters, 2(21):2696–2701, 2011.

30. T. M. Halasinski, F. Salama, and L. J. Allamandola. Investigation of the ultraviolet, visible, and near-infrared absorption spectra of hydrogenated polycyclic aromatic hydrocarbons and their cations. *The Astrophysical Journal*, 628(1):555–566, July 2005.
31. P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Physical review*, 136(3B):B864, 1964.
32. J. Ivanic. Direct configuration interaction and multiconfigurational self-consistent-field method for multiple active spaces with variable occupations. i. method. *The Journal of Chemical Physics*, 119(18):9364–9376, Nov. 2003.
33. S. Jurinovich, L. Cupellini, C. A. Guido, and B. Mennucci. EXAT: EXcitonic analysis tool. *Journal of Computational Chemistry*, 39(5):279–286, Nov. 2017.
34. W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Physical review*, 140(4A):A1133, 1965.
35. N. Krebs, I. Pugliesi, J. Hauer, and E. Riedle. Two-dimensional fourier transform spectroscopy in the ultraviolet with sub-20 fs pump pulses and 250–720 nm supercontinuum probe. *New Journal of Physics*, 15(8):085016, Aug. 2013.
36. S. Kümmel and L. Kronik. Orbital-dependent density functionals: Theory and applications. *Reviews of Modern Physics*, 80(1):3, 2008.
37. A. D. Laurent and D. Jacquemin. Td-dft benchmarks: a review. *International Journal of Quantum Chemistry*, 113(17):2019–2039, 2013.
38. M. K. Lee, P. Huo, and D. F. Coker. Semiclassical path integral dynamics: Photosynthetic energy transfer with realistic environment interactions. *Annual Review of Physical Chemistry*, 67(1):639–668, May 2016.
39. B. G. Levine, C. Ko, J. Quenneville, and T. J. Martínez. Conical intersections and double excitations in time-dependent density functional theory. *Molecular Physics*, 104(5–7):1039–1051, 2006.

40. B. Li, A. E. Johnson, S. Mukamel, and A. B. Myers. The brownian oscillator model for solvation effects in spontaneous light emission and their relationship to electron transfer. *Journal of the American Chemical Society*, 116(24):11039–11047, Nov. 1994.
41. F. Lipparini and B. Mennucci. Perspective: Polarizable continuum models for quantum-mechanical descriptions. *The Journal of Chemical Physics*, 144(16):160901, Apr. 2016.
42. D. Loco, L. Lagardère, G. A. Cisneros, G. Scalmani, M. Frisch, F. Lipparini, B. Mennucci, and J.-P. Piquemal. Towards large scale hybrid QM/MM dynamics of complex systems with advanced point dipole polarizable embeddings. *Chemical Science*, 10(30):7200–7211, 2019.
43. K. Lopata and N. Govind. Modeling fast electron dynamics with real-time time-dependent density functional theory: application to small molecules and chromophores. *Journal of chemical theory and computation*, 7(5):1344–1355, 2011.
44. K. Lopata, B. E. Van Kuiken, M. Khalil, and N. Govind. Linear-response and real-time time-dependent density functional theory studies of core-level near-edge x-ray absorption. *Journal of chemical theory and computation*, 8(9):3284–3292, 2012.
45. M. Lundberg and M. G. Delcey. Multiconfigurational approach to x-ray spectroscopy of transition metal complexes. In *Transition Metals in Coordination Environments*, pages 185–217. Springer International Publishing, 2019.
46. D. Ma, G. L. Manni, and L. Gagliardi. The generalized active space concept in multiconfigurational self-consistent field methods. *The Journal of Chemical Physics*, 135(4):044128, July 2011.
47. D. Ma, G. L. Manni, J. Olsen, and L. Gagliardi. Second-order perturbation theory for generalized active space self-consistent-field wave functions. *Journal of Chemical Theory and Computation*, 12(7):3208–3213, June 2016.
48. N. T. Maitra. Perspective: Fundamental aspects of time-dependent density functional theory. *The Journal of Chemical Physics*, 144(22):220901, 2016.

49. P. A. Malmqvist, A. Rendell, and B. O. Roos. The restricted active space self-consistent-field method, implemented with a split graph unitary group approach. *The Journal of Physical Chemistry*, 94(14):5477–5482, July 1990.
50. G. L. Manni, S. D. Smart, and A. Alavi. Combining the complete active space self-consistent field method and the full configuration interaction quantum monte carlo within a super-CI framework, with application to challenging metal-porphyrins. *Journal of Chemical Theory and Computation*, 12(3):1245–1258, Feb. 2016.
51. M. A. Marques, N. T. Maitra, F. M. Nogueira, E. K. Gross, and A. Rubio. *Fundamentals of time-dependent density functional theory*, volume 837. Springer Science & Business Media, 2012.
52. H. J. B. Marroux, A. P. Fidler, D. M. Neumark, and S. R. Leone. Multidimensional spectroscopy with attosecond extreme ultraviolet and shaped near-infrared pulses. *Science Advances*, 4(9):eaau3783, Sept. 2018.
53. S. Mukamel. *Principles of nonlinear optical spectroscopy*, volume 523. Oxford Univ. Press, 1995.
54. I. H. Nayyar, S. E. Chamberlin, T. C. Kaspar, N. Govind, S. A. Chambers, and P. V. Sushko. Effect of doping and chemical ordering on the optoelectronic properties of complex oxides: Fe₂O₃–V₂O₃ solid solutions and hetero-structures. *Physical Chemistry Chemical Physics*, 19(2):1097–1107, 2017.
55. A. Nenov, R. Borrego-Varillas, A. Oriana, L. Ganzer, F. Segatta, I. Conti, J. Segarra-Marti, J. Omachi, M. Dapor, S. Taioli, C. Manzoni, S. Mukamel, G. Cerullo, and M. Garavelli. UV-light-induced vibrational coherences: The key to understand kasha rule violation in trans-azobenzene. *The Journal of Physical Chemistry Letters*, 9(7):1534–1541, Mar. 2018.
56. A. Nenov, A. Giussani, B. P. Fingerhut, I. Rivalta, E. Dumont, S. Mukamel, and M. Garavelli. Spectral lineshapes in nonlinear electronic spectroscopy. *Physical Chemistry Chemical Physics*, 17(46):30925–30936, 2015.

57. A. Nenov, S. Mukamel, M. Garavelli, and I. Rivalta. Two-dimensional electronic spectroscopy of benzene, phenol, and their dimer: An efficient first-principles simulation protocol. *Journal of Chemical Theory and Computation*, 11(8):3755–3771, July 2015.
58. A. Nenov, I. Rivalta, G. Cerullo, S. Mukamel, and M. Garavelli. Disentangling peptide configurations via two-dimensional electronic spectroscopy: Ab initio simulations beyond the frenkel exciton hamiltonian. *The Journal of Physical Chemistry Letters*, 5(4):767–771, Feb. 2014.
59. A. Nenov, I. Rivalta, S. Mukamel, and M. Garavelli. Bidimensional electronic spectroscopy on indole in gas phase and in water from first principles. *Computational and Theoretical Chemistry*, 1040-1041:295–303, July 2014.
60. T. A. A. Oliver. Recent advances in multidimensional ultrafast spectroscopy. *Royal Society Open Science*, 5(1):171425, Jan. 2018.
61. J. Olsen, B. O. Roos, P. Jo/rsgensen, and H. J. A. Jensen. Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces. *The Journal of Chemical Physics*, 89(4):2185–2192, Aug. 1988.
62. G. Onida, L. Reining, and A. Rubio. Electronic excitations: density-functional versus many-body green’s-function approaches. *Reviews of modern physics*, 74(2):601, 2002.
63. S. Pandey, R. Bean, T. Sato, I. Poudyal, J. Bielecki, J. C. Villarreal, O. Yefanov, V. Mariani, T. A. White, C. Kupitz, M. Hunter, M. H. Abdellatif, S. Bajt, V. Bondar, A. Echelmeier, D. Doppler, M. Emons, M. Frank, R. Fromme, Y. Gevorkov, G. Giovanetti, M. Jiang, D. Kim, Y. Kim, H. Kirkwood, A. Klimovskaia, J. Knoska, F. H. M. Koua, R. Letrun, S. Lisova, L. Maia, V. Mazalova, D. Meza, T. Michelat, A. Ourmazd, G. Palmer, M. Ramilli, R. Schubert, P. Schwander, A. Silenzi, J. Sztuk-Dambietz, A. Tolstikova, H. N. Chapman, A. Ros, A. Barty, P. Fromme, A. P. Mancuso, and M. Schmidt. Time-resolved serial femtosecond crystallography at the european XFEL. *Nature Methods*, 17(1):73–78, Nov. 2019.

64. R. Parr and W. Yang. *Density-functional theory of atoms and molecules*. Oxford University Press, 1989.
65. J. P. Perdew and K. Schmidt. Jacob's ladder of density functional approximations for the exchange-correlation energy. In *AIP Conference Proceedings*, volume 577, pages 1–20. American Institute of Physics, 2001.
66. A. Picchiotti, A. Nenov, A. Giussani, V. I. Prokhorenko, R. J. D. Miller, S. Mukamel, and M. Garavelli. Pyrene, a test case for deep-ultraviolet molecular photophysics. *The Journal of Physical Chemistry Letters*, 10(12):3481–3487, May 2019.
67. J. R. Platt. Classification of spectra of cata-condensed hydrocarbons. *The Journal of Chemical Physics*, 17(5):484–495, May 1949.
68. A. Pribam-Jones, D. A. Gross, and K. Burke. Dft: A theory full of holes? *Annual review of physical chemistry*, 66:283–304, 2015.
69. M. Raytchev, E. Pandurski, I. Buchvarov, C. Modrakowski, and T. Fiebig. Bichromophoric interactions and time-dependent excited state mixing in pyrene derivatives. a femtosecond broad-band pump-probe study. *The Journal of Physical Chemistry A*, 107(23):4592–4600, June 2003.
70. D. Roca-Sanjuán, F. Aquilante, and R. Lindh. Multiconfiguration second-order perturbation theory approach to strong electron correlation in chemistry and photochemistry. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(4):585–603, Nov. 2011.
71. B. O. Roos. *Ab Initio Methods in Quantum Chemistry: Part II*. Wiley, Chichester, UK, 1987.
72. B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn. *Chem. Phys.*, 48:157, 1980.
73. E. Runge and E. K. Gross. Density-functional theory for time-dependent systems. *Physical Review Letters*, 52(12):997, 1984.

74. F. Santoro, R. Improta, A. Lami, J. Bloino, and V. Barone. Effective method to compute franck-condon integrals for optical spectra of large molecules in solution. *The Journal of Chemical Physics*, 126(8):084509, Feb. 2007.
75. F. P. Schwarz and S. P. Wasik. Fluorescence measurements of benzene, naphthalene, anthracene, pyrene, fluoranthene, and benzo[e]pyrene in water. *Analytical Chemistry*, 48(3):524–528, Mar. 1976.
76. J. Segarra-Martí, S. Mukamel, M. Garavelli, A. Nenov, and I. Rivalta. Towards accurate simulation of two-dimensional electronic spectroscopy. In *Topics in Current Chemistry Collections*, pages 63–112. Springer International Publishing, June 2018.
77. J. Segarra-Martí, F. Segatta, T. A. Mackenzie, A. Nenov, I. Rivalta, M. J. Bearpark, and M. Garavelli. Modeling multidimensional spectral lineshapes from first principles: application to water-solvated adenine. *Faraday Discussions*, 221:219–244, 2020.
78. H. Shinohara, Y. Yamakita, and K. Ohno. Raman spectra of polycyclic aromatic hydrocarbons. comparison of calculated raman intensity distributions with observed spectra for naphthalene, anthracene, pyrene, and perylene. *Journal of Molecular Structure*, 442(1-3):221–234, Feb. 1998.
79. C. L. Smallwood and S. T. Cundiff. Multidimensional coherent spectroscopy of semiconductors. *Laser & Photonics Reviews*, 12(12):1800171, Nov. 2018.
80. J. E. T. Smith, B. Mussard, A. A. Holmes, and S. Sharma. Cheap and near exact CASSCF with large active spaces. *Journal of Chemical Theory and Computation*, 13(11):5468–5478, Nov. 2017.
81. Y. Song, A. Konar, R. Sechrist, V. P. Roy, R. Duan, J. Dziurgot, V. Policht, Y. A. Matutes, K. J. Kubarych, and J. P. Ogilvie. Multispectral multidimensional spectrometer spanning the ultraviolet to the mid-infrared. *Review of Scientific Instruments*, 90(1):013108, Jan. 2019.
82. A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, Inc., New York, 1996.

83. C. A. Ullrich. *Time-dependent density-functional theory: concepts and applications*. OUP Oxford, 2011.
84. M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. Van Dam, D. Wang, J. Nieplocha, E. Aprà, T. L. Windus, et al. Nwchem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Computer Physics Communications*, 181(9):1477–1489, 2010.
85. Y. Wang, K. Lopata, S. A. Chambers, N. Govind, and P. V. Sushko. Optical absorption and band gap reduction in (fe_{1-x} cr _x)₂o₃ solid solutions: A first-principles study. *The Journal of Physical Chemistry C*, 117(48):25504–25512, 2013.
86. O. Weingart, A. Nenov, P. Altoè, I. Rivalta, J. Segarra-Martí, I. Dokukina, and M. Garavelli. COBRAMM 2.0 — a software interface for tailoring molecular electronic structure calculations and running nanoscale (QM/MM) simulations. *Journal of Molecular Modeling*, 24(9), Sept. 2018.
87. A. Wituschek, L. Bruder, E. Allaria, U. Bangert, M. Binz, R. Borghes, C. Callegari, G. Cerullo, P. Cinquegrana, L. Giannessi, M. Danailov, A. Demidovich, M. D. Fraia, M. Drabbels, R. Feifel, T. Laarmann, R. Michiels, N. S. Mirian, M. Mudrich, I. Nikolov, F. H. O’Shea, G. Penco, P. Piseri, O. Plekan, K. C. Prince, A. Przystawik, P. R. Ribič, G. Sansone, P. Sigalotti, S. Spampinati, C. Spezzani, R. J. Squibb, S. Stranges, D. Uhl, and F. Stienkemeier. Tracking attosecond electronic coherences using phase-manipulated extreme ultraviolet pulses. *Nature Communications*, 11(1), Feb. 2020.
88. A. H. Zewail and J. M. Thomas. *4D Electron Microscopy*. IMPERIAL COLLEGE PRESS, Dec. 2009.
89. Y. Zhang, J. D. Biggs, D. Healion, N. Govind, and S. Mukamel. Core and valence excitations in resonant x-ray spectroscopy using restricted excitation window time-dependent density functional theory. *The Journal of chemical physics*, 137(19):194306, 2012.

90. W. Zhuang, T. Hayashi, and S. Mukamel. Coherent multidimensional vibrational spectroscopy of biomolecules: Concepts, simulations, and challenges. *Angewandte Chemie International Edition*, 48(21):3750–3781, May 2009.