Article

Time-Resolved X-ray Absorption Spectroscopy: An MCTDH Quantum Dynamics Protocol

Francesco Segatta, Daniel Aranda, Flavia Aleotti, Francesco Montorsi, Shaul Mukamel, Marco Garavelli,* Fabrizio Santoro,* and Artur Nenov*

Cite This: *J. Chem. Theory Comput.* 2024, 20, 307−322

Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: Expressions for linear and nonlinear spectroscopy simulation in the X-ray window in which the time evolution of a photoexcited molecular system is treated via quantum dynamics are derived. By leveraging on the peculiar properties of core-excited/ionized states, first- and third-order response functions are recast in the limit of time-scale separation between the extremely short core-state lifetime and the (comparably longer) electronic-state transfer and nuclear vibrational motion. This work is a natural extension of Segatta et al. (*J. Chem. Theory Comput.* 2023, 19, 2075−2091), in which some of the present authors coupled MCTDH quantum dynamics to spectroscopy simulation at different levels of sophistication. Full quantum dynamics and approximate expressions are compared by simulating X-ray transient absorption spectroscopy at the carbon K-edge in the pyrene molecule.
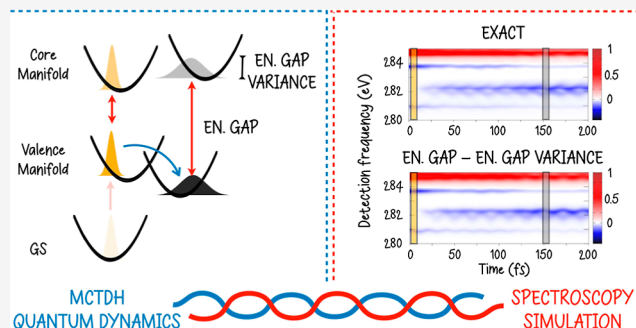
1. INTRODUCTION

Monitoring photoinduced events in real time is the primary goal of time-resolved (TR-) spectroscopy. Transient absorption (TA) is one of the most widespread techniques: an ultraviolet (UV)/visible (vis) *pump* pulse excites the system into the manifold of valence-excited states where the photophysics/photochemistry of interest occurs. After a controlled time delay, a *probe* pulse is sent to the sample to monitor the evolution of the molecular system that took place during the delay time. Techniques interrogating the system's temporal evolution with infrared (IR) or UV/vis light are routinely used to resolve events taking place on the sub-ps time scale,[2−4] demonstrating that a wealth of complementary information can be obtained by targeting vibrational levels (IR) and valence-excited and -ionized states (UV/vis). Recently, new sources of light, including X-rays in free-electron lasers (XFEL)[5,6] large facilities and table-top high harmonic generation (HHG)[7,8] setups, have facilitated access to a new spectral probing window that covers core-excited and -ionized states.[9−15] By employing X-rays as a probe, promotion of (atom-centered) core electrons into empty molecular orbitals (core excitations) or in the continuum (core ionization) reveals information about the photoinduced dynamics of the system, which is inaccessible by probing other spectral regimes. Such information is *local*, atom-dependent, and sensitive to the chemical environment and/or oxidation state of the core-excited atom.

The simulation of TA signals requires computation of the nonlinear (third-order) response of the system perturbed by the pump and probe laser pulses. The pump pulse creates an out-of-equilibrium wave packet (WP) in the manifold of valence states, whereas the probe resonantly couples the evolved WP to a manifold of higher vibrational and/or electronic excited states. Regardless of the energy window that different types of probe may employ to monitor the photoinduced process of interest, it is understood that the quality of the simulated TA spectra will be strongly affected by the quality of the WP evolution. While cost-effective semiclassical[16] or mixed quantum-classical[17,18] approaches have been developed in the last few decades to replace the WP propagation with a swarm of trajectories, the highest level of accuracy is obtained by means of quantum dynamics methods, in which the quantum nature of both electrons and nuclei is preserved. Interestingly, many approximate approaches are not capable of capturing subtle quantum effects such as the creation of electronic coherences along the dynamics, as, e.g., when the WP passes through a conical intersection. Novel spectroscopic techniques in the X-ray
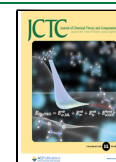
domain (such as TRUECARS and ultrafast X-ray diffraction)[19−21] that leverage on the detection of such coherences to pinpoint passage through a conical intersection require these to be carefully described, making it necessary to employ quantum dynamics WP propagation schemes. Recently, we have developed a simulation protocol for computing the third-order response from multidimensional fully quantum dynamical simulations based on the multiconfigurational time-dependent Hartree (MCTDH) method.[1] In the present contribution, the approach of ref 1 is further developed to conjugate the accuracy of quantum dynamics simulations with an efficient TA simulation protocol in the X-ray domain.

MCTDH[22−24] is a numerical protocol for solving the time-dependent Schröedinger equation which is particularly effective when the potential energy surfaces (PESs) on which the nuclei evolve can be approximated by a low-order Taylor expansion in normal mode coordinates. This is often possible in the diabatic picture, in which the molecular system is represented in a basis of states with well-defined (electronic) character. Hamiltonians that describe such simplified PESs are called model vibronic coupling Hamiltonians.[25,26] The simplest form of the Hamiltonian, known as the linear vibronic coupling (LVC) model, assumes that all PESs share the same normal modes and frequencies and that off-diagonal terms, emerging in the diabatic representation, are linear functions of the coordinates. This facilitates QD simulations with many tens of nuclear degrees of freedom. Quadratic vibronic coupling (QVC) models are also efficiently treated at the MCTDH level, also taking into account different curvatures for the different electronic states, and even Duschinsky mixing between normal modes.[27,28] MCTDH can treat more general PES functional forms, inevitably requiring reduction of the dimensionality of the problem due to the increased computational cost.

Despite its simplicity, the LVC Hamiltonian naturally captures passage through conical intersections, which is known to facilitate ultrafast internal conversion, as well as charge and energy transfer.[29] In our implementation, the LVC Hamiltonian is parametrized with a maximum-overlap diabatization protocol,[30,31] using multiconfigurational wave-function-based electronic structure methods such as the complete and restricted active space self-consistent field theory corrected by second-order perturbation theory, i.e., the CASSCF/CASPT2 and RASSCF/RASPT2 protocols.[32,33] The response function computation requires evaluation of the overlap between nuclear WP (WPO) evolving on different electronic surfaces during the pump–probe delay time, t_2, as well as during the time that elapses between the probe–matter interaction and the signal detection, termed t_3. This LVC/WPO protocol can be applied to simulate TA signals in the entire spectral window from the visible to the X-ray.

The short-lived nature of core-excited states compared to the time scales of electronic population dynamics and nuclear vibrational periods justifies treating the WP propagation following the interaction with the probe pulse in approximate ways. The short-time approximation (STA) invoked in ref 34 computes static absorption signals on top of a WP dynamics in the valence manifold, thereby completely neglecting the WP evolution projected by the probe pulse in the manifold of higher-lying states. Its validity for X-ray probe pulses has been studied by Freibert et al.[35] who compared LVC/STA results with the highest LVC/WPO approach. They report that STA accurately reproduces the positions of the transient signals and

their broadening, but as expected it lacks a fine vibronic structure, which becomes more visible the longer the lifetime of the core-excited states.

The exact STA response at each delay time t_2 can be obtained either by computing the nuclear probability density in coordinate space[36,37] at a cost which increases with the number of degrees of freedom, or, alternatively, in the time domain by explicitly propagating the WP during a time interval t_3, while omitting the kinetic operator, at a cost comparable to that of the WPO method.[35] An approximation often made is to utilize only the vertical energy gap at the centroid, i.e., at the expectation value of the nuclear position of the WP evolving in the manifold of valence states. While this approximation minimizes the overhead to the t_2 WP propagation in the valence manifold, peak broadening and/or asymmetries due to the form of the nuclear probability density are fully neglected. Instead, the stick spectra obtained are dressed with phenomenological broadening, set identical for all valence–core transitions and constant in time.[38] In the following, we will refer to this level of approximation as the (vertical) energy gap (at centroid) approximation (EGA).

In this contribution, we develop a protocol which improves the EGA by incorporating a time- and transition-dependent analytical signal broadening due to WP *width* in the valence states (accurately captured by the quantum dynamics evolution of the WP) by introducing an energy gap (pseudo)variance expression whose exact form is rigorously obtained by Taylor expanding the WP propagator during the t_3 time interval. We refer to this protocol as the energy gap/(pseudo)variance (at centroid) approximation (EGVA). The performance of this approach is assessed by comparing X-ray TA spectra obtained with the LVC/EGVA and LVC/WPO protocols at the carbon K-edge of pyrene, following the ultrafast (sub-100 fs) $S_2 \rightarrow S_1$ internal conversion (IC) in UV-excited pyrene.[31,39−43]

2. METHODS

2.1. Quantum Chemistry of Core Excitations. The required information for the simulation of X-ray pump–probe spectra within the LVC model includes state energies and gradients, as well as transition dipole moments (TDMs) between electronic states of different manifolds (see Figure 2a). The RASSCF/RASPT2 parametrization of the LVC model including the seven lowest valence-excited states has been described elsewhere.[31] The core-excited states were calculated following the protocol documented in ref 44. The high-lying core-excited states were directly obtained at the RASSCF level by putting one core orbital at a time in an orbital subspace (RAS1) and excluding from the configurational space configuration state functions (CFSs) in which it is doubly occupied (by means of the HEXS keyword); moreover, we avoided possible rotations of the core orbital out of the active space (by means of the SUPSYM keyword), which might occur as the RASSCF optimizer will favor their substitution with "higher-lying" inactive valence orbitals (to describe lower-lying valence-excited states), paying the price of freezing the considered core orbital in its SCF shape. Their energy was corrected by applying the multistate version of RASPT2 (with the FROZEN = 0 option).

Specifically, due to the D_{2h} symmetry of pyrene in its ground-state minimum, one only needs to consider core excitations from the five nonequivalent carbon centers (see Section S1 of the Supporting Information for the pyrene structure and the five carbon centers): for each of them, a 15-

state RASSCF/RASPT2 calculation was performed putting the corresponding 1s core orbital in RAS1, resulting in a total of 75 core-excited states. The rationale behind this choice is based on the spectral window we decided to simulate (the pre-edge region between 275 and 285 eV, which is expected to be less congested with respect to the region above the carbon IP energy, which lies at about 290 eV), making the computation of additional higher-lying core-excited states not relevant. The core-excited manifold was computed with C_s symmetry: this allows us to describe the relaxation effect that involves the valence orbitals upon core excitation, with a considerable electron density redistribution on the top of the created core-hole and a consequent stabilization of the core-excited state energies. Solutions that exactly respect the D_{2h} symmetry would not be capable of describing this orbital localization effect. Moreover, in pyrene, most of the equivalent (isoenergetic) cores lie in distant parts of the molecule, making their overlap negligible and any possible delocalization effect not relevant in this specific case. Within the C_s symmetry, all π and π^* orbitals belong to the a'' irreducible representation, while core orbitals belong to a'. The low-lying valence states of $\pi\pi^*$ nature will therefore be of A' symmetry, while the core-to-π^* excited states will be of A'' symmetry. The lower symmetry makes impracticable computations with the extended active space (i.e., full-π augmented with a set of virtual orbitals) reported in our previous study of the valence manifold (ref [31]); therefore, it was reduced to the (full-π) 16 frontier π and π^* orbitals (16 electrons, with maximum four excitations), plus the carbon 1s orbital bearing two additional electrons [with maximum one excitation; the final active space will be labeled as RAS(1,1|4,8|4,8)]. Moreover, the high number of core-excited states and the increased cost of RASPT2 without frozen orbitals make the calculation of RASPT2 gradients (either numerical or analytical) unfeasible; therefore, we relied on RASSCF gradients for these states.

In order to compute TDMs between valence and core states, they all have to be obtained at the same level of theory (active space size and composition); therefore, their calculation was repeated as outlined above, keeping the desired core orbital in RAS1 but including CSFs with doubly occupied core orbitals in the configurational space (i.e., not employing the HEXS keyword).

In passing, we note that a similar procedure can be employed to compute core-ionized states, required for simulating X-ray photoelectron spectroscopy (XPS) and time-resolved XPS (TR-XPS) spectra. In that case, instead of TDMs, one could evaluate Dyson orbital norms as the approximate ionization cross-section. The protocol is detailed in ref [44].

All electronic structure calculations were performed with OpenMolcas[45−47] using the generally contracted relativistic atomic natural orbital basis set (ANO-RCC[48] with 4s3p2d and 2s1p contraction for C and H atoms, respectively) and applying Cholesky decomposition in the calculation of two-electron integrals. In all RASPT2 calculations, the imaginary shift was set to 0.2 au and the IP/EA shift to zero.

Quantum dynamical propagation of vibronic WPs was performed with the MCTDH method as implemented in the code Quantics.[49] We adopted a primitive basis set of Hermite DVR functions and a multiset formulation for the single particle functions (SPFs) which make it easier to project the part of the WP residing on a specific valence state e to a

specific core-excited state c. As for the integrator section, we used a constant mean field approach, a Bulirsch−Stoer extrapolation integrator for SPFs, and a short iterative Lanczos for the multiconfigurational A coefficients.[22]

2.2. Spectroscopy Simulations. By leveraging the time-scale separation between core-excited state lifetime (from sub-fs to a few fs) and electronic-state transfer/vibrational motion of the nuclei (with hydrogen bonds exhibiting the shortest vibrational period of about 10 fs), we now derive the approximate energy-gap/energy-gap variance expressions for the evaluation of first- and third-order response functions, required for the simulation of linear and TA spectroscopy. The basic derivation steps are reported here; additional details are provided in the Supporting Information.

2.2.1. System Hamiltonian and Wave Functions. The molecular Hamiltonian is given by

$$\hat{H} = \sum_{a\in\{\mathcal{G},\mathcal{E},C\}} (E_a^{(\mathrm{ad})} + \hat{H}_a)|a\rangle\langle a| + \sum_{e\neq e'\in\mathcal{E}} \hat{U}_{ee'}|e\rangle\langle e'| \tag{1}$$

where a runs over three manifolds of electronic states: the ground-state $\mathcal{G}$ manifold (that typically only contains a single GS state, g), the $\mathcal{E}$ manifold in which the nonadiabatic dynamics occurs, and the C manifold of the core-excited states probed by the X-ray probe pulse; $E_a^{(\mathrm{ad})}$ is the adiabatic energy of the electronic state a. $\hat{H}_a$ and $\hat{U}_{ee'}$ represent the dependence over the nuclear coordinates, and in the LVC model (employing dimensionless coordinates) they read

$$\hat{H}_a = \sum_m \hat{H}_{a,m} = \sum_m \hat{T}_m + \hat{V}_{a,m}$$
$$= \sum_m \hat{T}_m + \frac{1}{2}\omega_m(\hat{Q}_m - d_{a,m})^2$$
$$\hat{U}_{ee'} = \sum_m \hat{U}_{ee',m} = \sum_m \lambda_{ee',m}\hat{Q}_m \tag{2}$$

where m runs over the normal modes (accounted for in the LVC model), a denotes a generic state of the three manifolds, while g, e/e', and c denote states that belong to the $\mathcal{G}$, $\mathcal{E}$, and C manifolds, respectively. $\hat{T}_m$ is the kinetic energy operator along mode m,[a] and the potential energy surface (PES) term $\hat{V}_{a,m}$ is given by a displaced harmonic oscillator (DHO) potential, whose displacement for the electronic state a and along the given mode m is given by $d_{a,m}$.[b] We further assume that all electronic states a have the same set of normal modes and frequencies (generally computed for the ground state g). Without loss of generality, we assume $d_{g,m} = 0\ \forall\ m$, so that

$$\hat{V}_{g,m} = \frac{1}{2}\omega_m\hat{Q}_m^{\ 2}$$
$$\hat{V}_{c,m} = \frac{1}{2}\omega_m(\hat{Q}_m - d_{c,m})^2 = \hat{V}_{g,m} + \lambda_{c,m} - \omega_m\hat{Q}_m d_{c,m} \tag{3}$$

where Q_m is the normal mode coordinate of mode m, $d_{c,m}$ is the displacement (along mode m) of the c state PES with respect to the ground state PES, and $\lambda_{c,m} = \frac{1}{2}\omega_m(d_{c,m})^2$ is the reorganization energy of state c along mode m (see Figure 1).

$\hat{U}_{ee'}$ promotes the nonadiabatic dynamics between electronic states of the $\mathcal{E}$ manifold. By leveraging on the time-scale separation between the extremely short core-excited-state lifetime and the longer population-transfer time, the term $\hat{U}_{ee'}$
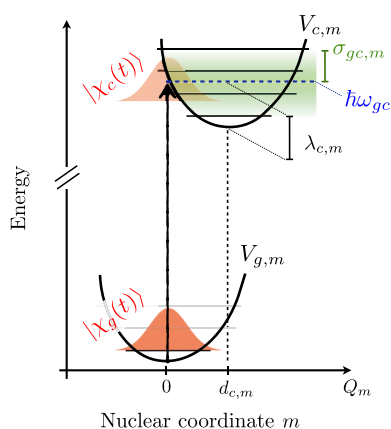
Figure 1. Schematic representation of the potentials $V_{g,m}$ and $V_{c,m}$ along a given mode m, the displacement $d_{c,m}$, the vertical energy gap $\hbar\omega_{gc}$, the reorganization energy $\lambda_{c,m}$, the GS and ES WPs ($|\chi_g(t)\rangle$ and $|\chi_c(t)\rangle$, reported here for $t = 0$), and the energy-gap standard deviation $\sigma_{gc,m}$.

can be neglected during core excitations. For the same reason, we neglect the nonadiabatic coupling between electronic states of the C manifold. We will consider projections of the Hamiltonian in eq 2 onto various electronic state manifolds, i.e.,

$$\hat{H}_{\mathcal{G}} = (E_g^{(ad)} + \hat{H}_g)|g\rangle\langle g|$$

$$\hat{H}_{\mathcal{E}} = \sum_{e \in \mathcal{E}} (E_e^{(ad)} + \hat{H}_e)|e\rangle\langle e| + \sum_{e \neq e' \in \mathcal{E}} \hat{U}_{ee'}|e\rangle\langle e'|$$

$$\hat{H}_C = \sum_{c \in C} (E_c^{(ad)} + \hat{H}_c)|c\rangle\langle c| \quad (4)$$

We assume that electronic states belonging to different manifolds are coupled only via dipole interactions (i.e., only the external electromagnetic field can promote a change of manifold),[c] which in the Condon approximation is given by

$$\hat{\boldsymbol{\mu}} = \sum_e \boldsymbol{\mu}_{ge}|g\rangle\langle e| + \sum_c \boldsymbol{\mu}_{gc}|g\rangle\langle c| + \sum_{e,c} \boldsymbol{\mu}_{ec}|e\rangle\langle c| + h.\,c. \quad (5)$$

It is useful here to define some projections of the dipole moment onto specific manifolds, i.e.,

$$\hat{\boldsymbol{\mu}}_{\mathcal{GE}} = \sum_e \boldsymbol{\mu}_{ge}|g\rangle\langle e|;\ \hat{\boldsymbol{\mu}}_{\mathcal{EG}} = \hat{\boldsymbol{\mu}}_{\mathcal{GE}}^\dagger$$

$$\hat{\boldsymbol{\mu}}_{\mathcal{GC}} = \sum_c \boldsymbol{\mu}_{gc}|g\rangle\langle c|;\ \hat{\boldsymbol{\mu}}_{\mathcal{CG}} = \hat{\boldsymbol{\mu}}_{\mathcal{GC}}^\dagger$$

$$\hat{\boldsymbol{\mu}}_{\mathcal{EC}} = \sum_{ec} \boldsymbol{\mu}_{ec}|e\rangle\langle c|;\ \hat{\boldsymbol{\mu}}_{\mathcal{CE}} = \hat{\boldsymbol{\mu}}_{\mathcal{EC}}^\dagger \quad (6)$$

By following the steps of ref 1, the molecular wave function at a given time t is defined as

$$|\Psi(Q, q, t)\rangle = \sum_{a \in \{\mathcal{G}, \mathcal{E}, C\}} |\phi_a(q;\, Q)\rangle|\chi_a(t,\, Q)\rangle \quad (7)$$

where $|\chi_a(t,\, Q)\rangle$ represents the nuclear WP evolving on the a-th electronic state $|\phi_a(q;\, Q)\rangle$ and q denotes the electronic degrees of freedom. By employing a *diabatic* formulation, the electronic wave function becomes (ideally) independent of the nuclear coordinates. Moreover, the condition $\langle\Psi(Q, q, t)|\Psi(Q, q, t)\rangle = 1$ holds at all times t. The

electronic wave function $|\phi_a(q)\rangle$ is conveniently denoted as $|g\rangle$, $|e\rangle$ and $|c\rangle$ for states in the three manifolds; hereafter, we will drop the explicit coordinate dependence, so that

$$|\Psi(t)\rangle = \sum_{a \in \{\mathcal{G}, \mathcal{E}, C\}} |a\rangle|\chi_a(t)\rangle \quad (8)$$

Finally, in what follows, we will use the notation $a \to b$ in the subscript of the nuclear WP to denote the transfer process that occurs between the two electronic states a and b; more precisely, $|\chi_{a \to b}(t)\rangle$ corresponds to that *fraction* of the nuclear WP initially prepared in a at $t = 0$, that is found in b at time t. This means that at each time, $|\chi_b(t)\rangle = \sum_a |\chi_{a \to b}(t)\rangle$, which is the sum over all the WPs that, initially prepared in various electronic states a at time $t = 0$, evolved driven by the nonadiabatic dynamics along t, and now contribute to shaping the WP of state b at time t.

2.2.2. Linear Spectra: X-ray Absorption Spectroscopy/XPS. The first-order response function within a QD formulation, employing the WPO approach and without making any assumption besides the Condon approximation, is given by (being interested in the X-ray response, we only consider transitions to C states and neglect those to $\mathcal{E}$ ones)[1]

$$R_{\mathrm{WPO}}^{(1)}(t) = \left(\frac{i}{\hbar}\right)\theta(t)\sum_{c,c'} \mu_{gc'}\mu_{cg}\langle\chi_g(t)|\chi_{c \to c'}(t)\rangle e^{-t/2\tau_c} \quad (9)$$

where $e^{-t/2\tau_c}$ is an exponential decay factor that accounts for the finite-core excited-state lifetime τ_c (here assumed to be identical for all the states in the C manifold),[d] which quickly eliminates the $g{-}c'$ coherence; $\langle\chi_g(t)|\chi_{c \to c'}(t)\rangle$ is the $g{-}c'$ nuclear WP overlap at time t. Note that in general both the Gaussian and exponential decay factors can be considered: the second accounts for the finite lifetime, the first, e.g., for the static disorder that could also contribute to the overall spectral broadening.

Assuming the c state lifetime to be shorter than the time-scale of the population transfer in the C manifold allows us to neglect $c \to c'$ transfer process; this means that the dynamics on each initially populated c state is purely adiabatic and unaffected by the other states. The double summation in eq 9 reduces to a single summation over the states in the C manifold, and we obtain

$$R_{\mathrm{WPO}^*}^{(1)}(t) = \left(\frac{i}{\hbar}\right)\theta(t)\sum_c \mu_{gc}\mu_{cg}\langle\chi_g(t)|\chi_{c^*}(t)\rangle e^{-t/2\tau_c} \quad (10)$$

The * symbol in the WPO* subscript implies that we have assumed that transfer is slow compared to the lifetime,[e] so that one can consider an adiabatic dynamics that remains in the given state c for all t times (as indicated by the symbol c^*). In what follows for brevity, we set $\hbar = 1$.

The nuclear overlap factor $\langle\chi_g(t)|\chi_{c^*}(t)\rangle$ can be rewritten as

$$\langle\chi_g(t)|\chi_{c^*}(t)\rangle = \langle\chi_g(0)|e^{+i(E_g^{(ad)} + \hat{H}_g)t}e^{-i(E_c^{(ad)} + \hat{H}_c)t}|\chi_c(0)\rangle =$$

$$= e^{-i\omega_{gc}^{(ad)}t}\prod_m \langle\chi_{g,m}(0)|e^{+i\hat{H}_{g,m}t}e^{-i\hat{H}_{c,m}t}|\chi_{c,m}(0)\rangle \quad (11)$$

where $\omega_{gc}^{(ad)} = \frac{\Delta E_{gc}^{(ad)}}{\hbar}$, with $\Delta E_{gc}^{(ad)}$ being the adiabatic transition energy between states g and c (that in the DHO

model corresponds to the 0−0 energy), and the propagators of the bra and ket WPs appear. Let us focus on one mode, m. At zero temperature, the GS WP is prepared in the lowest vibrational level of each mode; furthermore, by assuming impulsive interactions with the external electromagnetic field, we have $|\chi_{c,m}(0)\rangle = |\chi_{g,m}(0)\rangle = |v_{m,0}\rangle$, with $|v_{m,0}\rangle$ being the lowest GS vibrational eigenstate along mode m. This means that the laser field creates an exact copy of the GS WP in the ES. $\hat{H}_g |v_{m,0}\rangle = \epsilon_{m,0} |v_{m,0}\rangle$, with $\epsilon_{m,0} = \hbar \frac{\omega_m}{2}$ and is the m-th mode frequency. Considering all of that, we get

$$\langle \chi_g(t) | \chi_{c^*}(t) \rangle = e^{-i\omega_{gc}^{(ad)}t} \prod_m e^{+i\epsilon_{m,0}t} \langle v_{m,0} | e^{-i\hat{H}_{c,m}t} | v_{m,0} \rangle \tag{12}$$

There exist different approaches to evaluate the propagator of eq 11, both *time-independent* (expand the WPs in terms of the Hartree product of Harmonic basis functions) and *time-dependent* (compute the overlap explicitly, via a QD simulation, explored in ref 1 in the more general case of nonadiabatic dynamics, i.e., when the $c \rightarrow c'$ transfer could happen). In the present case of adiabatic dynamics, one may write analytical expressions for eq 12 which are exact.[50,51] These various approaches are summarized in Section S2 of the Supporting Information.

Even if an analytical evaluation of eq 12 is possible, we nonetheless consider a different approach, which is the focus of the present work and set the steps that will be followed in the TA case, where complete analytical expressions are not available (some derivations are reported in Section S2 of the Supporting Information). Let us define two quantities, the average energy of the projected GS WP onto the ES well (along mode m), $\langle \hat{H}_c \rangle_{g,m}$, and the energy-gap variance between states g and c, $\sigma_{gc;m}^2$. These read, respectively, as

$$\langle \hat{H}_c \rangle_{g,m} = \langle v_{m,0} | \hat{H}_c | v_{m,0} \rangle = \epsilon_{m,0} + \lambda_{c,m} \tag{13}$$

and

$$\sigma_{gc;m}^2 = \langle v_{m,0} | (\hat{H}_c - \langle \hat{H}_c \rangle_{g,m})^2 | v_{m,0} \rangle = \frac{1}{2} \omega_m^2 (d_{c,m})^2 = \omega_m \lambda_{c,m} \tag{14}$$

where these equalities, which hold for the DHO model, admit straightforward generalization to analytical expressions that also include Duschinsky and temperature effects.

Let us go back to the overlap expression of eq 12. By adding and subtracting $\langle \hat{H}_c \rangle_{g,m}$ to $\hat{H}_{c,m}$ in the exponent, we get

$$e^{+i\epsilon_{m,0}t} e^{-i\langle \hat{H}_c \rangle_{g,m}t} \langle v_{m,0} | e^{-i(\hat{H}_{c,m} - \langle \hat{H}_c \rangle_{g,m})t} | v_{m,0} \rangle =$$
$$= e^{-i\lambda_{c,m}t} \langle v_{m,0} | e^{-i(\hat{H}_{c,m} - \langle \hat{H}_c \rangle_{g,m})t} | v_{m,0} \rangle \tag{15}$$

Let us now expand the exponential operator in the overlap term in a Taylor series, obtaining

$$\langle v_{m,0} | e^{-i(\hat{H}_{c,m} - \langle \hat{H}_c \rangle_{g,m})t} | v_{m,0} \rangle = \langle v_{m,0} | 1 - i(\hat{H}_{c,m} - \langle \hat{H}_c \rangle_{g,m})t$$
$$- \frac{1}{2}(\hat{H}_{c,m} - \langle \hat{H}_c \rangle_{g,m})^2 t^2 + ... | v_{m,0} \rangle$$
$$= 1 - 0t - \frac{1}{2}\sigma_{gc;m}^2 t^2 + ...$$
$$= 1 - \frac{1}{2}\sigma_{gc;m}^2 t^2 + ... \tag{16}$$

At this point, we invoke the time-scale separation between vibrational motions on the (core-)excited state and its lifetime: we assume the c state lifetime to be shorter than the fastest (spectroscopically relevant) vibration in the system. This allows us to truncate the expansion to second order for short t. Equation 16 is then equal to the expansion of a Gaussian to the second order, so that

$$\langle v_{m,0} | e^{-i(\hat{H}_{c,m} - \langle \hat{H}_c \rangle_{g,m})t} | v_{m,0} \rangle = 1 - \frac{1}{2}\sigma_{gc;m}^2 t^2 ... \sim e^{-\frac{1}{2}\sigma_{gc;m}^2 t^2} \tag{17}$$

Putting all the pieces together finally gives

$$\langle \chi_{g,m}(t) | \chi_{c^*,m}(t) \rangle \sim e^{-i\lambda_{c,m}t} e^{-\frac{1}{2}\sigma_{gc;m}^2 t^2} \tag{18}$$

which is valid in the assumption of time-scale separation between vibrations and the c state lifetime.

Considering all N modes and following the same steps detailed above for a single mode m, one obtains an approximate expression for $R_{\text{WPO}^*}^{(1)}(t)$ in terms of the vertical energy gap and the energy-gap variance, i.e.,

$$R_{\text{WPO}^*}^{(1)}(t) \sim R_{\text{EGVA}}^{(1)}(t) = \left(\frac{i}{\hbar}\right)\theta(t)\sum_c \mu_{gc}\mu_{cg} e^{-i\omega_{gc}t - \frac{1}{2}\varsigma_{gc}^2 t^2 - t/2\tau_c} \tag{19}$$

where the subscript EGVA denotes the *energy-gap/energy-gap variance approximation*, $\omega_{gc} = \omega_{gc}^{(ad)} + \sum_m \lambda_{c,m}$ is the vertical energy-gap frequency, and $\varsigma_{gc}^2 = \sum_m \sigma_{gc;m}^2$ is the total variance, i.e., the sum of variances along all modes.

We already noted that an analytical expression for $R_{\text{WPO}^*}^{(1)}(t)$ can be written in terms of line shape functions.[50] In the adopted notation, this reads

$$R_{\text{WPO}^*}^{(1)}(t) = \left(\frac{i}{\hbar}\right)\theta(t)\sum_c \mu_{gc}\mu_{cg} e^{-i\omega_{gc}^{(ad)}t - g_{cc}(t) - t/2\tau_c} \tag{20}$$

where $g_{ab}(t)$ is the so-called line shape function (whose expression is given in Section S3 of the Supporting Information). Interestingly, the same result of eq 19 can be derived by a second-order Taylor expansion of the line shape function, obtaining $g_{cc}(t) \sim i\lambda_c t - \frac{1}{2}\varsigma_{gc}^2 t^2$ (see the Supporting Information).

In eq 19, all the pieces required for the linear absorption spectrum in the energy gap/energy-gap variance limit become apparent: the center of the transition (the first momentum of the spectrum) is determined by the vertical energy gap ω_{gc}, while the broadening (the second momentum of the spectrum) is given by the WP energy width term $e^{-\frac{1}{2}\varsigma_{gc}^2 t^2}$. The lifetime term $e^{-t/2\tau_c}$ further contributes to the line shape. By Fourier transformation of eq 19 along the time t, one get a so-called *Voigt* line shape profile. The Voigt profile is symmetric; thus, the derived expressions are not able to describe the asymmetry of the vibronic spectra. This follows from truncating the Taylor series at the second order, i.e., by computing the spectrum employing only its first two moments only. The band asymmetry indeed depends on the higher spectrum moments. Interestingly, the manipulations we have performed are closely related to those introduced to achieve a semiclassical approximation of the LA spectrum,[50,52,53] and the first two moments are the only two that can be evaluated exactly in such a framework (if the Franck−Condon approximation holds), whereas fully QD approaches are

mandatory to compute accurately higher-order moments of the spectrum.[52]

While closing this section, we remark the fact that even taking into account Duschinsky mixings and temperature effects it is possible to derive analytical expressions both for the energy gap and the energy-gap variance[52,54,55] as well as for the full correlation function (see, e.g., eq 10) whose Fourier transform gives the LA spectrum.[53,56,57]

In the Results and Discussion section, we will compare X-ray absorption spectroscopy (XAS) spectra obtained with the WPO* expression (eq 10), the approximated EGVA expression (eq 19), and the EGA expression, which completely neglects the broadening induced by the WP *width* and is obtained truncating the WP overlap Taylor series to the first order in t.

2.2.3. Nonlinear Spectra: TR-XAS/TR-XPS. We consider a TR-XAS experiment, in which the pump pulse (in the vis/UV spectral region) excites the system in the valence manifold of states $\mathcal{E}$ (at $t = 0$), and after a time interval t_2 an X-ray probe pulse further excites the system into the core-excited manifold of states C. Note that everything applies very similarly also to TR-XPS, for which the C manifold should include core-ionized states instead of core-excited states.

By following similar steps to those that led from eq 10 to eq 19, we now derive the approximate expression for the nonlinear (third-order) response. The main difference is that the WP that is projected (impulsively) by the probe pulse from the $\mathcal{E}$ to the C manifold, is no more the cold GS WP: it is rather the QD (nonadiabatically) evolved WP along t_2.

The third-order response function reads

$$R^{(3)}_{\mathrm{WPO}}(t_3, t_2)=$$

$$=+\langle\Psi_g(0)|\hat{\boldsymbol{\mu}}_{\mathcal{G}C}\hat{\boldsymbol{\mu}}_{C\mathcal{G}}\,e^{i\hat{H}_{\mathcal{G}}(t_2+t_3)}\hat{\boldsymbol{\mu}}_{\mathcal{G}C}\,e^{-i\hat{H}_C t_3}\hat{\boldsymbol{\mu}}_{C\mathcal{G}}\,e^{-i\hat{H}_{\mathcal{G}}t_2}|\Psi_g(0)\rangle e^{-t_3/2\tau_c}\,+$$

$$-\langle\Psi_g(0)|\hat{\boldsymbol{\mu}}_{\mathcal{G}\mathcal{E}}\,e^{i\hat{H}_{\mathcal{E}}(t_2+t_3)}\hat{\boldsymbol{\mu}}_{\mathcal{E}C}\,e^{-i\hat{H}_C t_3}\hat{\boldsymbol{\mu}}_{C\mathcal{E}}\,e^{-i\hat{H}_{\mathcal{E}}t_2}\hat{\boldsymbol{\mu}}_{\mathcal{E}\mathcal{G}}|\Psi_g(0)\rangle e^{-t_3/2\tau_c}\tag{21}$$

In writing eq 21, we have removed the common $(i/\hbar)^3\theta(t_2)\theta(t_3)$ prefactor and we have expressed both Hamiltonian and dipole operators as acting in/between specific manifolds. Note that only the ground-state bleaching (GSB, first term) and the excited-state absorption (ESA, second term) contributions are reported in eq 21, as no stimulated emission is possible in this vis/UV pump−X-ray probe setup. The Feynman diagrams for the GSB and ESA contributions are drawn in Figure 2.

We now express the response function equations in a form that naturally captures the *resulting* effect of the t_2 quantum dynamics without splitting all the pathways over the initial conditions, which is un-necessary. In order to do so, we incorporate the initial dipole−moment interaction in the wave function, so that

$$|\Psi_{\mathcal{E}}(0)\rangle = \frac{\hat{\boldsymbol{\mu}}_{\mathcal{E}\mathcal{G}}}{\sqrt{\mathcal{N}_\mu}}|\Psi_g(0)\rangle = \sum_{e\in\mathcal{E}}\frac{\mu^{\epsilon}_{eg}}{\sqrt{\mathcal{N}_\mu}}|e\rangle|\chi_e(0)\rangle$$

$$= \sum_{e\in\mathcal{E}}|e\rangle|\chi^{\mu}_e(0)\rangle\tag{22}$$

where $|\chi^{\mu}_e(0)\rangle = \mu^{\epsilon}_{eg}|\chi_e(0)\rangle/\sqrt{\mathcal{N}_\mu}$, with $\mathcal{N}_\mu = \sum_e(\mu^{\epsilon}_{ge})^2$ being a normalization factor to keep the wave function norm equal to 1; due to the impulsive nature of the laser field, we have that $|\chi_e(0)\rangle = |\chi_g(0)\rangle$, as previously specified. $\mu^{\epsilon}_{ge} = \boldsymbol{\mu}_{ge}\cdot\boldsymbol{\epsilon}$, with $\boldsymbol{\epsilon}$
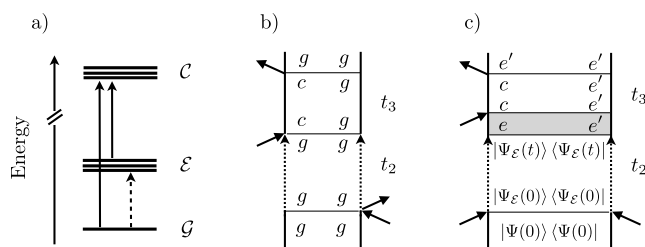

Figure 2. (a) Schematic of the electronic state manifolds and allowed transitions (*pump*: dashed arrow; *probe*: full arrows). (b) Feynman diagram for the GSB contribution. (c) Feynman diagram for the ESA contribution in which the initial conditions are implicitly accounted for in Ψ. In (c), the gray area highlights the fact that $\Psi_{\mathcal{E}}(t_2)$ is unpacked in its contributions, and then the various $\boldsymbol{\mu}_{ec}$ transitions and the subsequent $|c\rangle\langle e'|$ coherences are considered.

being the pump-pulse field polarization. Note that the μ symbol over $|\chi^{\mu}_e\rangle$ is used to keep track of the fact that the WP is prepared in various PESs according to the field−matter interaction. We then rewrite eq 21 as follows

$$R^{(3)}_{\mathrm{WPO}}(t_3, t_2)=$$

$$=+\sum_e\mu_{ge}\mu_{eg}\sum_{c,c'}\mu_{gc}\mu_{cg}\langle\chi_g(t_3)|\chi_{c\to c'}(t_3)\rangle e^{-t_3/2\tau_c}+$$

$$-\,\mathcal{N}_\mu\langle\Psi_{\mathcal{E}}(t_2)|e^{i\hat{H}_{\mathcal{E}}t_3}\hat{\boldsymbol{\mu}}_{\mathcal{E}C}\,e^{-i\hat{H}_C t_3}\hat{\boldsymbol{\mu}}_{C\mathcal{E}}|\Psi_{\mathcal{E}}(t_2)\rangle e^{-t_3/2\tau_c}\tag{23}$$

The GSB contribution closely resembles the first-order contribution presented in the previous section and therefore needs no further discussion. The ESA contribution, instead, requires that the WP, prepared at time $t_2 = 0$ in (a given electronic state, or in a linear combination of electronic states in) the $\mathcal{E}$ manifold, has evolved along t_2. We thus focus on this latter contribution, labeled as $R^{(3)\mathrm{ESA}}_{\mathrm{WPO}}(t_3,t_2)$.

Let us take a closer look at the t_2 evolved $|\Psi_{\mathcal{E}}(0)\rangle$ wave function

$$|\Psi_{\mathcal{E}}(t_2)\rangle = e^{-i\hat{H}_{\mathcal{E}}t_2}|\Psi_{\mathcal{E}}(0)\rangle=$$

$$= e^{-i\hat{H}_{\mathcal{E}}t_2}\sum_{e\in\mathcal{E}}|e\rangle|\chi^{\mu}_e(0)\rangle=$$

$$= \sum_{e'\in\mathcal{E}}|e'\rangle[\sum_{e\in\mathcal{E}}|\chi^{\mu}_{e\to e'}(t_2)\rangle]=$$

$$= \sum_{e'\in\mathcal{E}}|e'\rangle|\chi^{\mu}_{e'}(t_2)\rangle\tag{24}$$

Notice that in the last step the effect of the sum over all the initial states $e\in\mathcal{E}$ is captured by the *final* amplitudes/WPs in each electronic state $e'\in\mathcal{E}$, i.e., $|\chi^{\mu}_{e'}(t_2)\rangle$. The idea here is that the propagation of a wave function prepared in $\mathcal{E}$ along t_2 will remain in $\mathcal{E}$ but will reshuffle the initially prepared WP, as dictated by the nonadiabatic dynamics, so that every contribution that was on $|e\rangle$ at time 0 can be redistributed among all other members $|e'\rangle$ of $\mathcal{E}$ along t_2, and the final wave function can therefore be written as $|\Psi_{\mathcal{E}}(t_2)\rangle=\sum_{e\in\mathcal{E}}|e\rangle|\chi^{\mu}_e(t_2)\rangle$. See Figure 3 for a pictorial representation of this process.

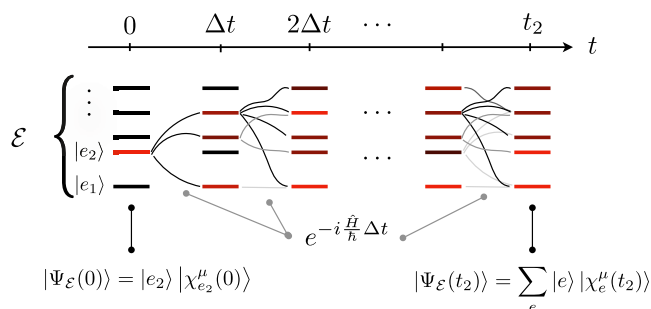The ESA contribution to the WPO third-order response therefore reads

Figure 3. Schematic representation of the WF evolution in the $\mathcal{E}$ manifold under the action of the Hamiltonian $\hat{H}$ that contains nonadiabatic coupling terms between the electronic states e. In the example reported in the figure, the wave function is initially prepared in the electronic state e_2 (i.e., $c_{e_2}(0) = 1$, while all the other amplitudes are set to zero).

$$R^{(3)\text{ESA}}_{\text{WPO}}(t_3, t_2)=$$

$$= -\,\mathcal{N}_\mu \langle \Psi_\mathcal{E}(t_2)|e^{i\hat{H}_\mathcal{E}t_3}\hat{\boldsymbol{\mu}}_{\mathcal{E}C}\,e^{-i\hat{H}_C t_3}\hat{\boldsymbol{\mu}}_{C\mathcal{E}}|\Psi_\mathcal{E}(t_2)\rangle e^{-t_3/2\tau_c}$$

$$= -\,\mathcal{N}_\mu \sum_{e,e'} \langle \chi^\mu_e(t_2)|\langle e|e^{i\hat{H}_\mathcal{E}t_3}\hat{\boldsymbol{\mu}}_{\mathcal{E}C}\,e^{-i\hat{H}_C t_3}\hat{\boldsymbol{\mu}}_{C\mathcal{E}}|e'\rangle|\chi^\mu_{e'}(t_2)\rangle e^{-t_3/2\tau_c}$$

$$= -\,\mathcal{N}_\mu \sum_{e,e',e'',e''',c} \boldsymbol{\mu}_{e''c}\boldsymbol{\mu}_{ce'''}\langle \chi^\mu_{e,e\to e''}(t_2,t_3)|\chi^\mu_{e',c*}(t_2,t_3)\rangle e^{-t_3/2\tau_c}$$

$$(25)$$

where we employ a notation that keeps track of the history of the WP: the bra side evolved on the $\mathcal{E}$ manifold for the full $t_2 + t_3$ time interval, while the ket side first evolved in the $\mathcal{E}$ manifold along t_2 was projected into the C manifold and thus evolved again in state c along t_3.[f]

By invoking the time-scale separation between the lifetime of the c state (τ_σ, fast) and transfer (slow), transfer along t_3 can be neglected (what we referred to as the WPO* level of theory), and therefore an electronic state (in whatever manifold) will not change along this time interval. The ESA contribution to eq 23 therefore becomes

$$R^{(3)\text{ESA}}_{\text{WPO}*}(t_3, t_2)=$$

$$= -\,\mathcal{N}_\mu \sum_{e,e',c} \langle \chi^\mu_e(t_2)|e^{i(E_e^{(\text{ad})}+\hat{H}_e)t_3}\boldsymbol{\mu}_{ec}\,e^{-i(E_c^{(\text{ad})}+\hat{H}_c)t_3}\boldsymbol{\mu}_{ce'}|\chi^\mu_{e'}(t_2)\rangle e^{-t_3/2\tau_c}$$

$$= -\,\mathcal{N}_\mu \sum_{e,e',c} \boldsymbol{\mu}_{ec}\boldsymbol{\mu}_{ce'}e^{-i\omega_{ec}^{(\text{ad})}t_3}\langle \chi^\mu_e(t_2)|e^{i\hat{H}_e t_3}e^{-i\hat{H}_c t_3}|\chi^\mu_{e'}(t_2)\rangle e^{-t_3/2\tau_c}$$

$$(26)$$

where the * symbol indicates that we have invoked the above-mentioned time scale separation, for which the dynamics of a WP on a given electronic state is driven only by the Hamiltonian of such state (i.e., $\hat{U}_{ee'=0}$ in this time interval).

We have also made use of the fact that the WP on the state c are identical to those of e' at $t_3 = 0$.

To implement the energy-gap approximation, we rewrite eq 26 by splitting *population* and *coherence* contributions (i.e., terms for which the bra and ket WPs are identical or different, respectively), obtaining

$$R^{(3)\text{ESA}}_{\text{WPO}*}(t_3, t_2)=$$

$$= -\,\mathcal{N}_\mu \sum_{e,c} \boldsymbol{\mu}_{ec}\boldsymbol{\mu}_{ce}e^{-i\omega_{ec}^{(\text{ad})}t_3}\langle \chi^\mu_e(t_2)|e^{i\hat{H}_e t_3}e^{-i\hat{H}_c t_3}|\chi^\mu_e(t_2)\rangle e^{-t_3/2\tau_c}+$$

$$- \,\mathcal{N}_\mu \sum_{e\neq e',c} \boldsymbol{\mu}_{ec}\boldsymbol{\mu}_{ce'}e^{-i\omega_{ec}^{(\text{ad})}t_3}\langle \chi^\mu_e(t_2)|e^{i\hat{H}_e t_3}e^{-i\hat{H}_c t_3}|\chi^\mu_{e'}(t_2)\rangle e^{-t_3/2\tau_c}$$

$$(27)$$

We now consider a normalized nuclear WP at each time t_2, i.e., $|\bar{\chi}^\mu_e\rangle = |\chi^\mu_e\rangle / \sqrt{\langle \chi^\mu_e|\chi^\mu_e\rangle}$.[g] The square of the normalization factor has a simple physical interpretation: it gives the probability to find the WP in a given electronic state e, i.e., it is the *population* of the electronic state e, hereafter referred to as $\rho_e(t_2)$. Therefore, we have

$$R^{(3)\text{ESA}}_{\text{WPO}*}(t_3, t_2)=$$

$$= -\,\mathcal{N}_\mu \sum_{e,c} \boldsymbol{\mu}_{ec}\boldsymbol{\mu}_{ce}\rho_e(t_2)e^{-i\omega_{ec}^{(\text{ad})}t_3}\langle \bar{\chi}^\mu_e(t_2)|e^{i\hat{H}_e t_3}e^{-i\hat{H}_c t_3}|\bar{\chi}^\mu_e(t_2)\rangle e^{-t_3/2\tau_c}$$

$$- \,\mathcal{N}_\mu \sum_{e\neq e',c} \sqrt{\rho_e(t_2)\rho_{e'}(t_2)}\,\boldsymbol{\mu}_{ec}\boldsymbol{\mu}_{ce'}e^{-i\omega_{ec}^{(\text{ad})}t_3}\langle \bar{\chi}^\mu_e(t_2)|e^{i\hat{H}_e t_3}e^{-i\hat{H}_c t_3}|\bar{\chi}^\mu_{e'}(t_2)\rangle e^{-t_3/2\tau_c}$$

$$(28)$$

We next turn to the bracket term. Let us again consider a single mode m; also, we here derived the population contribution ($e = e'$): a similar derivation for the coherence contribution ($e \neq e'$) is reported in Section S5 of the Supporting Information.

First, we perform a similar trick previously done, by introducing $\langle \hat{H}_e\rangle_{t_2} = \langle \bar{\chi}^\mu_{e,m}(t_2)|\hat{H}_e|\bar{\chi}^\mu_{e,m}(t_2)\rangle$ and $\langle \hat{H}_c\rangle_{t_2} = \langle \bar{\chi}^\mu_{e,m}(t_2)|\hat{H}_c|\bar{\chi}^\mu_{e,m}(t_2)\rangle$,[h] where the t_2 subscript highlights the fact that these quantities depend on t_2 (i.e., on the shape of the WPs on the e PESs, just before projecting it onto C states); we then rewrite the exponential operators as

$$e^{+i\hat{H}_e t_3} = e^{+i\langle \hat{H}_e\rangle_{t_2}t_3}e^{+i(\hat{H}_e-\langle \hat{H}_e\rangle_{t_2})t_3}$$

$$e^{-i\hat{H}_c t_3} = e^{-i(\hat{H}_c-\langle \hat{H}_c\rangle_{t_2})t_3}e^{-i\langle \hat{H}_c\rangle_{t_2}t_3}$$

$$(29)$$

The constant exponential terms can be moved outside of the bracket; we then perform the Taylor expansion of the exponentials within the $\langle \cdots \rangle$ symbol, retaining only second-order terms in t_3, which is justified when the separation of time scales between the fast lifetime slow vibrations is invoked. $\left\langle e^{+i(\hat{H}_e-\langle \hat{H}_e\rangle_{t_2})t_3}e^{-i(\hat{H}_c-\langle \hat{H}_c\rangle_{t_2})t_3}\right\rangle$ is then replaced by

$$\left\langle \left[1 + i(\hat{H}_e - \langle \hat{H}_e \rangle_{t_2})t_3 - \frac{1}{2}(\hat{H}_e - \langle \hat{H}_e \rangle_{t_2})^2 t_3^2 \right]\left[1 - i(\hat{H}_c - \langle \hat{H}_c \rangle_{t_2})t_3 - \frac{1}{2}(\hat{H}_c - \langle \hat{H}_c \rangle_{t_2})^2 t_3^2 \right] \right\rangle$$

$$= 1 + i\left[\langle (\hat{H}_e - \langle \hat{H}_e \rangle_{t_2}) \rangle - \langle (\hat{H}_c - \langle \hat{H}_c \rangle_{t_2}) \rangle \right]t_3 + \langle (\hat{H}_e - \langle \hat{H}_e \rangle_{t_2})(\hat{H}_c - \langle \hat{H}_c \rangle_{t_2}) \rangle t_3^2$$

$$- \frac{1}{2}\langle (\hat{H}_e - \langle \hat{H}_e \rangle_{t_2})^2 \rangle t_3^2 - \frac{1}{2}\langle (\hat{H}_c - \langle \hat{H}_c \rangle_{t_2})^2 \rangle t_3^2$$

$$= 1 - \frac{1}{2}\left[\langle (\hat{H}_e - \langle \hat{H}_e \rangle_{t_2})^2 \rangle + \langle (\hat{H}_c - \langle \hat{H}_c \rangle_{t_2})^2 \rangle - 2\langle (\hat{H}_e - \langle \hat{H}_e \rangle_{t_2})(\hat{H}_c - \langle \hat{H}_c \rangle_{t_2}) \rangle \right]t_3^2 \tag{30}$$

where in the derivation we have employed the fact that $\langle \overline{\chi}_e^\mu(t_2)|\overline{\chi}_e^\mu(t_2) \rangle = 1 \forall t_2$. This is not the case for the coherence contribution. One would be tempted to rewrite the term in front of t_3^2 as the matrix element of the $e-c$ energy-gap variance computed on the t_2 evolved WP on both e and e' PESs, which is given by

$$\sigma_{ec,m}^2(t_2) = \left\langle \left[(\hat{H}_e - \hat{H}_c) - (\langle \hat{H}_e \rangle_{t_2} - \langle \hat{H}_c \rangle_{t_2}) \right]^2 \right\rangle$$

$$= \left\langle \left[(\hat{H}_e - \langle \hat{H}_e \rangle_{t_2}) - (\hat{H}_c - \langle \hat{H}_c \rangle_{t_2}) \right]^2 \right\rangle$$

$$= \langle (\hat{H}_e - \langle \hat{H}_e \rangle_{t_2})^2 \rangle + \langle (\hat{H}_c - \langle \hat{H}_c \rangle_{t_2})^2 \rangle$$

$$- \langle (\hat{H}_e - \langle \hat{H}_e \rangle_{t_2})(\hat{H}_c - \langle \hat{H}_c \rangle_{t_2}) \rangle$$

$$- \langle (\hat{H}_c - \langle \hat{H}_c \rangle_{t_2})(\hat{H}_e - \langle \hat{H}_e \rangle_{t_2}) \rangle \tag{31}$$

Nonetheless, this is not the same result reported in eq 30, as $[\hat{H}_e, \hat{H}_c] \neq 0$, i.e., the two operators do not commute. Therefore, the t_3^2 term in square brackets in eq 30 does not simplify to the energy gap variance, and we will refer to it as the *pseudovariance* $\tilde{\sigma}_{ec,m}^2(t_2)$ (where the t_2 dependence is made explicit), given by

$$\tilde{\sigma}_{ec,m}^2(t_2) = \sigma_{ec,m}^2(t_2) - \langle [\hat{H}_{e,m}, \hat{H}_{c,m}] \rangle_{t_2}$$

$$= \sigma_{ec,m}^2(t_2) + i2\Xi_{ec,m}(t_2) \tag{32}$$

The commutator term of the pseudovariance turns out to be a purely imaginary t_2 dependent term, here called $i2\Xi_{ec,m}(t_2)$, which we demonstrate to be the expectation value of the nuclear momentum along the m-th mode (see Section S4 of the Supporting Information for a derivation of such term).

Following the same step performed for the linear response $\left(1 - \frac{1}{2}\tilde{\sigma}_{ec,m}^2 t^2 \sim e^{-\frac{1}{2}\tilde{\sigma}_{ec,m}^2 t^2} \right)$ we get

$$\langle \overline{\chi}_{e,m}^\mu(t_2)|e^{+i\hat{H}_e t_3}e^{-i\hat{H}_c t_3}|\overline{\chi}_{e,m}^\mu(t_2) \rangle \sim e^{-i(\langle \hat{H}_c \rangle_{t_2} - \langle \hat{H}_e \rangle_{t_2})t_3}e^{-\frac{1}{2}\tilde{\sigma}_{ec,m}^2(t_2)t_3^2}$$

$$= e^{-i(\langle \hat{H}_c \rangle_{t_2} - \langle \hat{H}_e \rangle_{t_2})t_3}e^{-i\Xi_{ec,m}(t_2)t_3^2}e^{-\frac{1}{2}\sigma_{ec,m}^2(t_2)t_3^2} \tag{33}$$

The population term of the ESA response function in the EGVA approximation, taking into account all N modes, therefore reads

$$R_{EGVA}^{(3)ESA}(t_3, t_2) = -\mathcal{N}_\mu \sum_{e,c} \boldsymbol{\mu}_{ec}\boldsymbol{\mu}_{ce}\, e^{-i(\omega_{ec}(t_2) + \Xi_{ec}(t_2)t_3)t_3 - \frac{1}{2}\varsigma_{ec}^2(t_2)t_3^2 - t_3/2\tau_c} \tag{34}$$

In the last expression, the relevant terms to be computed are the t_2-dependent energy-gap frequency $\omega_{ec}(t_2) = \omega_{ec}^{(ad)} + (\langle \hat{H}_c \rangle_{t_2} - \langle \hat{H}_e \rangle_{t_2})$, the variances $\varsigma_{ec}^2(t_2)$, which are the sums of all the (t_2 dependent) ec-energy-gap variances along the N modes, and the $\Xi_{ec}(t_2) = \sum_m \Xi_{ec,m}(t_2)$ term, that acts as a linear chirp term along t_3. Notice that

$$\langle \hat{H}_c \rangle_{t_2} - \langle \hat{H}_e \rangle_{t_2} = \langle \hat{H}_c - \hat{H}_e \rangle_{t_2}$$

$$= \langle \hat{V}_c - \hat{V}_e \rangle_{t_2}$$

$$= \sum_m \lambda_{c,m} - \lambda_{e,m} - \omega_m \langle \hat{Q}_m \rangle_{e,t_2}(d_{c,m} - d_{e,m})$$

$$= V_c(\langle Q \rangle_{e,t_2}) - V_e(\langle Q \rangle_{e,t_2}) \tag{35}$$

implying that the energy-gap expectation value can be evaluated as the difference of the two potential energies computed at $\langle \hat{Q}_m \rangle_{e,t_2}$, where $\langle \hat{Q}_m \rangle_{e,t_2} = \langle \chi_{e,m}(t_2)|\hat{Q}_m|\chi_{e,m}(t_2) \rangle$ is the centroid of the WP on state e along mode m. Similarly, the t_2-dependent energy-gap variance can be recast in terms of the $\hat{Q}$ and $\hat{Q}^2$ matrix elements of the WP on state e, i.e.,

$$\left\langle (\hat{H}_c - \hat{H}_e - \langle \hat{H}_c - \hat{H}_e \rangle_{t_2})^2 \right\rangle_{t_2}$$

$$= \sum_m \left\langle (\omega_m(d_{c,m} - d_{e,m})\hat{Q}_m - \omega_m(d_{c,m} - d_{e,m})\langle \hat{Q}_m \rangle_{e,t_2})^2 \right\rangle_{t_2}$$

$$= \sum_m \omega_m^2(d_{c,m} - d_{e,m})^2(\langle \hat{Q}_m^2 \rangle_{e,t_2} - \langle \hat{Q}_m \rangle_{e,t_2}^2) \tag{36}$$

and the WP coordinate variance now shows up.

Moreover, we have

$$\Xi_{ec}(t_2) = \frac{1}{2}\sum_m \omega_m^2(d_{e,m} - d_{c,m})\langle \hat{P}_m \rangle_{e,t_2} \tag{37}$$

where $\langle \hat{P}_m \rangle_{e,t_2}$ is the expectation value of the momentum operator along mode m (see Section S4 of the Supporting Information). The complete pseudovariance expression therefore reads

$$\tilde{\varsigma}_{ec}^2(t_2) = \varsigma_{ec}^2(t_2) + 2i\Xi_{ec} =$$

$$= \sum_m \omega_m^2(d_{e,m} - d_{c,m})[(d_{e,m} - d_{c,m})(\langle \hat{Q}_m^2 \rangle_{e,t_2}$$

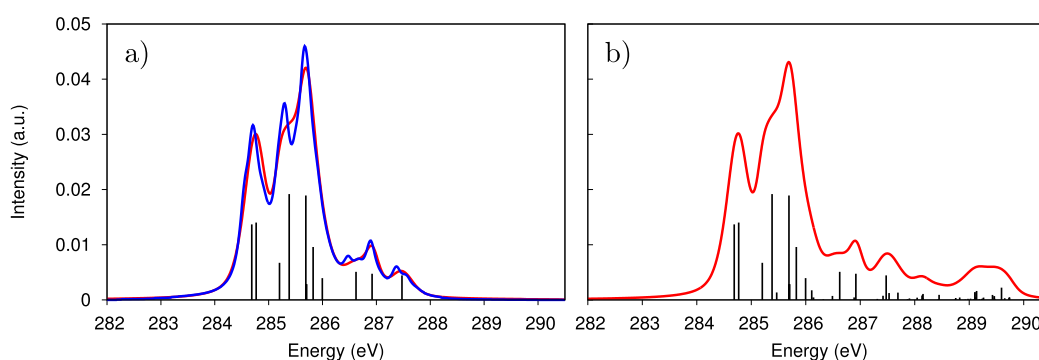$$- \langle \hat{Q}_m \rangle_{e,t_2}^2) + i\langle \hat{P}_m \rangle_{e,t_2}] \tag{38}$$

Figure 4. (a) XAS spectrum for the 11 brightest $g \rightarrow c$ transitions: EGVA spectrum (red line) and exact spectrum (blue line). (b) Total XAS spectrum for the full set of 75 $g \rightarrow c$ transitions, computed only at the EGVA level. In both plots, the stick spectrum is also reported (black bars), and τ_c is set to 3 fs.

A clear physical interpretation of the obtained quantities can be drawn: $V_c(\langle \hat{Q} \rangle_{e,t_2}) - V_e(\langle \hat{Q} \rangle_{e,t_2})$ is the vibrational contribution to the energy-gap between the c and e PESs evaluated at the WP centroid and $\omega_{ec}(t_2)$ gives simply the vertical energy-gap frequency computed for each time t_2 (at the WP centroid). As prescribed by eq 36, the energy-gap variance can be obtained through the WP variance along each nuclear coordinate Q_m.

To summarize, the EGVA expression of the third-order response, reported in eq 34, allows all quantities to be obtained from a QD only along the t_2 time (i.e., without running QD and WP overlap evaluation along t_3).

3. RESULTS AND DISCUSSION

In this section, we compare linear and nonlinear spectroscopy simulations of the pyrene molecule in the X-ray window, obtained by employing the LVC/WPO* and LVC/EGVA approaches. Comparison with LVC/WPO is also considered to assess the impact of neglecting the $\mathcal{E}$ nonadiabatic dynamics in the short-limit time-scale dictated by τ_c.

An LVC model Hamiltonian of pyrene, parametrized via ab initio CASSCF/CASPT2 inputs, was employed as a benchmark. The lowest 7 excited singlet states were considered in the $\mathcal{E}$ manifold, accounting for the dipole moments coupling to the GS, and the nonadiabatic coupling between them. We have considered the internal conversion processes taking place after photoexcitation to the first bright state (S_2), that mainly involve the transition to the lower dark S_1 state. The other electronic states serve as mediators of the energy-transfer process, as shown by our previous studies.[31] Our model was used to run quantum dynamics on pyrene along several nuclear coordinates with the MCTDH method.[1,31] In order to speed ease the computation of the WP overlaps along the simulations, we used MCTDH and restricted the number of nuclear DOFs to 15 out of a total of 49 that were included in the original ML-MCTDH dynamics,[31] see ref 1 for details on DOF selection.[i]

First, we report, for a restricted number of transitions (i.e., the brightest ones) the comparison of WPO* and EGVA XAS spectra (obtained as the Fourier transform of the first-order response functions; see expressions reported in eq 39). The total XAS spectrum (that accounts for contributions from all 75 $g \rightarrow c$ transitions, with $c \in C$) is also reported at the EGVA level. We then discuss the results and the appropriateness of the EGVA approximation. In particular, we assessed the quality of the XAS spectra in the EGVA limit for different τ_c's, and for

high- and low-frequency vibrational modes (this last study is reported in Section S6 of the Supporting Information). One expects the EGVA approach to be optimal for short lifetimes (for which the second-order truncation of the overlap Taylor expansion is justified) and for low vibrational frequencies (the lower the mode frequency, the longer the mode period, the better justified the time-scale separation between the lifetime and vibrational motion of the WP).
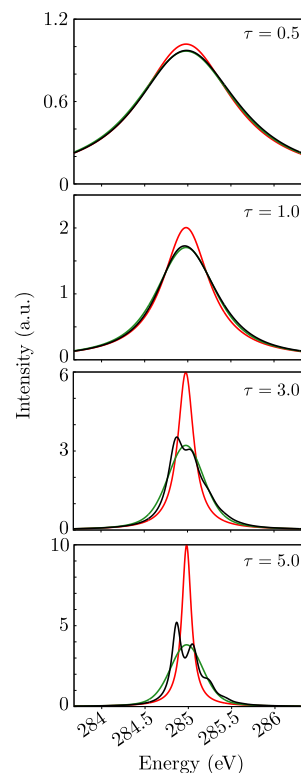

Figure 5. Comparison of carbon K-edge XAS spectra obtained via WPO* (black line), EGVA approximation (green line), and EGA, which simply sets the WP overlap to 1 (red line) and thus gives a Lorentzian line width to the considered transition. Different lifetime values τ_c (in fs) are considered: as expected, the shorter the lifetime, the higher the quality of the EGVA approximation. Accounting for the overlap via the energy-gap variance greatly improves the spectrum with respect to just setting the overlap term to 1 (EGA). For the considered transition, the vertical energy gap is 284.7 eV, and the energy-gap variance is $\varsigma_{gc}^2 = 0.026$ eV^2.
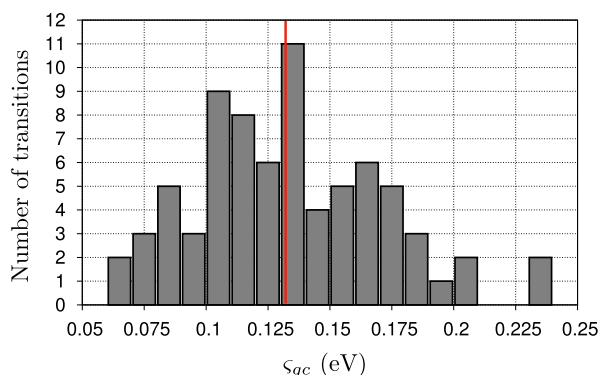
Figure 6. Histogram of ς_{gc} values for the 75 $g \rightarrow C$ transitions that builds the XAS spectrum. The binning size was set to 0.01 eV; the vertical red bar indicates the average ς_{gc} value of about 0.13 eV.

Figure 7. Intensities from products of dipoles for various transient absorption contributions ($\boldsymbol{\mu}_{ac}\boldsymbol{\mu}_{ca'}$, with a and a' being equal to g, S_1 and/or S_2). GSB (core excitations from g) are depicted in violet; ESA (population) contributions from S_1 are depicted in green; ESA (population) contributions from S_2 (less intense than the former) are depicted in cyan; ESA (coherence) contributions that involve both S_1 and S_2 are depicted in orange. Note that the spectra are dominated by population contributions: coherence ESA contributions are therefore neglected. The symbols *, ○, and § denote the transitions (from g, S_1, and S_2, respectively) that have been selected for the evaluation of the spectra via the overlap method.

We then compare the WPO* and EGVA nonlinear (transient absorption) spectra (by Fourier transformation along the time t_3 of the third-order responses reported in eq 40). Note that no coherence ESA contributions need to be computed for pyrene, which simplifies the calculations. Two levels of theory could be considered: the complete EGVA approach, which accounts for the pseudovariance in the response function expressions and for which one needs to evaluate the purely imaginary term $i2\Xi_{ec}(t_2)$ along the t_2 quantum dynamics, and the reduced rEGVA approach, which sets to zero the commutator term and thus only requires computing the (standard) energy-gap variance along the t_2 quantum dynamics. We also examined the accuracy of the approximation of neglecting population transfer in the $\mathcal{E}$ manifold along t_3, i.e., when the probe creates a $\mathcal{E} - C$ coherence, by comparing the WPO and WPO* TA spectra.

3.1. Linear Absorption XAS Spectrum. We first rewrite the two expressions for the first-order response function, i.e., WPO* and EGVA. While we already highlighted that, in the approximation of neglecting nonadiabatic dynamics, an analytical expression for $R^{(1)}(t)$ can be written,[50] here we formulate the linear WPO* response function in terms of the time-dependent WP overlap to remain consistent with the definition of the nonlinear response discussed in the next section. Thus, we have

$$R_{\text{WPO*}}^{(1)}(t) = \left(\frac{i}{\hbar}\right)\theta(t)\sum_c \boldsymbol{\mu}_{gc}\boldsymbol{\mu}_{cg}\, e^{-i\omega_{gc}^{(\text{ad})}t - t/2\tau_c}\langle\chi_g(t)|\chi_{c^*}(t)\rangle$$

$$R_{\text{EGVA}}^{(1)}(t) = \left(\frac{i}{\hbar}\right)\theta(t)\sum_c \boldsymbol{\mu}_{gc}\boldsymbol{\mu}_{cg}\, e^{-i\omega_{gc}t - \frac{1}{2}\varsigma_{gc}^2 t^2 - t/2\tau_c}$$

$$(39)$$

3.1.1. WPO versus EGVA Linear Absorption.* Figure 4a compares the XAS spectrum computed at both levels of theory for the 11 brightest transitions (vertical energy gaps, TDMs, and variances of these transitions are reported in Section S9 of the Supporting Information). The core-excited-state lifetime was set to $\tau_c = 3$ fs for all states (while we also report results with longer lifetimes, $\tau_c = 5$ and $\tau_c = 7$ fs, in Section S7 of the Supporting Information).

The WPO* spectrum is quite accurately reproduced by the EGVA approach with minor differences that can be ascribed to the inability of the EGVA to describe the band asymmetry and possible vibronic side bands.
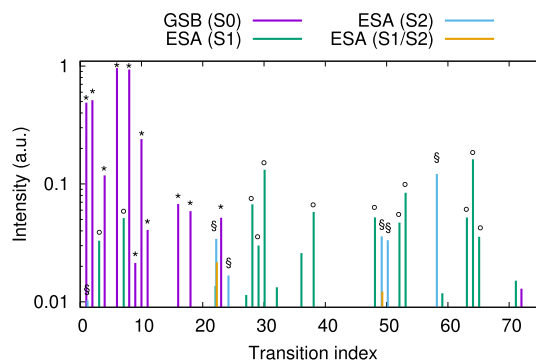
The total spectrum (comprising 75 transitions) is reported in Figure 4b. Note that, in the high-energy region of the spectrum, a large number of low-intensity transitions accumulate and form an additional shoulder centered at 289.5 eV. We should also note that experimental XAS spectra (not yet available) would also include transitions to quasi/bound and unbound states in the continuum, resulting in a broad and unstructured contribution to the spectrum rising around the carbon K-edge ionization potential. These are not accounted for here.

3.1.2. WPO versus EGVA Linear Spectra at Various τ_c's.* Figure 5 compares the WPO*, the EGVA, and the EGA spectra for the $g \rightarrow c_1$ transition, for different τ_c's. The EGVA spectrum has a Voigt line shape (Fourier transform of $R_{\text{EGVA}}^{(1)}(t)$), while the EGA one possesses a Lorentzian profile (which is what is obtained when setting the WP overlap term to 1, or, equivalently, when truncating the Taylor expansion to the first order in t). We note that the EGA/Lorentzian spectrum is always narrower than the other two; the EGVA/Voigt spectrum is much more accurate in capturing the band as a whole and compares well with the WPO* spectrum: indeed, WPO* and EGVA spectra have the same variance. Note also that, as expected, both EGA and EGVA approaches miss the asymmetry of the WPO* band, as both the Lorentzian and Voigt line shapes are symmetric. It is interesting to note how even for the very short lifetimes expected for X-rays, including the variance in the spectrum, greatly improves the line shape with respect to the EGA approach, which means that vibrational broadening is of the same order or larger than the (homogeneous) lifetime broadening already at $\tau_c \sim 2$ or 3 fs.

The WPO*/EGVA linear spectra for high- and low-frequency vibrational modes are reported in the Supporting Information (Section S6).

3.1.3. EGVA Transition Width. In closing this section, we show the extent of the variability of the standard deviation ς_{gc} (in eV) for the 75 considered $g \rightarrow C$ transitions. In Figure 6, we reported the histogram of the ς_{gc} values, employing a binning size of 0.01 eV. The average ς_{gc} is at about 0.13 eV with
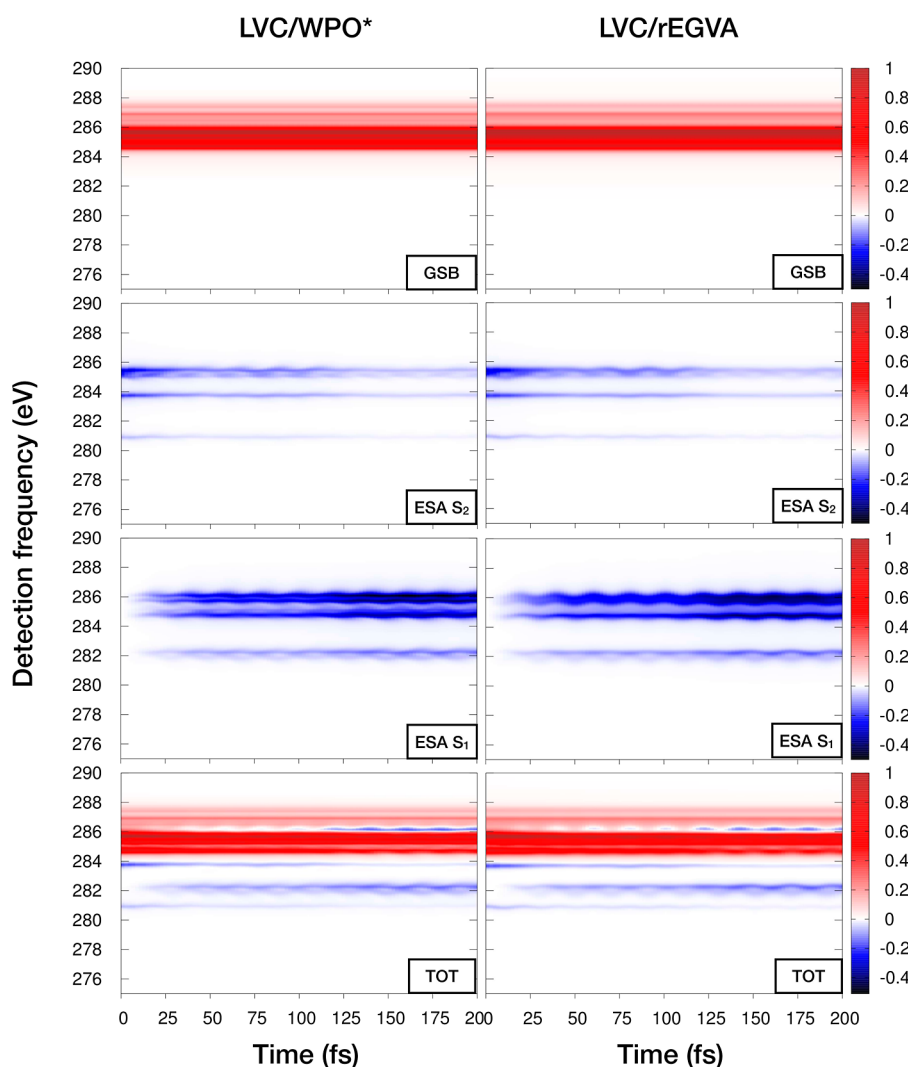
Figure 8. Comparison of WPO* (*left*) and rEGVA (*right*) TA spectra for the *brightest* $\mathcal{E} \rightarrow C$ transitions. The total spectrum and various contributions (GSB, ESA from S_1, and ESA from S_2) are shown. The core-excited-state lifetime is set to $\tau_c = 3$ fs.

variations between 0.06 and 0.23 eV (which, would it be the only contribution to the broadening would correspond to a band full width at half-maximum of about 0.14 and 0.54 eV, respectively). We note that assigning the same phenomenological broadening to all the transitions would deteriorate the spectrum shape, especially when states of very different nature are present in the spectrum (as, e.g., 1s → π^* and 1s → σ^* transitions). Figure S11 of the Supporting Information shows the comparison of total XAS obtained employing the calculated ς_{gc}^2 values and a constant ς_{gc}^2 value (set to the average of all the calculated variances). As expected, the spectrum obtained by employing a constant variance value tends to smooth out some spectral features. This is particularly relevant in TA, where ESA signals may arise from a few relatively isolated bright transitions, therefore showing significant differences if computed with their own variance or a constant phenomenological value.

3.2. Transient Absorption. In the previous section, we tested the quality of the EGVA approximation for the static XAS spectrum. We now turn to transient spectroscopy: having the possibility of avoiding the t_3 propagation of the WP in the $\mathcal{E} + C$ manifolds and to compute the t_3-dependent overlap, while still accounting for the nonadiabatic dynamics during t_2

evolution of the wave function at the quantum dynamics level, is the main goal of our approach which exploits the EGVA approximation.

For pyrene transient absorption (TR-XAS), we consider the $S_2 \rightarrow S_1$ internal conversion process: the pump is assumed to be resonant only to the bright S_2 state (and the system initial state—after the pump interaction—is $|\Psi(0)\rangle = \mu_{S_{2g}}^{\epsilon}|\chi_g(0)\rangle|S_2\rangle$).

Moreover, pyrene has no core-excited states c that are simultaneously dipole-coupled to both S_2 and S_1 (see Figure 7): therefore, the coherence term of eq 28 can be dropped which results in the following working equations

$$R_{\text{WPO*}}^{(3)}(t_3, t_2) \propto + \sum_c \mu_{gc}\mu_{cg} e^{-i\omega_{gc}^{(\text{ad})}t_3 - t_3/2\tau_c}\langle \chi_g^\mu(t_3)|\chi_c^\mu(t_3)\rangle +$$

$$- \sum_{e,c} \mu_{ec}\mu_{ce} e^{-i\omega_{ec}^{(\text{ad})}t_3 - t_3/2\tau_c}\langle \chi_e^\mu(t_2)|e^{i\hat{H}_e t_3}e^{-i\hat{H}_c t_3}|\chi_e^\mu(t_2)\rangle$$

$$R_{\text{rEGVA}}^{(3)}(t_3, t_2) \propto + \sum_c \mu_{gc}\mu_{cg} e^{-i\omega_{gc}t_3 - \frac{1}{2}\varsigma_{gc}^2 t_3^2 - t_3/2\tau_c} +$$

$$- \sum_{e,c} \mu_{ec}\mu_{ce}\rho_e(t_2) e^{-i\omega_{ec}(t_2)t_3 - \frac{1}{2}\varsigma_{ec}^2(t_2)t_3^2 - t_3/2\tau_c}$$
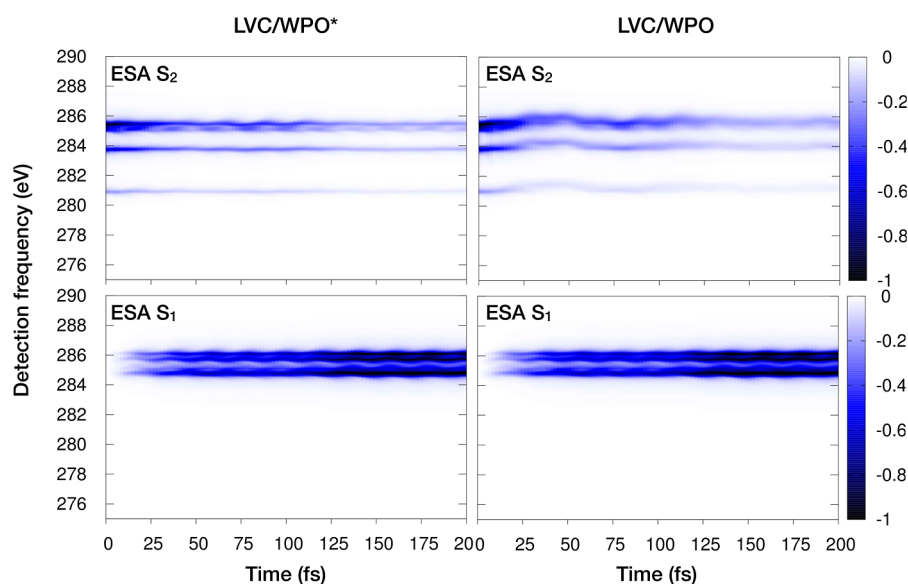
$$(40)$$

Figure 9. Comparison of selected WPO* (*left*) and WPO (*right*) ESA contributions in the TA. In WPO*, population transfer is assumed to occur on longer time scales than the core-excited-state lifetime (here set to $\tau = 3$ fs) and therefore the nonadiabatic dynamics in the $\mathcal{E}$ manifold is switched off along the t_3 time. The WPO level of theory considers population transfer in the $\mathcal{E}$ at all times. Only slight deviations are noted in the S_2 ESA signals (with WPO signals generally broader than the WPO* ones and with a slight blue shift of the former signal with respect to the latter level of theory), while an extremely accurate reproduction is shown for S_1 ESA.

In both formulations, the first term is the GSB term (which again in the WPO* formulation can be evaluated analytically) and the second is the ESA term. Moreover, for simplicity, we have removed the common $(i/\hbar)^3\theta(t_2)\theta(t_3)$ factor, as well as the common $\mathcal{N}_\mu$ normalization factor.

To assess the validity of the approximation, we compare results for the WPO and WPO* expressions, which explicitly compute the overlap, with the equations obtained employing the energy-gap/energy-gap variance approximation without computing the commutator term, a level of theory that we refer to as *reduced* EGVA or rEGVA. The states were selected by a brightness criterion. A total of 28 core-excited states have been considered: 11 dipole coupled to g, 12 dipole coupled to S_1, and 6 dipole coupled to S_2 vertical energy gaps, and TDMs are given in the Supporting Information (Section S9).

Figure 8 shows the WPO* and rEGVA TA spectra. The agreement is good, even though the standard variance has been used instead of the prescribed pseudovariance. This suggests that in the present system neglecting the purely imaginary term $i\Xi_{ec}(t_2)$ does not significantly affect the results. Note that the energy-gap variation along t_2 captures the oscillatory behavior of the signals along the detection frequency axis at different t_2 times. The GSB signal covers almost half of the considered window. Nonetheless, the spectrum shows a background-free region around 280−283 eV in which the dying out of the $S_2 \to c$ signals and the concomitant growing of the $S_1 \to c$ ones are clearly observed; therefore, these signals can be used to follow the photoinduced dynamics on the $\mathcal{E}$ manifold.

To assess the impact of neglecting nonadiabatic dynamics along t_3, in Figure 9, we compare WPO and WPO* TA spectra. The main difference between the two approach is that in WPO one does not switch off the nonadiabatic coupling in the $\mathcal{E}$ manifold along t_3. Interestingly, as noted in ref 1, the nonadiabatic dynamics along t_3 affects the shape of the WP on the S_2 surface, so that ESA signals from S_2 experience a slight blue shift and a clear enhancement of the band broadening. At

variance, the WPO S_1 ESA signals are completely reproduced at the WPO* level.

To get additional insight, in Figure 10, we study cuts of the WPO, WPO* and rEGVA TA spectra for two transitions (namely, $S_2 \to c_{30}$, which appears on the low-energy side of the spectrum, and $S_1 \to c_1$, on the high-energy side of the spectrum) at selected t_2 times: $t_2 = 0$ fs (where all the population is on S_2), $t_2 = 25$ and 100 fs (where the population is split between S_2 and S_1), and $t_2 = 200$ fs (where most of the population eventually accumulates on S_1). A few relevant observations are as follows: as expected, the $S_2 \to c_{30}$ ESA diminishes with time, while the $S_1 \to c_1$ ESA grows, which mirrors the population dynamics; WPO and WPO* can assume an asymmetric line shape with vibronic side bands, while rEGVA (as well as EGVA) miss these finer details; finally, the impact of the nonadiabatic dynamics along t_3 is clearly assessed by comparing WPO and WPO* cuts of the $S_2 \to c_{30}$ ESA signal which is considerably broader at the WPO level.

Figure 11 shows the rEGVA TA map computed by including all of the transitions obtained from quantum chemistry calculations (and not just the brightest ones). The spectrum closely resembles the total spectrum in Figure 8.

To conclude, we note that the (reduced) 15-mode model considered here, which was fine-tuned to describe the dynamics in the $\mathcal{E}$ manifold, might have left-out modes that are relevant to describe the spectroscopy of the core-excited states, such as strongly displaced modes along some of the C states. Therefore, spectroscopy simulations of a full-mode model might show some differences, with the left-out modes possibly contributing to the additional broadening/vibronic structure of the transitions. This is clearly illustrated in Section S10 of the Supporting Information, where we compare the $g \to c_1$ linear absorption spectrum in the reduced (15 modes) and the full-mode model. To overcome this issue, we mention that a strategy that mixes numerical QD propagation along a restricted number of modes with analytical evaluation of the
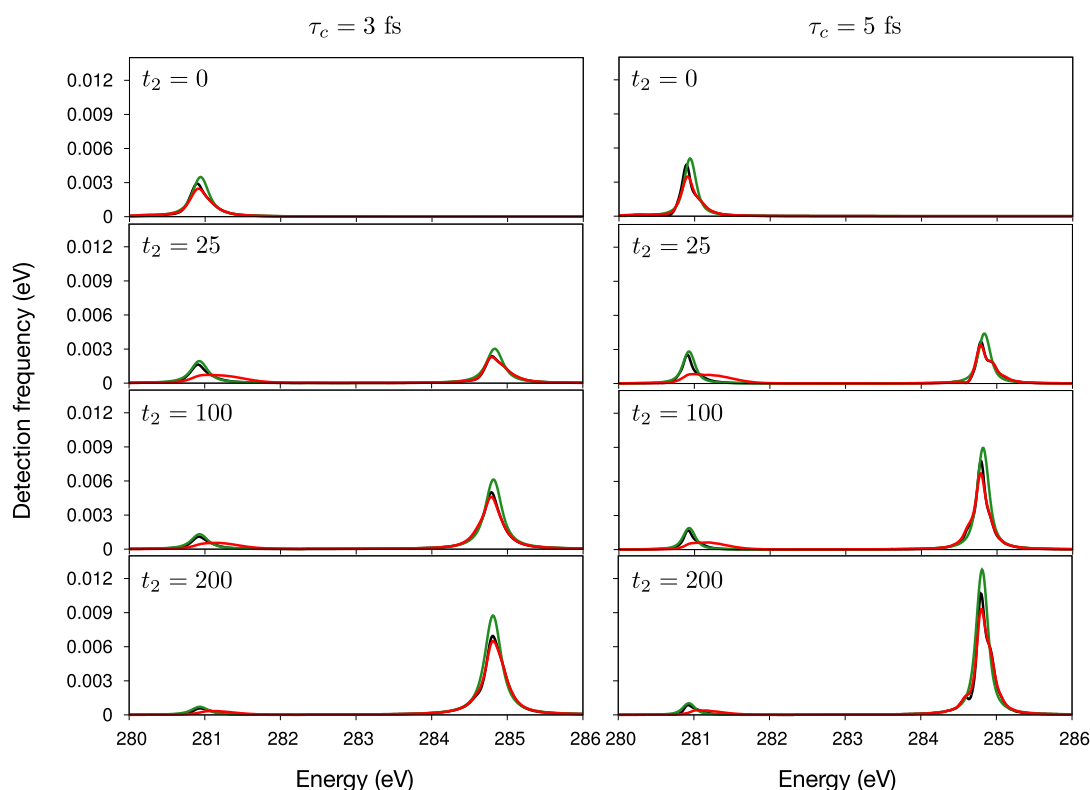
Figure 10. Comparison of TA cuts at selected t_2 times and for different values of τ_c. Only two transitions (ESA) are considered: $S_2 \rightarrow c_{30}$ (on the left-hand side of the spectrum) and $S_1 \rightarrow c_1$ (on the right-hand side of the spectrum). WPO, WPO*, and rEGVA levels of theory are depicted with red, black, and green curves, respectively. Note that we reported the ESA signal multiplied by -1.

remaining (*spectroscopically relevant*) modes has been recently reported in ref 58.

4. CONCLUSIONS

An efficient approach to simulate nonlinear X-ray spectroscopy has been demonstrated. Approximate expressions that exploit the accuracy of quantum dynamics along the interpulse waiting time t_2 while taking advantage of the extremely short lifetime of core-excited/ionized states along the detection time t_3 were derived by means of Taylor expanding the WP propagator in time and truncating at the second order. The approximation is physically motivated by the time-scale separation between states' lifetime τ_c which is assumed to be shorter than any other relevant dynamics time scale (i.e., that of electronic population dynamics and that of nuclear vibrational motion). This allows omission of explicit propagation of the WP on both the valence- and core-excited states manifold along t_3 as done in the exact protocol termed WPO. The obtained expressions require computing a few relevant quantities only along the t_2 dynamics in the manifold of valence-excited states: the energy gap at the centroid of the evolving WP and the energy-gap pseudovariance (EGVA) or standard variance (rEGVA, or reduced EGVA). Our protocol positions itself between the EGA, which neglects the energy-gap variance, and the exact WPO protocols, as it incorporates a higher degree of physical detail than EGA without significant additional computational demand (essentially requiring a QD simulation only in the manifold of valence states) but is unable to recover finer details such as asymmetrical broadening and vibronic structures seen in WPO simulations which can be only recovered by considering terms beyond the quadratic in the Taylor expansion, with an increase of the computational cost.

We presented XAS and TR-XAS for the ultrafast $S_2 \rightarrow S_1$ internal conversion in pyrene at both the LVC/WPO and approximate LVC/rEGVA levels of theory, showing excellent agreement that validates the approximation underlying the rEGVA. The limits of the time-scale separation and, thus the validity of the rEVGA, were explored by analyzing the dependence of the spectral line shapes on the core-excited state lifetimes τ_c, as well as on the mode frequencies. Following the validation of the approximation, complete TR-XAS at the carbon K-edge was simulated at the LVC/rEGVA level, thereby incorporating into the Hamiltonian nearly 100 core-excited states. The spectra reveal signatures of the S_2 disappearing with t_2 accompanied by emerging signatures of the S_1 state accumulating the excited-state population in a background free spectral window between 280 and 284 eV.
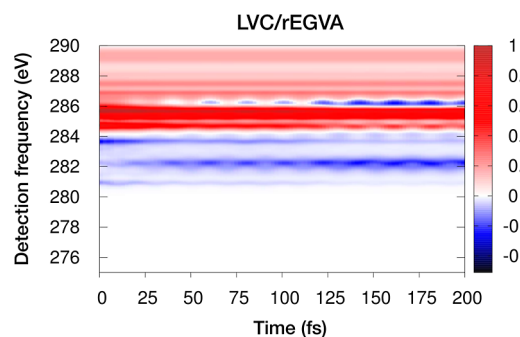

Figure 11. TA spectrum obtained at the rEGVA level, considering all the transitions from the valence-excited states (g, S_1, and S_2) to all of the core-excited states (c_1–c_{75}). Only population contributions have been computed. $\tau_c = 3$ fs.

The signals are modulated by oscillatory features due to the coherent vibrational motion in the excited states. Despite not having an experimental counterpart, we are confident that the simulated spectra will agree with future TR-XAS experiments, which will be the ultimate validation of separating the time scales on which core-excited states decay and nuclei move for the purpose of accelerating simulations in the X-ray regime.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.3c00953.

> Active spaces for core-excited states; different approaches to simulate the LA spectrum; line shape functions second-order Taylor expansion; the commutator $[\hat{H}_e, \hat{H}_c]$; the coherence $R_{\mathrm{WPO}*}^{(3)\mathrm{ESA}}(t_3,t_2)$ term in the EGVA approximation; WPO* vs EGVA LA: high- and low-frequency modes; WPO* vs EGVA LA at various values of τ_c; EGVA XAS spectrum: variable vs constant variance; energy gaps, energy-gap variances, and TDMs; pyrene LA—full-mode models vs reduced-mode model; and response function Fourier transform—technical details (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Marco Garavelli − *Dipartimento di Chimica Industriale "Toso Montanari", University of Bologna, 40136 Bologna, Italy;* ⓞ orcid.org/0000-0002-0796-289X; Email: marco.garavelli@unibo.it

Fabrizio Santoro − *Istituto di Chimica dei Composti Organometallici (ICCOM-CNR), I-56124 Pisa, Italy;* ⓞ orcid.org/0000-0003-4402-2685; Email: fabrizio.santoro@pi.iccom.cnr.it

Artur Nenov − *Dipartimento di Chimica Industriale "Toso Montanari", University of Bologna, 40136 Bologna, Italy;* ⓞ orcid.org/0000-0003-3071-5341; Email: artur.nenov@unibo.it

Authors

Francesco Segatta − *Dipartimento di Chimica Industriale "Toso Montanari", University of Bologna, 40136 Bologna, Italy;* ⓞ orcid.org/0000-0003-4150-6676

Daniel Aranda − *ICMol, Universidad de Valencia, 46980 Paterna, Spain; Istituto di Chimica dei Composti Organometallici (ICCOM-CNR), I-56124 Pisa, Italy;* ⓞ orcid.org/0000-0003-0747-6266

Flavia Aleotti − *Dipartimento di Chimica Industriale "Toso Montanari", University of Bologna, 40136 Bologna, Italy;* ⓞ orcid.org/0000-0002-7176-5305

Francesco Montorsi − *Dipartimento di Chimica Industriale "Toso Montanari", University of Bologna, 40136 Bologna, Italy*

Shaul Mukamel − *Department of Chemistry and Department of Physics and Astronomy, University of California, Irvine 92697 California, United States;* ⓞ orcid.org/0000-0002-6015-3135

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.3c00953

Notes

The authors declare no competing financial interest.

■ ADDITIONAL NOTES

[a]In dimensionless coordinates, it reads $\hat{T}_m = \frac{1}{2}\omega_m \hat{P}_m^2$, with $\hat{\boldsymbol{P}}_m$ being the momentum operator along the mth coordinate.

[b]Note that such $d_{a,m}$ displacements can be obtained either at the Franck–Condon geometry (simply computing the a-th state gradient, which, along the given mode m, is given by $-\omega_m d_{a,m}$) or via an excited-state geometry optimization. Here, we employ the (computationally cheaper) first approach. The two results are equivalent only when the surfaces are perfectly harmonic and obey the DHO model.

[c]If, on the time scale of interest, the process that brings an excited electronic state back to the ground electronic state is relevant, one may add a term in the Hamiltonian that couples these two manifolds.

[d]Note that one could in principle consider different τ_c values for dissociative and nondissociative core-excited states.

[e]Note that the WPO* first-order response function of eq 10 can be evaluated analytically by means of line shape functions.[1]

[f]Note that the most general formulation would also consider the $c \rightarrow c'$ transfer process. We still assume that accounting for such a process is immaterial as, in the dense manifold of core-excited states, the nonadiabatic dynamics is fast, and its effect can be accounted for by further shortening the lifetime term τ_c.

[g]Note that such a procedure is correct only when $\langle \chi_e^u | \chi_e^u \rangle \neq 0$. At early times, for dark states, this might not be the case.

[h]Note that we simplify the notation: it is understood that the $\langle \cdots \rangle$ symbol refers to that specific matrix element; using a similar but more elaborate notation, one could have specified that by writing $\langle \hat{H}_a \rangle_{bc,m}(t_2)$, with $a = b$ or c. Since this is clear from the context, we will use the simpler notation.

[i]Even though the total number of pyrene normal modes is 72, for symmetry reasons, 23 of them have zero coupling and gradient, so that only 49 are required.

■ REFERENCES

(1) Segatta, F.; Ruiz, D. A.; Aleotti, F.; Yaghoubi, M.; Mukamel, S.; Garavelli, M.; Santoro, F.; Nenov, A. Nonlinear Molecular Electronic Spectroscopy via MCTDH Quantum Dynamics: From Exact to Approximate Expressions. *J. Chem. Theory Comput.* **2023**, *19*, 2075−2091.

(2) Paz, Y. Transient IR spectroscopy as a tool for studying photocatalytic materials. *J. Phys.: Condens. Matter* **2019**, *31*, 503004.

(3) Maiuri, M.; Garavelli, M.; Cerullo, G. Ultrafast Spectroscopy: State of the Art and Open Challenges. *J. Am. Chem. Soc.* **2020**, *142*, 3−15.

(4) Conti, I.; Cerullo, G.; Nenov, A.; Garavelli, M. Ultrafast Spectroscopy of Photoactive Molecular Systems from First Principles: Where We Stand Today and Where We Are Going. *J. Am. Chem. Soc.* **2020**, *142*, 16117−16139.

(5) Pellegrini, C.; Marinelli, A.; Reiche, S. The physics of x-ray free-electron lasers. *Rev. Mod. Phys.* **2016**, *88*, 015006.

(6) Zhao, Z.; Wang, D.; Gu, Q.; Yin, L.; Gu, M.; Leng, Y.; Liu, B. Status of the SXFEL Facility. *Appl. Sci.* **2017**, *7*, 607.

(7) Boutu, W.; Ducousso, M.; Hergott, J.-F.; Merdji, H. *Optical Technologies for Extreme-Ultraviolet and Soft X-ray Coherent Sources*; Canova, F., Poletto, L., Eds.; Springer: Berlin, Germany, 2008; p 63.

(8) Calegari, F.; Sansone, G.; Stagira, S.; Vozzi, C.; Nisoli, M. Advances in attosecond science. *J. Phys. B: At., Mol. Opt. Phys.* **2016**, *49*, 062001.

(9) Chen, L. X.; Zhang, X.; Shelby, M. L. Recent advances on ultrafast X-ray spectroscopy in the chemical sciences. *Chem. Sci.* **2014**, *5*, 4136−4152.

(10) Milne, C.; Penfold, T.; Chergui, M. Recent experimental and theoretical developments in time-resolved X-ray spectroscopies. *Coord. Chem. Rev.* **2014**, *277−278*, 44−68 Following Chemical Structures using Synchrotron Radiation.

(11) Chergui, M.; Collet, E. Photoinduced Structural Dynamics of Molecular Systems Mapped by Time-Resolved X-ray Methods. *Chem. Rev.* **2017**, *117*, 11025−11065.

(12) Bhattacherjee, A.; Leone, S. R. Ultrafast X-ray Transient Absorption Spectroscopy of Gas-Phase Photochemical Reactions: A New Universal Probe of Photoinduced Molecular Dynamics. *Acc. Chem. Res.* **2018**, *51*, 3203−3211.

(13) Young, L.; Ueda, K.; Gühr, M.; Bucksbaum, P. H.; Simon, M.; Mukamel, S.; Rohringer, N.; Prince, K. C.; Masciovecchio, C.; Meyer, M.; et al. Roadmap of ultrafast x-ray atomic and molecular physics. *J. Phys. B: At., Mol. Opt. Phys.* **2018**, *51*, 032003.

(14) Kraus, P. M.; Zürch, M.; Cushing, S. K.; Neumark, D. M.; Leone, S. R. The ultrafast X-ray spectroscopic revolution in chemical dynamics. *Nat. Rev. Chem* **2018**, *2*, 82−94.

(15) Geneaux, R.; Marroux, H. J. B.; Guggenmos, A.; Neumark, D. M.; Leone, S. R. Transient absorption spectroscopy using high harmonic generation: a review of ultrafast X-ray dynamics in molecules and solids. *Philos. Trans. R. Soc., A* **2019**, *377*, 20170463.

(16) List, N. H.; Dempwolff, A. L.; Dreuw, A.; Norman, P.; Martínez, T. J. Probing competing relaxation pathways in malonaldehyde with transient X-ray absorption spectroscopy. *Chem. Sci.* **2020**, *11*, 4180−4193.

(17) Scutelnic, V.; Tsuru, S.; Pápai, M.; Yang, Z.; Epshtein, M.; Xue, T.; Haugen, E.; Kobayashi, Y.; Krylov, A. I.; Møller, K. B.; Coriani, S.; Leone, S. R. X-ray transient absorption reveals the 1Au (nπ*) state of pyrazine in electronic relaxation. *Nat. Commun.* **2021**, *12*, 5003.

(18) Kaczun, T.; Dempwolff, A. L.; Huang, X.; Gelin, M. F.; Domcke, W.; Dreuw, A. Tuning UV Pump X-ray Probe Spectroscopy on the Nitrogen K Edge Reveals the Radiationless Relaxation of Pyrazine: Ab Initio Simulations Using the Quasiclassical Doorway−Window Approximation. *J. Phys. Chem. Lett.* **2023**, *14*, 5648−5656.

(19) Keefer, D.; Schnappinger, T.; de Vivie-Riedle, R.; Mukamel, S. Visualizing conical intersection passages via vibronic coherence maps generated by stimulated ultrafast X-ray Raman signals. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 24069−24075.

(20) Keefer, D.; Aleotti, F.; Rouxel, J. R.; Segatta, F.; Gu, B.; Nenov, A.; Garavelli, M.; Mukamel, S. Imaging conical intersection dynamics during azobenzene photoisomerization by ultrafast X-ray diffraction. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, 118.

(21) Nam, Y.; Keefer, D.; Nenov, A.; Conti, I.; Aleotti, F.; Segatta, F.; Lee, J. Y.; Garavelli, M.; Mukamel, S. Conical Intersection Passages of Molecules Probed by X-ray Diffraction and Stimulated Raman Spectroscopy. *J. Phys. Chem. Lett.* **2021**, *12*, 12300−12309.

(22) Beck, M. H.; Jäckle, A.; Worth, G. A.; Meyer, H.-D. The multiconfiguration time-dependent Hartree (MCTDH) method: a highly efficient algorithm for propagating wavepackets. *Phys. Rep.* **2000**, *324*, 1−105.

(23) *Multidimensional Quantum Dynamics: MCTDH Theory and Applications*; Meyer, H.-D., Gatti, F., Worth, G. A., Eds.; Wiley VCH: Weinheim, 2009.

(24) Gatti, F.; Lasorne, B.; Meyer, H.-D.; Nauts, A. *Lecture Notes in Chemistry*; Springer International Publishing, 2017; Vol. *98*.

(25) Köppel, H.; Domcke, W.; Cederbaum, L. The Multi-mode vibronic-coupling approach. *Conical Intersections; Electronic Structure, Dynamics and Spectroscopy*; World Scientific, 2004; pp 323−368.

(26) Worth, G. A.; Meyer, H.-D.; Köppel, H.; Cederbaum, L. S.; Burghardt, I. Using the MCTDH wavepacket propagation method to describe multimode non-adiabatic dynamics. *Int. Rev. Phys. Chem.* **2008**, *27*, 569−606.

(27) Zuehlsdorff, T. J.; Shedge, S. V.; Lu, S.-Y.; Hong, H.; Aguirre, V. P.; Shi, L.; Isborn, C. M. Vibronic and Environmental Effects in Simulations of Optical Spectroscopy. *Annu. Rev. Phys. Chem.* **2021**, *72*, 165−188.

(28) Anda, A.; De Vico, L.; Hansen, T.; Abramavičius, D. Absorption and Fluorescence Lineshape Theory for Polynomial Potentials. *J. Chem. Theory Comput.* **2016**, *12*, 5979−5989.

(29) Köppel, H. Diabatic Representation: Methods for the Construction of Diabatic Electronic State. *Conical Intersections; Electronic Structure, Dynamics and Spectroscopy*; World Scientific, 2004; pp 175−204.

(30) Yaghoubi Jouybari, M.; Liu, Y.; Improta, R.; Santoro, F. Ultrafast Dynamics of the Two Lowest Bright Excited States of Cytosine and 1-Methylcytosine: A Quantum Dynamical Study. *J. Chem. Theory Comput.* **2020**, *16*, 5792−5808.

(31) Aleotti, F.; Aranda, D.; Yaghoubi Jouybari, M.; Garavelli, M.; Nenov, A.; Santoro, F. Parameterization of a linear vibronic coupling model with multiconfigurational electronic structure methods to study the quantum dynamics of photoexcited pyrene. *J. Chem. Phys.* **2021**, *154*, 104106.

(32) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* **1992**, *96*, 1218−1226.

(33) Sauri, V.; Serrano-Andrés, L.; Shahi, A. R. M.; Gagliardi, L.; Vancoillie, S.; Pierloot, K. Multiconfigurational Second-Order Perturbation Theory Restricted Active Space (RASPT2) Method for Electronic Excited States: A Benchmark Study. *J. Chem. Theory Comput.* **2011**, *7*, 153−168.

(34) Lee, S.-Y.; Pollard, W.; Mathies, R. A. Quantum theory for transition state absorption. *Chem. Phys. Lett.* **1989**, *160*, 531−537.

(35) Freibert, A.; Mendive-Tapia, D.; Huse, N.; Vendrell, O. Femtosecond x-ray absorption spectroscopy of pyrazine at the nitrogen K-edge: on the validity of the Lorentzian limit. *J. Phys. B: At., Mol. Opt. Phys.* **2021**, *54*, 244003.

(36) Penfold, T. J.; Pápai, M.; Rozgonyi, T.; Møller, K. B.; Vankó, G. Probing spin−vibronic dynamics using femtosecond X-ray spectroscopy. *Faraday Discuss.* **2016**, *194*, 731−746.

(37) Capano, G.; Milne, C. J.; Chergui, M.; Rothlisberger, U.; Tavernelli, I.; Penfold, T. J. Probing wavepacket dynamics using ultrafast x-ray spectroscopy. *J. Phys. B: At., Mol. Opt. Phys.* **2015**, *48*, 214001.

(38) Tsuru, S.; Vidal, M. L.; Pápai, M.; Krylov, A. I.; Møller, K. B.; Coriani, S. Time-resolved near-edge X-ray absorption fine structure of pyrazine from electronic structure and nuclear wave packet dynamics simulations. *J. Chem. Phys.* **2019**, *151*, 124114.

(39) Foggi, P.; Pettini, L.; Santa, I.; Righini, R.; Califano, S. Transient absorption and vibrational relaxation dynamics of the lowest excited singlet state of pyrene in solution. *J. Phys. Chem.* **1995**, *99*, 7439−7445.

(40) Raytchev, M.; Pandurski, E.; Buchvarov, I.; Modrakowski, C.; Fiebig, T. Bichromophoric Interactions and Time-Dependent Excited State Mixing in Pyrene Derivatives. A Femtosecond Broad-Band Pump-Probe Study. *J. Phys. Chem. A* **2003**, *107*, 4592−4600.

(41) Neuwahl, F. V. R.; Foggi, P. Direct Observation of S2−S1 Internal Conversion in Pyrene by Femtosecond Transient Absorption. *Laser Chem.* **1999**, *19*, 375−379.

(42) Nenov, A.; Giussani, A.; Fingerhut, B. P.; Rivalta, I.; Dumont, E.; Mukamel, S.; Garavelli, M. Spectral lineshapes in nonlinear electronic spectroscopy. *Phys. Chem. Chem. Phys.* **2015**, *17*, 30925−30936.

(43) Roos, M. K.; Reiter, S.; de Vivie-Riedle, R. Ultrafast relaxation from 1La to 1Lb in pyrene: a theoretical study. *Chem. Phys.* **2018**, *515*, 586−595.

(44) Montorsi, F.; Segatta, F.; Nenov, A.; Mukamel, S.; Garavelli, M. Soft X-ray Spectroscopy Simulations with Multiconfigurational Wave Function Theory: Spectrum Completeness, Sub-eV Accuracy, and Quantitative Reproduction of Line Shapes. *J. Chem. Theory Comput.* **2022**, *18*, 1003−1016.

(45) Fdez Galván, I.; Vacher, M.; Alavi, A.; Angeli, C.; Aquilante, F.; Autschbach, J.; Bao, J. J.; Bokarev, S. I.; Bogdanov, N. A.; Carlson, R. K.; et al. OpenMolcas: From Source Code to Insight. *J. Chem. Theory Comput.* **2019**, *15*, 5925−5964.

(46) Aquilante, F.; Autschbach, J.; Baiardi, A.; Battaglia, S.; Borin, V. A.; Chibotaru, L. F.; Conti, I.; De Vico, L.; Delcey, M.; Fdez Galván, I.; et al. Modern quantum chemistry with [Open]Molcas. *J. Chem. Phys.* **2020**, *152*, 214117.

(47) Li Manni, G.; Fdez Galván, I.; Alavi, A.; Aleotti, F.; Aquilante, F.; Autschbach, J.; Avagliano, D.; Baiardi, A.; Bao, J. J.; Battaglia, S.; et al. The OpenMolcas Web: A Community-Driven Approach to Advancing Computational Chemistry. *J. Chem. Theory Comput.* **2023**, *19*, 6933−6991.

(48) Ma, D.; Li Manni, G.; Gagliardi, L. The generalized active space concept in multiconfigurational self-consistent field methods. *J. Chem. Phys.* **2011**, *135*, 044128.

(49) Worth, G. Quantics: A general purpose package for quantum molecular dynamics simulations. *Comput. Phys. Commun.* **2020**, *248*, 107040.

(50) Mukamel, S. *Principles of Nonlinear Optical Spectroscopy*; Oxford University Press: New York, 1995.

(51) Abramavicius, D.; Palmieri, B.; Voronine, D. V.; Šanda, F.; Mukamel, S. Coherent Multidimensional Optical Spectroscopy of Excitons in Molecular Aggregates Quasiparticle versus Supermolecule Perspectives. *Chem. Rev.* **2009**, *109*, 2350−2408.

(52) Lax, M. The Franck-Condon Principle and Its Application to Crystals. *J. Chem. Phys.* **1952**, *20*, 1752−1760.

(53) Lami, A.; Santoro, F. *Computational Strategies for Spectroscopy*; Barone, V., Ed.; John Wiley & Sons, Inc., 2011; Chapter 10, pp 475−516.

(54) Markham, J. J. Interaction of Normal Modes with Electron Traps. *Rev. Mod. Phys.* **1959**, *31*, 956−989.

(55) Biczysko, M.; Bloino, J.; Santoro, F.; Barone, V. *Computational Strategies for Spectroscopy*; Barone, V., Ed.; John Wiley & Sons, Inc., 2011; Chapter 8, pp 361−443.

(56) Ianconescu, R.; Pollak, E. Photoinduced Cooling of Polyatomic Molecules in an Electronically Excited State in the Presence of Dushinskii Rotations. *J. Phys. Chem. A* **2004**, *108*, 7778−7784.

(57) Huh, J.; Berger, R. Coherent state-based generating function approach for Franck-Condon transitions and beyond. *J. Phys.: Conf. Ser.* **2012**, *380*, 012019.

(58) Montorsi, F.; Aranda, D.; Garavelli, M.; Santoro, F.; Segatta, F. Spectroscopy from quantum dynamics: a mixed wave function/ analytical line shape functions approach. *Theor. Chem. Acc.* **2023**, *142*, 108.

Time-resolved X-ray absorption spectroscopy:

an MCTDH quantum dynamics protocol

Francesco Segatta,[†] Daniel Aranda,[‡,¶] Flavia Aleotti,[†] Francesco Montorsi,[†] Shaul

Mukamel,[§] Marco Garavelli,[*,†] Fabrizio Santoro,[*,¶] and Artur Nenov[*,†]

[†]*Dipartimento di Chimica Industriale "Toso Montanari", University of Bologna, Viale del Risorgimento, 4, 40136 Bologna, Italy*

[‡]*ICMol, Universidad de Valencia, c/Catedrático José Beltrán, 2, 46980 Paterna, Spain*

[¶]*Istituto di Chimica dei Composti Organometallici (ICCOM-CNR), Area della Ricerca del CNR, Via Moruzzi 1, I-56124 Pisa, Italy*

[§]*Department of Chemistry and Department of Physics and Astronomy, University of California, Irvine, 92697, USA*

E-mail: marco.garavelli@unibo.it; fabrizio.santoro@pi.iccom.cnr.it; artur.nenov@unibo.it

This PDF includes:

- Active spaces for core-excited state;

- Different approaches to simulate the LA spectrum;

- Lineshape functions second order Taylor expansion;

- The commutator $\left[\hat{H}_e, \hat{H}_c \right]$;

- The *coherence* $R^{(3)ESA}_{WPO^*}(t_3, t_2)$ term in the EGVA approximation;

- WPO* vs EGVA LA: high and low frequency modes;

- WPO* vs EGVA LA: at various values of τ_c;

- EGVA XAS spectrum: variable vs constant variance;

- Energy gaps, energy-gap variances and TDMs;

- Pyrine LA: full mode models vs reduced (15) modes model;

- Response function Fourier transform: technical details.

S1   Active spaces for core-excited states
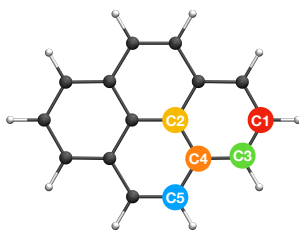
Figure s1: Labeling of the five types of C atoms (multiplicity: 2 for C1 and C2, 4 for C3-C5) identified in the S_0 minimum of pyrene (D_{2h} point group).
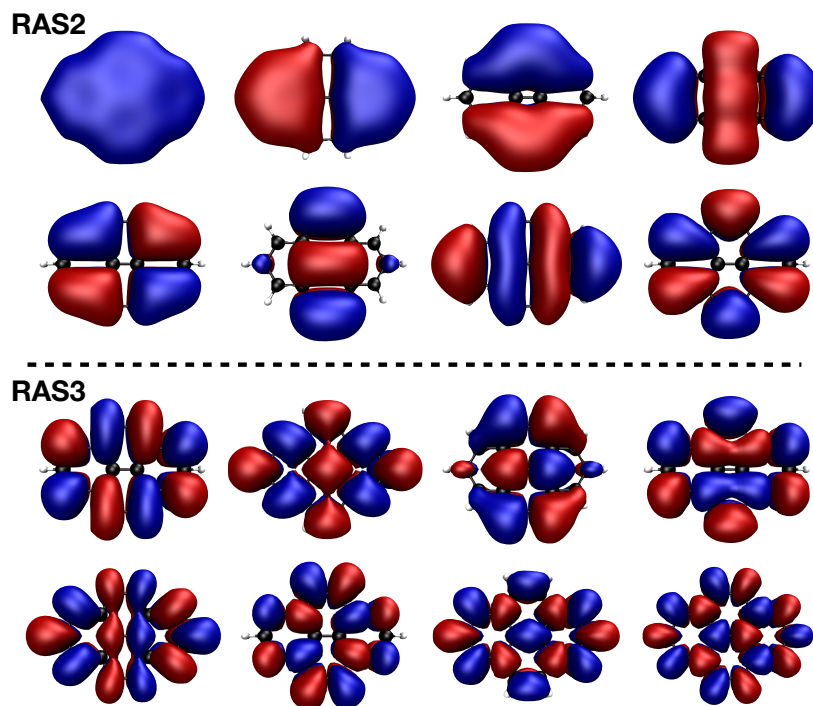
Figure s2: π (RAS2) and π^* (RAS3) orbitals used for the RASSCF/ANO-RCC calculation of the valence excited states (e manifold). For each carbon atom type identified in Figure s1 the active space was completed by putting the corresponding 1s orbital in RAS1.

Figure s3: 1s (RAS1), π (RAS2) and π^* (RAS3) orbitals used for the RASSCF/ANO-RCC calculation of the core excited states associated with carbon atom type C1 (see labels in Figure s1).
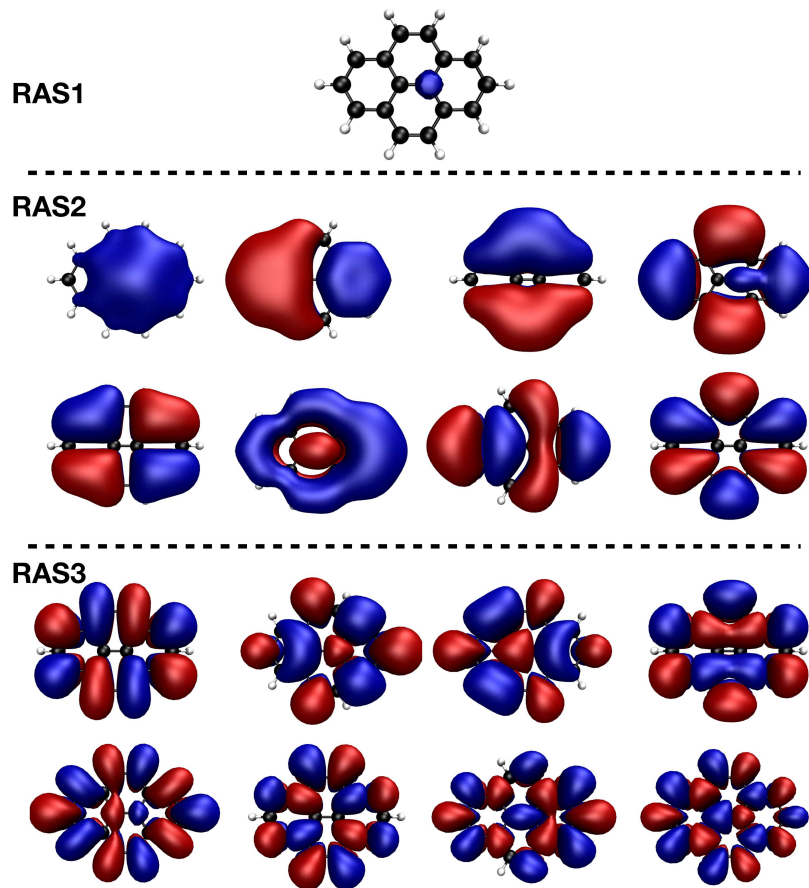
Figure s4: 1s (RAS1), π (RAS2) and π^* (RAS3) orbitals used for the RASSCF/ANO-RCC calculation of the core excited states associated with carbon atom type C2 (see labels in Figure s1).

Figure s5: 1s (RAS1), π (RAS2) and π^* (RAS3) orbitals used for the RASSCF/ANO-RCC calculation of the core excited states associated with carbon atom type C3 (see labels in Figure s1).

Figure s6: 1s (RAS1), π (RAS2) and π^* (RAS3) orbitals used for the RASSCF/ANO-RCC calculation of the core excited states associated with carbon atom type C4 (see labels in Figure s1).

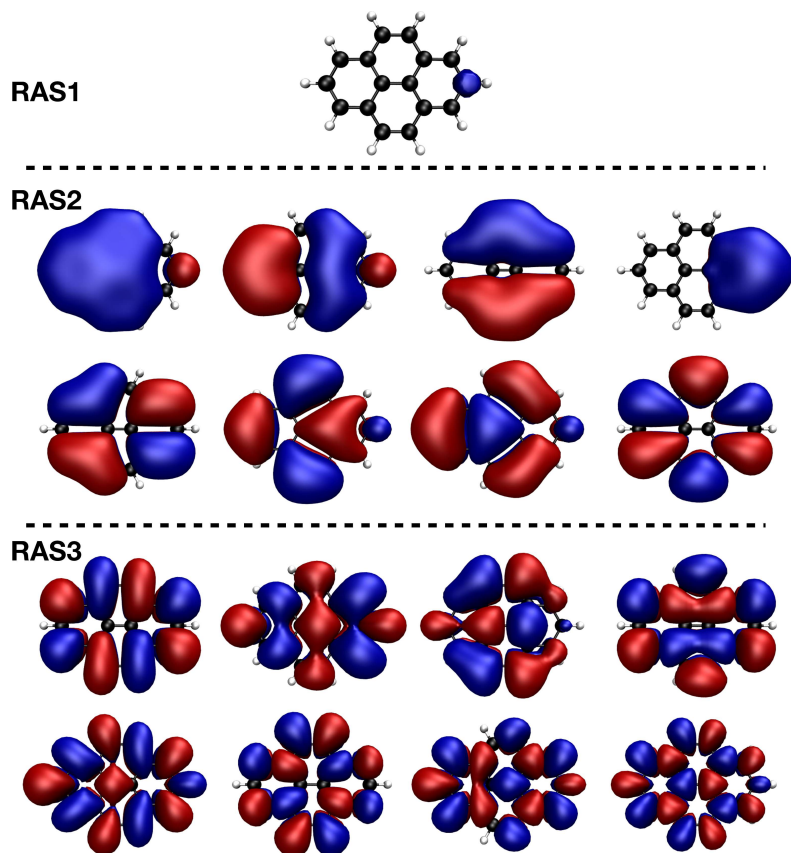Figure s7: 1s (RAS1), π (RAS2) and π^* (RAS3) orbitals used for the RASSCF/ANO-RCC calculation of the core excited states associated with carbon atom type C5 (see labels in Figure s1).
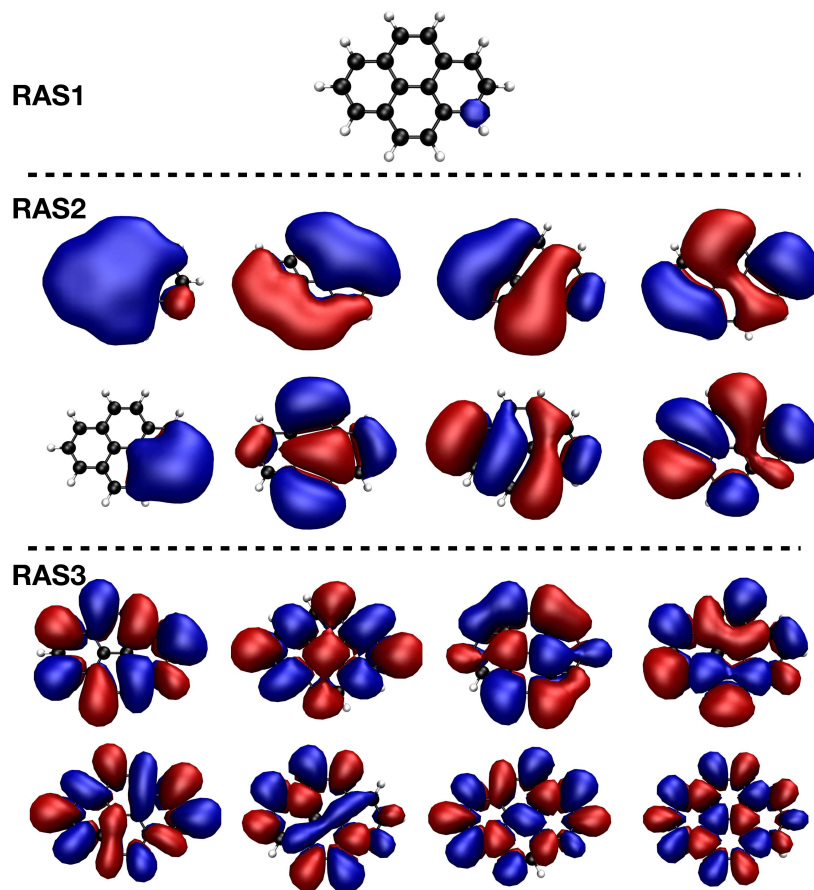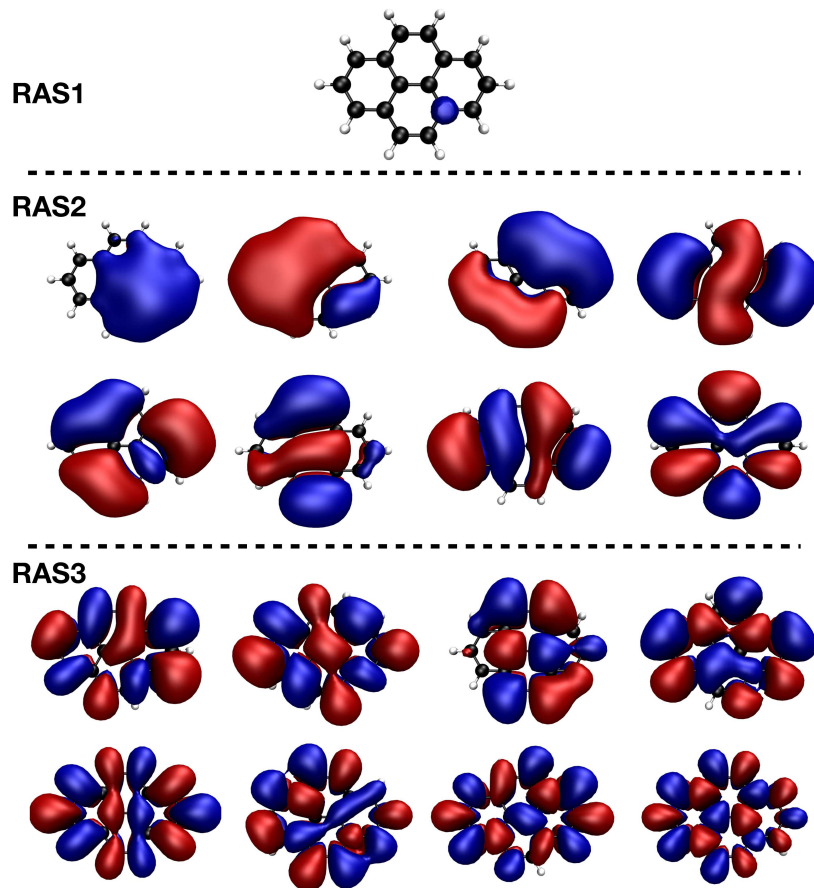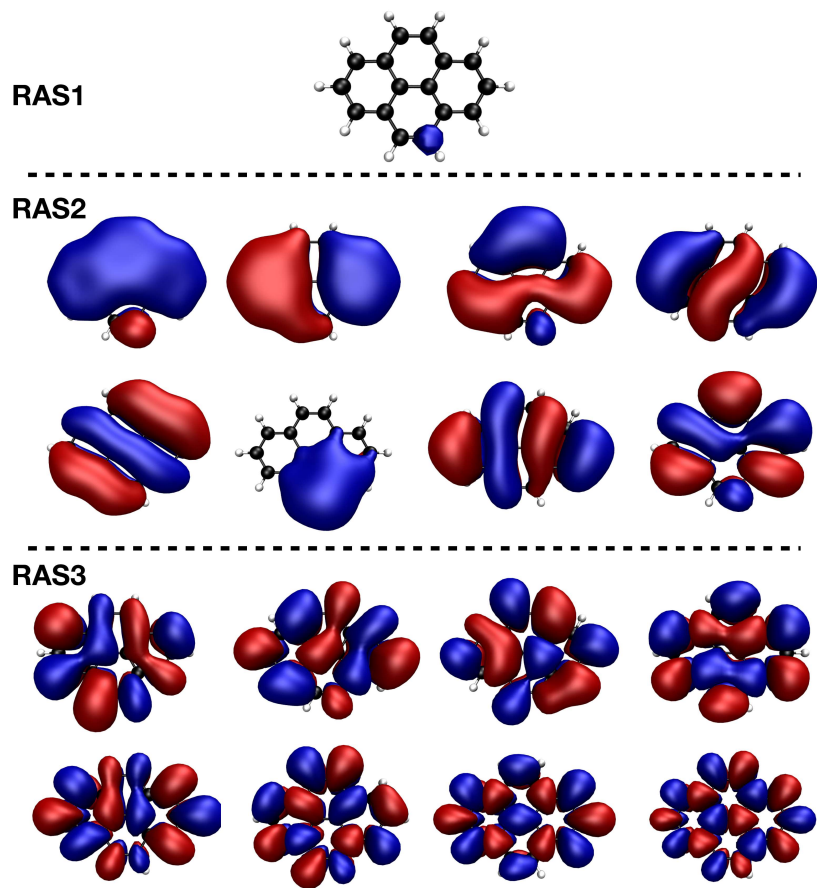
S2 Different approaches to simulate the LA spectrum

We seek to evaluate the time-dependent overlap term given by

$$\langle \chi_g(t)|\, \chi_{c^*}(t)\rangle = \langle \chi_g(0)|\, e^{+i\hat{H}_g t} e^{-i\hat{H}_c t}\, |\chi_c(0)\rangle \tag{S1}$$

where the $*$ indicates that the ket WP will evolve adiabatically on the potential energy surface of the state c (i.e., it will stay on the electronic state onto which it was projected at time $t = 0$).

S2.1 *Time-independent* approach

We introduce the eigenstates of $\hat{H}_g$ and $\hat{H}_c$, which are the vibrational states of the two wells. We refer to the GS vibrational eigenstates as $|v_{m,i}\rangle$ (where the index m denotes the mode, and the index i denotes the vibrational level on that mode), and the excited state eigenstates as $|u_{m,i}\rangle$. Therefore

$$\hat{H}_g\, |v_{m,i}\rangle = \epsilon_{m,i}\, |v_{m,i}\rangle \tag{S2}$$

$$\hat{H}_c\, |u_{m,i}\rangle = \epsilon_{m,i}\, |u_{m,i}\rangle \tag{S3}$$

$\epsilon_{m,i}$ is the vibrational energy of state $|v_{m,i}\rangle$ (and $|u_{m,i}\rangle$), i.e. $\epsilon_{m,k} = \frac{1}{2}\hbar\omega_m + k\hbar\omega_m$. One can jump from the GS vibrational basis to the ES vibrational basis by a simple basis transformation, given by:

$$|v_{m,i}\rangle = \sum_j \langle u_{m,j}|v_{m,i}\rangle\, |u_{m,j}\rangle \tag{S4}$$

Therefore, for our single mode m, we have $|\chi_{g,m}(0)\rangle = |\chi_{c,m}(0)\rangle = |v_{m,0}\rangle$ (where we have also added the subscript m to the GS and ES WP to specify that this part of the WP is along mode m; the complete WP is the Hartree product $|\chi_g(0)\rangle = \prod_m |\chi_{g,m}(0)\rangle$).[1]

---

[1] From this it follows that the overlap may be factorized into a product of contributions of each mode, since the total Hamiltonian can be written as a sum of commuting mode-specific Hamiltonians.

We can now return to our original problem, which is to derive an expression for the overlap term as a function of t (again in absence of transport). We first recast the overlap in the form:

$$\langle \chi_{g,m}(0)| e^{+i\hat{H}_g t} e^{-i\hat{H}_c t} |\chi_{c,m}(0)\rangle = \langle v_{m,0}| e^{+i\hat{H}_g t} e^{-i\hat{H}_c t} |v_{m,0}\rangle = e^{+i\frac{\epsilon_{m,0}}{\hbar}t} \langle v_{m,0}| e^{-i\hat{H}_c t} |v_{m,0}\rangle \quad (S5)$$

where in the last step we employed the fact that $|v_{m,0}\rangle$ is an eigenstate of $\hat{H}_g$, and $\frac{\omega_m}{2} = \frac{\epsilon_{m,0}}{\hbar}$. The problem then turns into the evaluation of $\langle v_{m,0}| e^{-i\hat{H}_c t} |v_{m,0}\rangle$.

We can now expand the bra and ket sides in terms vibrational eigenstates of the ES well, which gives:

$$\begin{aligned}
\langle v_{m,0}| e^{-i\hat{H}_c t} |v_{m,0}\rangle &= \sum_{jk} \langle v_{m,0}|u_{m,j}\rangle \langle u_{m,j}|e^{-i\hat{H}_c t}|u_{m,k}\rangle \langle u_{m,k}|v_{m,0}\rangle \\
&= \sum_k |\langle v_{m,0}|u_{m,k}\rangle|^2 e^{-i\frac{\epsilon_{m,k}}{\hbar}t} \quad (S6)
\end{aligned}$$

where in the last step we took advantage of the orthogonality of the vibrational states of the ES. Note that the Franck-Condon factor, $|\langle v_{m,0}|u_{m,k}\rangle|^2$ appears in the expression. We finally get a weighted sum of imaginary exponentials, where each exponential oscillates with the frequency of the respective vibrational state, and is weighted by the projection of the initial GS WP onto the ES set of vibrational states.

S2.2   *Time-dependent* (WPO) approach

In Ref. 1 we explored how the time-dependent overlap can be computed explicitly via a QD simulation. This involves propagating the two wave-packets on the g and c surfaces, and computing their overlap at regular time intervals (which are set by the fastest time-scale considered). The interested reader can refer to Ref. 1 for additional details.

S2.3   *Lineshape functions* approach

Mukamel et al. have shown how, in absence of transport (i.e., when the dynamics is adiabatic), this overlap term can be computed analytically. Briefly, one expresses the ES potential in terms of the GS one, rewrites the expression in terms of time-ordered exponential, expand to second-order (the so-called *cumulant expansion*) and obtain the energy fluctuation autocorrelation functions, spectral density, and eventually the line-shape functions.[2,3]

Note that an approach that combines explicit WP overlap calculations along some modes, with analytical expressions via cumulant expansion has been recently demonstrated in Ref. 4.

S2.4   *Energy-gap/Energy-gap variance* approach

The main equations are reported in the main text. Here we provide additional details.

Definition of energy-gap and energy-gap variance

$$
\begin{aligned}
\langle \hat{H}_c \rangle_{g,m} = \langle v_{m,0}|\hat{H}_c|v_{m,0}\rangle = \\
= \langle v_{m,0}|\hat{T} + \hat{V}_c|v_{m,0}\rangle = \\
= \langle v_{m,0}|\hat{T} + \hat{V}_g + \lambda_{c,m} - \omega_m \hat{Q}_m d_{c,m}|v_{m,0}\rangle \\
= \langle v_{m,0}|\hat{H}_g|v_{m,0}\rangle + \lambda_{c,m} - \omega_m d_{c,m} \langle v_{m,0}|\hat{Q}_m|v_{m,0}\rangle \\
= \epsilon_{m,0} + \lambda_{c,m}
\end{aligned}
\tag{S7}
$$

where $\langle v_{m,0}|\hat{Q}_m|v_{m,0}\rangle$ (the centroid of the initial WP) is null. By adding the adiabatic $g-c$ energy gap, $\Delta E_{gc}^{(ad)}$ to this expression, one obtains:

$$
\langle \hat{H}_c \rangle_{g,m} + \Delta E_{gc}^{(ad)} = (\Delta E_{gc}^{(ad)} + \lambda_{c,m}) + \epsilon_{m,0} = \Delta E_{gc} + \epsilon_{m,0}
\tag{S8}
$$

where ΔE_{gc} is the vertical transition, which summed with the zero point energy $\epsilon_{m,0}$ gives the first moment of the spectrum.

$$\sigma_{gc;m}^2 = \langle v_{m,0}|(\hat{H}_c - \langle \hat{H}_c \rangle_{g,m})^2|v_{m,0} \rangle =$$

$$= \langle v_{m,0}| \left(\hat{T} + \hat{V}_g + \lambda_{c,m} - \omega_m \hat{Q}_m d_{c,m} \right)^2 |v_{m,0} \rangle - (\epsilon_{m,0} + \lambda_{c,m})^2 =$$

$$= \omega_m^2 (d_{c,m})^2 \langle \hat{Q}_m^2 \rangle_{g,m} = \tag{S9}$$

$$= \frac{1}{2} \omega_m^2 (d_{c,m})^2 =$$

$$= \omega_m \lambda_{c,m}$$

where in the last equality we made use of the fact that $\langle \hat{Q}_m^2 \rangle_{g,m} = 1/2$ (in dimensionless coordinates). $\sigma_{gc;m}^2$ is easily recognized as the energy-gap variance: in fact,

$$\sigma_{gc;m}^2 = \langle v_{m,0}|(\hat{H}_c - \langle \hat{H}_c \rangle_{g,m})^2|v_{m,0} \rangle =$$
$$= \langle v_{m,0}| \left(\left[\hat{H}_c + \Delta E_{gc}^{(ad)} \right] - \left[\langle \hat{H}_c \rangle_{g,m} + \Delta E_{gc}^{(ad)} \right] \right)^2 |v_{m,0} \rangle \tag{S10}$$

would is the (vertical) energy-gap variance.

WP overlap: from one mode, to all modes

$$\langle \chi_g(t)|\chi_c(t) \rangle = e^{-i\omega_{gc}^{(ad)}t} \prod_m \langle \chi_{g,m}(0)|e^{+i\hat{H}_{g,m}t} e^{-i\hat{H}_{c,m}t}|\chi_{c,m}(0) \rangle \sim$$

$$\sim e^{-i\omega_{gc}^{(ad)}t} \prod_m e^{-i\lambda_{c,m}t} e^{-\frac{1}{2}\sigma_{gc;m}^2 t^2} =$$

$$= e^{-i\omega_{gc}^{(ad)}t} \prod_m e^{-i\lambda_{c,m}t} \prod_m e^{-\frac{1}{2}\sigma_{gc;m}^2 t^2} = \tag{S11}$$

$$= e^{-i\omega_{gc}^{(ad)}t} e^{-i\sum_m \lambda_{c,m}t} e^{-\frac{1}{2}\sum_m \sigma_{gc;m}^2 t^2} =$$

$$= e^{-i(\omega_{gc}^{(ad)}+\lambda_c)t - \frac{1}{2}\varsigma_{gc}^2 t^2} =$$

$$= e^{-i\omega_{gc}t - \frac{1}{2}\varsigma_{gc}^2 t^2}$$

where $\lambda_c = \sum_m \lambda_{c,m}$ is the total reorganization energy, $\omega_{gc} = \omega_{gc}^{(ad)} + \lambda_c$ is the vertical energy gap, and $\varsigma_{gc}^2 = \sum_m \sigma_{gc;m}^2$ is the total variance, i.e., the sum of the variances along all modes.

Voigt line-shape profile

Eq. 19 in the main text gives the EGVA response function, from which the spectrum can be obtained by Fourier transformation. Since the Fourier transform of a Gaussian and a decaying exponential gives, respectively, a Gaussian and a Lorentzian function, we obtain the following expression for the linear spectrum:

$$S_{EGVA}^{(1)}(\Omega) \quad \propto \quad \sum_c \boldsymbol{\mu}_{gc} \boldsymbol{\mu}_{cg} \; G_{gc}(\Omega - \omega_{gc}) * L_{gc}(\Omega) \tag{S12}$$

where

$$L_{gc}(\Omega) = \frac{2\tau_c}{(2\tau_c\Omega)^2 + 1}$$

$$G_{gc}(\Omega - \omega_{gc}) = \exp\left[-\frac{(\Omega - \omega_{gc})^2}{2\varsigma_{gc}^2}\right] \tag{S13}$$

and the symbol $*$ indicates the convolution giving rise to the *Voigt* line-shape profile.

S3 Lineshape functions second order Taylor expansion

The lineshape function $g_{ab}(t)$, involved in all expressions, reads

$$g_{ab}(t) = \frac{1}{2} \sum_m d_{a,m} d_{b,m} \left[\coth\left(\frac{\hbar \omega_m}{2 k_B T} \right) (1 - \cos(\omega_m t)) + i \sin(\omega_m t) \right] \tag{S14}$$

In the zero temperature limit the term $\coth \frac{\hbar \omega_m}{2 k_B T} \approx 1$, so that the above expression simplifies to

$$g_{ab}(t) = \frac{1}{2} \sum_m d_{a,m} d_{b,m} \left[1 - \cos(\omega_m t) + i \sin(\omega_m t) \right] \tag{S15}$$

The line-shape functions are integral transformation of the autocorrelation function describing the undamped oscillatory dynamics on the m-th normal mode harmonic potential with frequency ω_m and relative displacement with respect to the ground state equilibrium $d_{a,m}$ and $d_{b,m}$ in the a-th and b-th electronic state, respectively.[2]

S3.1 First-order response

The response function of the linear absorption involves $g_{cc}(t)$, i.e.

$$g_{cc}(t) = \frac{1}{2} \sum_m d_{c,m}^2 \left[1 - \cos(\omega_m t) + i \sin(\omega_m t) \right] \tag{S16}$$

The second order Taylor expansion of $g_{cc}(t)$ reads

$$\begin{aligned}
g_{cc}(t) &\sim \sum_m \left[i \frac{1}{2} \omega_m d_{c,m}^2 t - \frac{1}{2} \left(\frac{1}{2} \omega_m^2 d_{c,m}^2 \right) t^2 \right] = \\
&= i \sum_m \lambda_{c,m} t - \frac{1}{2} \sum_m \sigma_{gc,m}^2 t^2 = \\
&= i \lambda_c t - \frac{1}{2} \varsigma_{gc}^2 t^2
\end{aligned} \tag{S17}$$

which exactly matches what we have found in the main text.

---

[2]In the present formulation of the line shape function, the reorganization energy $\lambda_{e,m}$ in the e-th electronic state is merged in the phase function.

S3.2   Third-order response

Let us now consider the second order Taylor expansion of the (exact) third-order response lineshape function expression for the ESA contribution, to demonstrate that this is equivalent to the one obtained in the main text in the case of adiabatic dynamics.

Under the assumption of adiabatic dynamics, the response function can be expressed as [2,3]

$$
\begin{aligned}
R^{(3)ESA}(t_3, t_2) = & -\sum_{e,c} \boldsymbol{\mu}_{ec}\boldsymbol{\mu}_{ce}\rho_e(t_2)e^{-i(\omega_{ec}^{(ad)}-i/2\tau_c)t_3+\varphi_{ece}(0,t_2,t_2+t_3,0)} \\
& -\sum_{e\neq e',c} \boldsymbol{\mu}_{ec}\boldsymbol{\mu}_{ce'}e^{-i(\omega_{ee'}^{(ad)}-i/2\tau_{ee'})t_2-i(\omega_{ec}-i/2\tau_c)t_3+\varphi_{e'ce}(0,t_2,t_2+t_3,0)}
\end{aligned}
\tag{S18}
$$

where the ESA contribution has been separated for *population* ($e = e'$ during t_2) and *coherence* ($e = e'$ during t_2) evolution. $\varphi_{cba}(\tau_4, \tau_3, \tau_2, \tau_1)$ are the multidimensional phase-functions encoding the time-dependent overlaps of wave-packets moving on different adiabatic surfaces. These are given by:

$$
\begin{aligned}
\varphi_{cba}(\tau_4, \tau_3, \tau_2, \tau_1) = & -g_{cc}(\tau_{43}) - g_{bb}(\tau_{32}) - g_{aa}(\tau_{21}) \\
& - g_{cb}(\tau_{42}) + g_{cb}(\tau_{43}) + g_{cb}(\tau_{32}) \\
& - g_{ca}(\tau_{41}) + g_{ca}(\tau_{42}) + g_{ca}(\tau_{31}) - g_{ca}(\tau_{32}) \\
& - g_{ba}(\tau_{31}) + g_{ba}(\tau_{32}) + g_{ba}(\tau_{21})
\end{aligned}
\tag{S19}
$$

where $\tau_{ij} = \tau_i - \tau_j$.

Applied to the *population* ESA the phase function reads

$$
\begin{aligned}
\varphi_{ece}(0, t_2, t_2 + t_3, 0) = & 2g_{ec}(t_3) - g_{ee}(t_3) - g_{cc}(t_3) + 2\Im\left[g_{ec}(t_2) - g_{ee}(t_2) + \right. \\
& \left. -g_{ec}(t_2 + t_3) + g_{ee}(t_2 + t_3)\right]
\end{aligned}
\tag{S20}
$$

Inserting the expression for the line-shape function $g_{ab}(t)$ one obtains

$$\varphi_{ece}(0, t_2, t_2 + t_3, 0) = \sum_m \left(i \sin(\omega_m t_2) (d_{e,m} d_{c,m} - d_{e,m} d_{e,m}) - i \sin(\omega_m (t_2 + t_3)) (d_{e,m} d_{c,m} - d_{e,m} d_{e,m}) \right.$$
$$\left. -\frac{1}{2} (1 - \cos(\omega_m t_3) + i \sin(\omega_m t_3)) (d_{e,m} - d_{c,m})^2 \right)$$

$$\text{(S21)}$$

We now apply the short time approximation with respect to t_3 in eq. S21, i.e., we consider the second order Taylor expansion of the phase function, obtaining

$$\varphi_{ece}(0, t_2, t_2 + t_3, 0) \sim \sum_m \left(i\omega_m \left[-\cos(\omega_m t_2) (d_{e,m} d_{c,m} - d_{e,m} d_{e,m}) + \frac{1}{2} (d_{e,m} - d_{c,m})^2 \right] t_3 \right.$$
$$\left. + \frac{1}{2} i\omega_m^2 \left[\sin(\omega_m t_2) (d_{e,m} d_{c,m} - d_{e,m} d_{e,m}) + \frac{1}{2} i (d_{e,m} - d_{c,m})^2 \right] t_3^2 \right)$$

$$\text{(S22)}$$

where the t_2 dependent first term of eq. S21 is cancelled by the zeroth order term in the Taylor expansion of the $t_2 + t_3$ dependent term.

At this point we can compare, order by order in t_3, the expressions for $R^{(3)ESA}(t_3, t_2)$ obtained here with those derived in the main text (eq. 40). The t_3 dependent term in eq. S22 should be compared with $-i \left(\langle \hat{H}_c \rangle_{t_2} - \langle \hat{H}_e \rangle_{t_2} \right) t_3$ (main text, eq. 35). Indeed, by expressing the reorganization energy in terms of the displacements, and by considering that in the case of adiabatic dynamics the centroid of the nuclear wave-packet evolves along a classical trajectory, for which $\langle Q_m \rangle_{e,t_2} = d_{e,m} - d_{e,m} \cos(\omega_m t_2)$, we obtain

$$-i \left(\langle \hat{H}_c \rangle_{t_2} - \langle \hat{H}_e \rangle_{t_2} \right) t_3 = \sum_m i\omega_m \left[-\frac{1}{2} (d_{c,m} d_{c,m} - d_{e,m} d_{e,m}) + \langle Q_m \rangle_{t_2} (d_{c,m} - d_{e,m}) \right] t_3$$
$$= \sum_m i\omega_m \left[-\cos(\omega_m t_2) (d_{e,m} d_{c,m} - d_{e,m} d_{e,m}) + \frac{1}{2} (d_{e,m} - d_{c,m})^2 \right] t_3$$

$$\text{(S23)}$$

The t_3^2 dependent term in eq. S22, instead, can be obtained from the energy gap *pseudo-variance* $\tilde{\varsigma}_{ec}^2(t_2) = \varsigma_{ec}^2(t_2) + i2\Xi_{ec}(t_2)$. We rewrite $\varsigma_{ec}^2(t_2)$ in terms of the WP position operator variance (main text, eq. 36), and express $i2\Xi_{ec}(t_2)$ as the commutator $\langle [\hat{T}_m, \hat{Q}_m] \rangle_{t_2}$ (see

next Section), so that

$$-\frac{1}{2}\tilde{\varsigma}_{ec}^2(t_2)t_3^2 = \sum_m -\frac{1}{2}\left[\omega_m \left\langle\left[\hat{T}_m,\hat{Q}_m\right]\right\rangle_{t_2}(d_{c,m}-d_{e,m}) + \omega_m^2(d_{c,m}-d_{e,m})^2\left(\langle\hat{Q}_m^2\rangle_{e,t_2} - \langle\hat{Q}_m\rangle_{e,t_2}^2\right)\right]t_3^2$$

$$= \sum_m \frac{1}{2}i\omega_m^2\left[\sin(\omega_m t_2)(d_{e,m}d_{c,m}-d_{e,m}d_{e,m}) - \frac{1}{2}i(d_{c,m}-d_{e.m})^2\right]t_3^2$$

(S24)

where we have made use of the fact that

$$\left\langle\left[\hat{T}_m,\hat{Q}_m\right]\right\rangle_{t_2} = \left\langle\left[\hat{H}_e,\hat{Q}_m\right]\right\rangle_{t_2} = i\frac{d\langle\hat{Q}_m\rangle_{e,t_2}}{dt} = -i\omega_m d_{e,m}\sin(\omega_m t_2)$$

(S25)

and the last step again follows from the assumed adiabatic dynamics, that gives an analytical expression for the time dependence of the WP centroid. Moreover, we also made use of the fact that the position operator variance of the adiabatically evolving wave-packet is constant, i.e.,

$$\langle\hat{Q}_m^2\rangle_{e,t_2} - \langle\hat{Q}_m\rangle_{e,t_2}^2 = 1/2$$

(S26)

An analytical expression for the linear chirp $\Xi_{ec}(t_2)$ term is thus obtained, in the assumption that the system always follows an adiabatic dynamics

$$\Xi_{ec}(t_2) = \frac{1}{2}\omega_m^2 d_{e,m}\sin(\omega_m t_2)(d_{c,m}-d_{e,m}) =$$

$$= \frac{1}{2}\omega_m^2 \langle P_m\rangle_{e,t_2}(d_{c,m}-d_{e,m})$$

(S27)

where $\langle P_m\rangle_{e,t_2} = d_{e,m}\sin(\omega t_2)$ is the expectation value of the momentum operator in dimensionless coordinates.

S4 The commutator $\left[\hat{H}_e, \hat{H}_c\right]$

Let us first rewrite $\hat{H}_{c,m}$ as $\hat{H}_{c,m} = \hat{H}_{e,m} + \frac{1}{2}\omega_m \left(d_{e,m} - d_{c,m}\right)^2 - \omega_m d_{e,m} \left(d_{e,m} - d_{c,m}\right) + \omega_m \hat{Q}_m \left(d_{e,m} - d_{c,m}\right)$, so that $\left[\hat{H}_{e,m}, \hat{H}_{c,m}\right]$ simplifies to $\omega_m \left(d_{e,m} - d_{c,m}\right)\left[\hat{T}_m, \hat{Q}_m\right]$. Therefore, for the population contribution, we have

$$\langle \chi_{e,m}(t_2)| \left[\hat{H}_{e,m}, \hat{H}_{c,m}\right] |\chi_{e,m}(t_2)\rangle = \sum_m \omega_m (d_{e,m} - d_{c,m}) \langle \chi_{e,m}(t_2)| \left[\hat{T}_m, \hat{Q}_m\right] |\chi_{e,m}(t_2)\rangle =$$

$$= \sum_m \omega_m (d_{e,m} - d_{c,m}) \langle \chi_{e,m}(t_2)| \left[-\frac{\omega_m}{2}\frac{\partial^2}{\partial Q_m^2}, \hat{Q}_m\right] |\chi_{e,m}(t_2)\rangle =$$

$$= -\sum_m \omega_m^2 (d_{e,m} - d_{c,m}) \langle \chi_{e,m}(t_2)| \frac{\partial}{\partial Q_m} |\chi_{e,m}(t_2)\rangle =$$

$$= -i \sum_m \omega_m^2 (d_{e,m} - d_{c,m}) \langle \hat{P}_m \rangle_{e,t_2}$$

(S28)

where $\langle \hat{P}_m \rangle_{t_2}$ is the expectation value of the momentum operator.

S5 $R_{WPO^*}^{(3)ESA}(t_3, t_2)$ *coherence* term in the EGVA approach

In what follows we consider the *coherence* term of the $R_{WPO^*}^{(3)ESA}(t_3, t_2)$, labled here as $R_{WPO^*}^{(3)ESA,coh}(t_3, t_2)$, aimed at deriving the corresponding EGVA expression. First, we start from eq. 28 of the main text, which gives

$$R_{WPO^*}^{(3)ESA,coh}(t_3, t_2) = - \mathcal{N}_\mu \sum_{e \neq e', c} \sqrt{\rho_e(t_2)\rho_{e'}(t_2)} \boldsymbol{\mu}_{ec} \boldsymbol{\mu}_{ce'} e^{-i\omega_{ec}^{(ad)} t_3} \langle \bar{\chi}_e^\mu(t_2)| e^{i\hat{H}_e t_3} e^{-i\hat{H}_c t_3} |\bar{\chi}_{e'}^\mu(t_2)\rangle e^{-t_3/2\tau_c}$$

$$(S29)$$

A significant difference with respect to the $R_{WPO^*}^{(3)ESA,pop}(t_3, t_2)$ term, is that the overlap between the bra and ket (normalized) wave-packet is not equal to 1. In fact

$$\langle \bar{\chi}_e^\mu(t_2)|\bar{\chi}_{e'}^\mu(t_2)\rangle = \begin{cases} 1 & \text{if } e = e' \\ \langle 1 \rangle_{t_2} \in \mathbb{C} & \text{if } e \neq e' \end{cases} \tag{S30}$$

where $\langle 1 \rangle_{t_2}$ is a generic t_2 function which can assume whatever value between -1 and 1.

Let us focus on the t_3 wave-packet overlap $\langle \bar{\chi}_e^\mu(t_2)| e^{i\hat{H}_e t_3} e^{-i\hat{H}_c t_3} |\bar{\chi}_{e'}^\mu(t_2)\rangle$. Following similar steps to those shown in the main text for the corresponding *population* term, we get

$$e^{i\hat{H}_e t_3} = e^{i\frac{\langle \hat{H}_e \rangle_{t2}}{\langle 1 \rangle_{t2}} t_3} e^{i\left(\hat{H}_e - \frac{\langle \hat{H}_e \rangle_{t2}}{\langle 1 \rangle_{t2}}\right) t_3}$$
$$e^{-i\hat{H}_c t_3} = e^{-i(\hat{H}_c - \frac{\langle \hat{H}_c \rangle_{t2}}{\langle 1 \rangle_{t2}}) t_3} e^{-i\frac{\langle \hat{H}_c \rangle_{t2}}{\langle 1 \rangle_{t2}} t_3} \tag{S31}$$

where we have introduced the matrix element $\langle \hat{H}_a \rangle_{t_2} = \langle \bar{\chi}_e^\mu(t_2)| \hat{H}_a |\bar{\chi}_{e'}^\mu(t_2)\rangle$ (with $a \in \mathcal{E}, \mathcal{C}$) weighted by the time dependent overlap $\langle 1 \rangle_{t_2} = \langle \bar{\chi}_e^\mu(t_2)|\bar{\chi}_{e'}^\mu(t_2)\rangle$. The need of this weighting factor will become apparent later.[3]

---

[3]Note that $\langle 1 \rangle_{t_2}$ can in principle be also equal to zero. This can happen when: at least one of the two electronic states e and e' is not populated; the bra and ket WP reside in completely different regions of the coordinate space, and the overlap is therefore null. In both cases one can safely consider the full S29 term to be null, and thus avoiding considering the derivation that follows. Thus, for such derivation, we assume $\langle 1 \rangle_{t_2} \neq 0$.

We thus rewrite the wave function overlap of eq. S29 as:

$$e^{-i\frac{(\langle \hat{H}_c\rangle_{t2} - \langle \hat{H}_e\rangle_{t2})}{\langle 1\rangle_{t2}}t_3} \left\langle e^{i\left(\hat{H}_e - \frac{\langle \hat{H}_e\rangle_{t2}}{\langle 1\rangle_{t2}}\right)t_3} e^{-i\left(\hat{H}_c - \frac{\langle \hat{H}_c\rangle_{t2}}{\langle 1\rangle_{t2}}\right)t_3}\right\rangle_{t_2}$$

(S32)

where we have have adopted the same notation of the main paper in which $\langle \cdots \rangle$ is used to highlight the matrix element: $\langle \bar{\chi}_e^\mu(t_2)| \cdots |\bar{\chi}_{e'}^\mu(t_2)\rangle$. When the maximum value assumed by t_3 is very small (*short time approximation*) the propagator in the last equation is well approximated by the it's second order Taylor expansion in t_3. This leads to:

$$\left\langle \left[1 + i\left(\hat{H}_e - \frac{\langle \hat{H}_e\rangle_{t_2}}{\langle 1\rangle_{t_2}}\right)t_3 - \frac{1}{2}\left(\hat{H}_e - \frac{\langle \hat{H}_e\rangle_{t_2}}{\langle 1\rangle_{t_2}}\right)^2 t_3^2\right] \times \right.$$

$$\left. \times \left[1 - i\left(\hat{H}_c - \frac{\langle \hat{H}_c\rangle_{t_2}}{\langle 1\rangle_{t_2}}\right)t_3 - \frac{1}{2}\left(\hat{H}_c - \frac{\langle \hat{H}_c\rangle_{t_2}}{\langle 1\rangle_{t_2}}\right)^2 t_3^2\right]\right\rangle =$$

$$= \langle 1\rangle_{t_2} - \frac{1}{2}\left[\langle\left(\hat{H}_e - \frac{\langle \hat{H}_e\rangle_{t_2}}{\langle 1\rangle_{t_2}}\right)^2\rangle + \langle\left(\hat{H}_c - \frac{\langle \hat{H}_c\rangle_{t_2}}{\langle 1\rangle_{t_2}}\right)^2\rangle - \right.$$

(S33)

$$\left. -2\langle\left(\hat{H}_e - \frac{\langle \hat{H}_e\rangle_{t_2}}{\langle 1\rangle_{t_2}}\right)\left(\hat{H}_c - \frac{\langle \hat{H}_c\rangle_{t_2}}{\langle 1\rangle_{t_2}}\right)\rangle\right]t_3^2 =$$

$$= \langle 1\rangle_{t_2} - \frac{1}{2}\langle\left[\left(\hat{H}_e - \hat{H}_c\right) - \left(\frac{\langle \hat{H}_e\rangle_{t_2} - \langle \hat{H}_c\rangle_{t_2}}{\langle 1\rangle_{t_2}}\right)\right]^2\rangle t_3^2 - \frac{1}{2}\langle\left[\hat{H}_e, \hat{H}_c\right]\rangle t_3^2$$

where first order term in t_3 have been canceled thanks to the $\langle 1\rangle_{t_2}$ weighting factor. At this point, one would be tempted to express the the t_3^2 dependent terms in analogy to the expression derived for the *population* contribution of the response function. To do so, we define the following quantities:

$$\sigma_{ee'c,m}^2(t_2) = \langle\left[\left(\hat{H}_e - \hat{H}_c\right) - \left(\frac{\langle \hat{H}_e\rangle_{t_2} - \langle \hat{H}_c\rangle_{t_2}}{\langle 1\rangle_{t_2}}\right)\right]^2\rangle$$

(S34)

where the subscript $ee'c$ denotes the fact that this term is computed on the ee' coherence, and to the final state accessed via the interaction with the probe pulse is c. $\sigma_{ee'c,m}^2(t_2)$ is the

equivalent of the variance defined for the ESA population term, and we similarly define the equivalent of the pseudovariance as

$$\tilde{\sigma}^2_{ee'c,m}(t_2) = \sigma^2_{ee'c,m}(t_2) + \left\langle \left[\hat{H}_{e,m}, \hat{H}_{c,m} \right] \right\rangle \tag{S35}$$

At this point, by noticing that:

$$\langle 1 \rangle_{t_2} \left(1 - \frac{\tilde{\sigma}^2_{ee'c,m}(t_2)}{2 \langle 1 \rangle_{t_2}} t_3^2 \right) \sim \langle 1 \rangle_{t_2} e^{-\frac{\tilde{\sigma}^2_{ee'c,m}(t_2)}{2 \langle 1 \rangle_{t_2}} t_3^2} \tag{S36}$$

the overlap of equation S29 can be finally rewritten as:

$$\langle \bar{\chi}^\mu_e(t_2) | e^{i\hat{H}_e t_3} e^{-i\hat{H}_c t_3} | \bar{\chi}^\mu_{e'}(t_2) \rangle = \langle 1 \rangle_{t_2} e^{-i \frac{\langle \hat{H}_c \rangle_{t2} - \langle \hat{H}_e \rangle_{t2}}{\langle 1 \rangle_{t2}} t_3 - \frac{\tilde{\sigma}^2_{ee'c,m}(t_2)}{2 \langle 1 \rangle_{t2}} t_3^2} \tag{S37}$$

which is reminescent of the similar expression obtained for the ESA population term, even if much more involved. Interestingly, if the WP overlap approaches zero, the signal converges to zero. This highlights the fact that the ESA signal of the *coherence* can contribute to the total spectra only when a significant overlap between the e and e' WP is observed.

S6 WPO* vs EGVA LA: high and low frequency modes

Figure s8: Comparison of the EGVA (dashed black line) and analytical LA spectra for a simple model system with the excited state c coupled to a single vibrational mode of frequency ω_m and displacement d_m. ω_m ranges between 0.04 eV and 0.16 eV. The excited state lifetime is set to $\tau_c = 3$ fs. Note how the EGVA approximation improves for lower mode frequencies. Note also that at larger displacements, a quasi-symmetric shape of the spectrum is recovered, which improves the quality of the EGVA approximation also for higher mode frequencies.

In order to assess the quality of the EGVA approximation for high- and low-frequency modes, we built a simple model system with two dipole-coupled electronic states, g and c, with the latter coupled to a single mode, m, with frequency ω_m and displacement (with respect to the ground state minimum) d_m. The mode frequency was varied between 0.04 and 0.16 eV, which correspond to a mode period of ca. 103 and 26 fs, respectively.[4] We show spectra with both $\tau_c = 3$ and $\tau_c = 5$ fs.

The rationale behind this analysis is the following: since the EGVA approximation is based on the time scale separation between core-excited state lifetime (supposed short) and period of the vibrations (supposed longer), the lower the mode frequency, the better the quality of the EGVA approximation should be. Indeed, within the $g - c$ coherence lifetime (whose Fourier transform gives rise to the -linear- spectrum) the system does not present any significant motion of the WP along the low-frequency modes (which might not be strictly true for high frequency modes).

---

[4] Note that in generic molecular systems the highest frequency modes are $C - H$, $N - H$ and $O - H$ stretching, with frequencies of about 3000 cm^{-1}, i.e., $\sim$0.4 eV. We choose not to consider such extremely high frequencies here as these modes do not typically affect the spectra (i.e., the various electronic state PESs are not displaced along these modes).
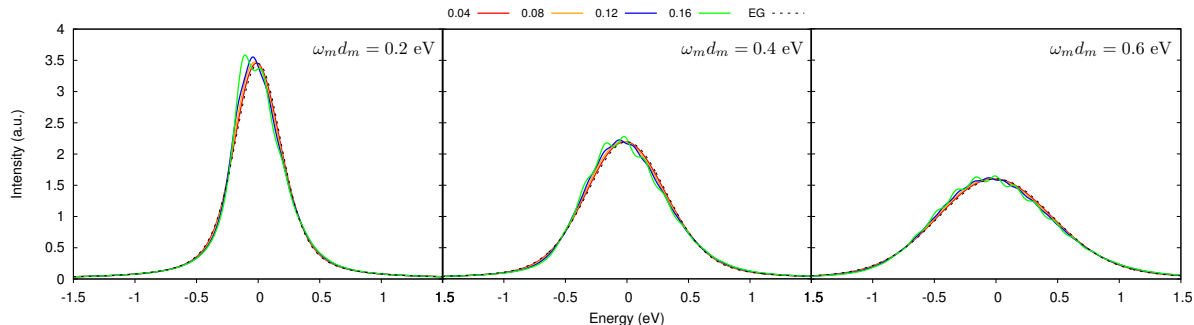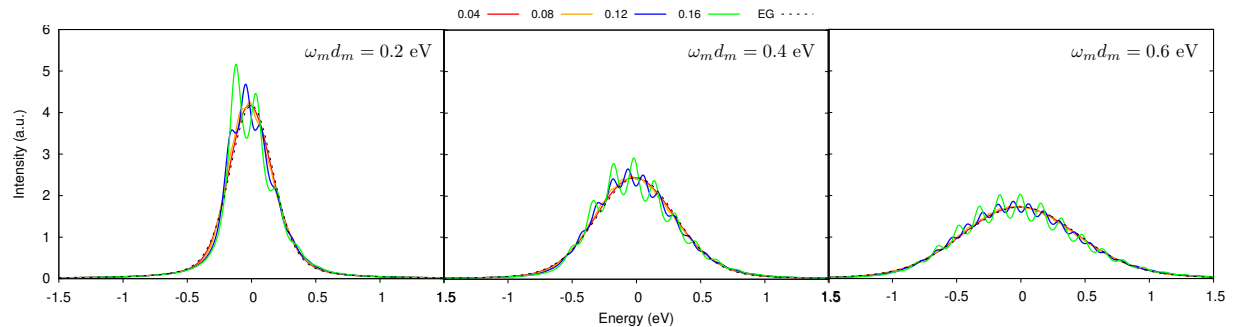
Figure s9: Comparison of the EG (dashed black line) and analytical LA spectra for a simple model system with the excited state c coupled to a single vibrational mode of frequency ω_m and displacement d_m. The excited state lifetime is set to $\tau_c = 5$ fs. Note how the EG approximation improves for lower mode frequencies. Note also that at larger displacements, a quasi-symmetric shape of the spectrum is recovered, which improves the quality of the EG approximation also for higher mode frequencies. The frequency values are reported in eV.

The results are shown in Figure s8 and s9 for three gradient values, $\omega_m d_m$. Note that, for a given gradient value, the EGVA spectrum is independent on the mode frequency. As expected, for $\omega_m = 0.04$ and 0.08 eV, the EGVA linear spectrum is on top of the exact one; $\omega_m = 0.12$ eV (i.e., a period of 34 fs) is nicely reproduced, while significant deviations in the spectral shape are observed for $\omega_m = 0.16$ eV. In particular, the EGVA approximation in unable to reproduce the asymmetry of the spectral shape and the vibronic peaks. Interestingly, for large displacements, whereby a large number of high-energy vibrational states are active, the band-shape symmetry is recovered, and the quality of the EGVA spectra improves at higher mode frequencies. A similar analysis is reported in Figure s9, employing a lifetime $\tau_c = 5$ fs. The differences between the EGVA and the exact spectra are indeed more prominent.

S7 WPO* vs EGVA LA at various values of τ_c

Figure s10 shows the comparison of the XAS spectrum from the 11 brightest $g \to \mathcal{C}$ transitions, for various values of τ_c: 3, 5 and 7 fs. In the main text we presented results that always employ $\tau_c = 3$ fs. It is clear that $\tau_c = 7$ fs represent a limit of extremely high vibronic resolution: since we are not explicitly accounting for static disorder, choosing such a large τ_c value gives spectra that are unrealistically resolved.
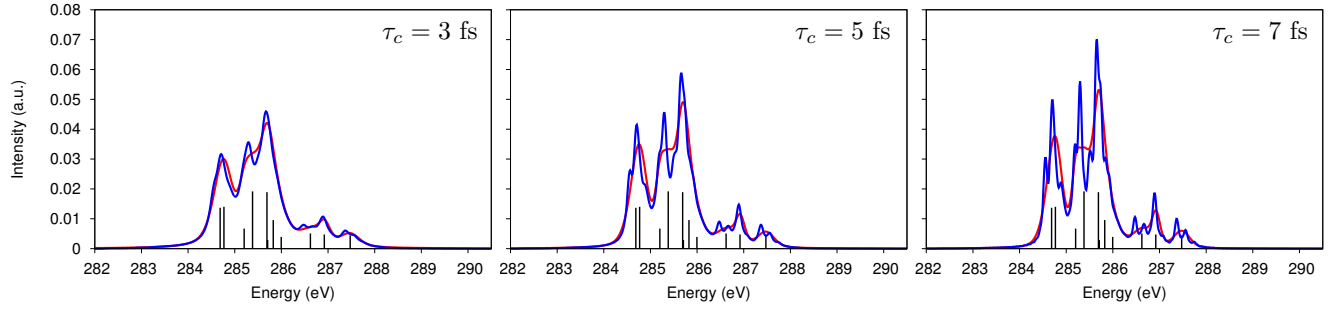
Figure s10: Comparison of WPO* (blue curve) and EGVA (red curve) XAS spectra setting τ_c to a) 3 fs, b) 5 fs and c) 7 fs.

S8   EGVA XAS spectrum: variable vs constant variance

$\tau_c = 3$ fs

$\tau_c = 5$ fs
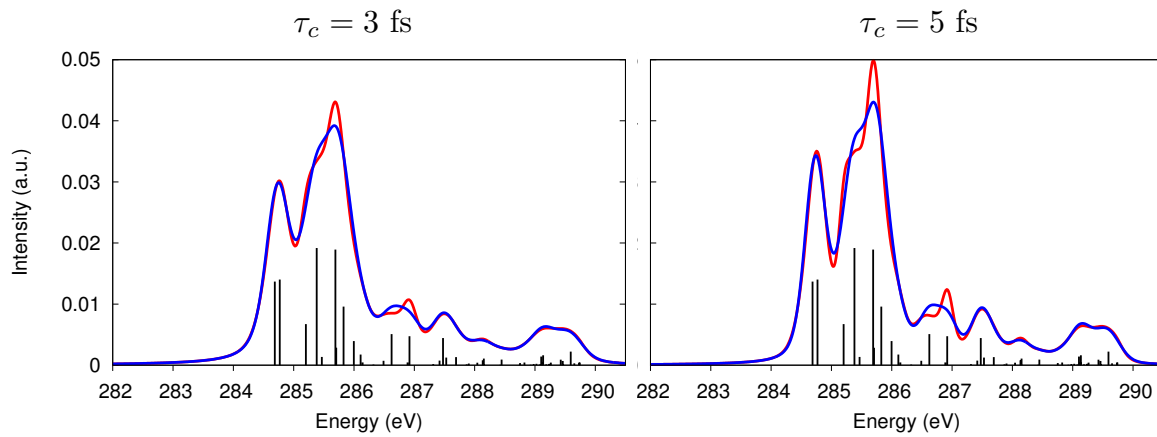
Figure s11: Comparison of total XAS spectra obtained at the EGVA level of theory, employing: the calculated ς_{gc}^2 value (red curve), and a constant ς_{gc}^2 value, set to the average of the variances values, i.e., to 0.019 eV^2. Note that the latter tends to smooth out some spectral features. $\tau_c = 3$ fs (*left*) and 5 fs (*right*).

S9   Energy gaps, energy-gap variances and TDMs

Table S1: Energy gap (eV), Energy gap variance (eV^2) and TDMs for the 75 $g \rightarrow c$ transitions that builds the total pyrene XAS spectrum. First part.

State	EG (eV)	EGV (eV^2)	TDM^2 (a.u.)
c1	284.69	0.026	3.9×10^{-03}
c2	284.77	0.013	4.0×10^{-03}
c3	285.15	0.029	2.3×10^{-13}
c4	285.20	0.004	1.9×10^{-03}
c5	285.32	0.017	1.2×10^{-14}
c6	285.38	0.017	5.5×10^{-03}
c7	285.47	0.011	3.8×10^{-04}
c8	285.69	0.008	5.4×10^{-03}
c9	285.71	0.034	8.2×10^{-04}
c10	285.83	0.018	2.7×10^{-03}
c11	286.00	0.030	1.1×10^{-03}
c12	286.11	0.006	5.0×10^{-04}
c13	286.14	0.035	1.3×10^{-04}
c14	286.32	0.012	3.2×10^{-05}
c15	286.49	0.008	2.1×10^{-04}
c16	286.62	0.030	1.5×10^{-03}
c17	286.89	0.031	1.3×10^{-04}
c18	286.92	0.004	1.4×10^{-03}
c19	286.95	0.009	4.3×10^{-16}
c20	286.95	0.013	3.0×10^{-15}
c21	287.31	0.017	5.0×10^{-05}
c22	287.42	0.013	2.1×10^{-04}
c23	287.48	0.022	1.3×10^{-03}
c24	287.53	0.018	3.6×10^{-04}
c25	287.69	0.014	3.8×10^{-04}
c26	287.74	0.043	3.8×10^{-07}
c27	287.87	0.018	3.9×10^{-05}
c28	287.90	0.014	7.8×10^{-05}
c29	288.04	0.030	1.1×10^{-04}
c30	288.06	0.020	3.5×10^{-13}

Table S2: Energy gap (eV), Energy gap variance (eV^2) and TDMs for the 75 $g \rightarrow c$ transitions that builds the total pyrene XAS spectrum. Second part.

State	EG (eV)	EGV (eV^2)	TDM^2 (a.u.)
c31	288.12	0.007	3.9×10^{-06}
c32	288.13	0.026	4.1×10^{-13}
c33	288.13	0.011	2.3×10^{-04}
c34	288.15	0.029	5.4×10^{-05}
c35	288.15	0.010	3.1×10^{-04}
c36	288.18	0.039	1.1×10^{-13}
c37	288.40	0.027	3.0×10^{-16}
c38	288.42	0.026	1.6×10^{-05}
c39	288.45	0.013	2.6×10^{-04}
c40	288.58	0.021	1.8×10^{-08}
c41	288.71	0.025	4.6×10^{-06}
c42	288.74	0.010	1.2×10^{-05}
c43	288.75	0.012	9.7×10^{-05}
c44	288.78	0.010	4.8×10^{-16}
c45	288.78	0.054	1.3×10^{-05}
c46	288.82	0.026	1.3×10^{-04}
c47	288.84	0.019	9.1×10^{-06}
c48	288.91	0.015	2.0×10^{-15}
c49	288.93	0.042	2.1×10^{-06}
c50	288.93	0.018	2.3×10^{-05}
c51	288.95	0.007	2.1×10^{-06}
c52	288.98	0.014	5.4×10^{-05}
c53	289.02	0.018	6.5×10^{-05}
c54	289.08	0.032	3.4×10^{-05}
c55	289.10	0.022	4.0×10^{-04}
c56	289.13	0.017	1.3×10^{-04}
c57	289.14	0.026	4.6×10^{-04}
c58	289.19	0.016	4.0×10^{-05}
c59	289.23	0.015	3.5×10^{-05}
c60	289.24	0.054	6.6×10^{-05}

Table S3: Energy gap (eV), Energy gap variance (eV^2) and TDMs for the 75 $g \rightarrow c$ transitions that builds the total pyrene XAS spectrum. Third part.

State	EG (eV)	EGV (eV^2)	TDM^2 (a.u.)
c61	289.26	0.012	1.2×10^{-04}
c62	289.27	0.031	7.6×10^{-09}
c63	289.29	0.006	6.3×10^{-13}
c64	289.32	0.010	7.2×10^{-13}
c65	289.34	0.025	6.0×10^{-07}
c66	289.40	0.018	2.6×10^{-06}
c67	289.42	0.012	2.3×10^{-13}
c68	289.43	0.010	2.6×10^{-04}
c69	289.46	0.023	2.0×10^{-04}
c70	289.47	0.014	5.7×10^{-05}
c71	289.54	0.006	2.5×10^{-13}
c72	289.59	0.016	6.4×10^{-04}
c73	289.65	0.011	8.7×10^{-05}
c74	289.73	0.018	1.2×10^{-04}
c75	289.74	0.005	1.3×10^{-04}

Table S4: Energy gap (eV), Energy gap variance (eV^2) (at $t_2 = 0$ fs) and TDMs for the 28 $\mathcal{G}/\mathcal{E} \to \mathcal{C}$ transitions that builds the total pyrene TRXAS spectrum (11 dipole coupled to g, 12 dipole coupled to S_1, and 6 dipole coupled to S_2). Note that the EGV for the $S_1 \to \mathcal{C}$ transitions is not reported as, at time 0, the WP does not reside on the S_1 state (yet).

In. State	Fin. State	EG (eV)	EGV (eV^2)	TDM^2 (a.u.)
g	c_1	284.69	0.026	3.9×10^{-03}
g	c_2	284.77	0.013	4.0×10^{-03}
g	c_4	285.20	0.004	1.9×10^{-03}
g	c_6	285.38	0.017	5.5×10^{-03}
g	c_8	285.69	0.008	5.4×10^{-03}
g	c_9	285.71	0.034	8.2×10^{-04}
g	c_{10}	285.83	0.018	2.7×10^{-03}
g	c_{11}	286.00	0.030	1.1×10^{-03}
g	c_{16}	286.62	0.030	1.5×10^{-03}
g	c_{18}	286.92	0.004	1.4×10^{-03}
g	c_{23}	287.48	0.022	1.3×10^{-03}
S_1	c_3	281.92	-	1.02×10^{-03}
S_1	c_7	282.23	-	1.27×10^{-03}
S_1	c_{28}	284.67	-	1.45×10^{-03}
S_1	c_{29}	284.81	-	9.69×10^{-04}
S_1	c_{30}	284.83	-	2.03×10^{-03}
S_1	c_{38}	285.19	-	1.34×10^{-03}
S_1	c_{48}	285.68	-	1.28×10^{-03}
S_1	c_{52}	285.75	-	1.21×10^{-03}
S_1	c_{53}	285.79	-	1.62×10^{-03}
S_1	c_{63}	286.06	-	1.27×10^{-03}
S_1	c_{64}	286.09	-	2.25×10^{-03}
S_1	c_{65}	286.10	-	1.06×10^{-03}
S_2	c_1	280.94	0.002	5.67×10^{-04}
S_2	c_{22}	283.67	0.001	1.03×10^{-03}
S_2	c_{24}	283.78	0.006	7.24×10^{-04}
S_2	c_{49}	285.18	0.012	1.06×10^{-03}
S_2	c_{50}	285.18	0.019	1.02×10^{-03}
S_2	c_{58}	285.44	0.007	1.95×10^{-03}

S10 Pyrene LA: full mode models vs reduced (15) modes model

Figure s12 shows the comparison between the $g \rightarrow c_1$ transition for the full modes pyrene model comprising 49 normal modes, in red, and the reduced pyrene model comprising 15 normal modes, in blue. The 15 modes were selected to describe the nonadiabatic dynamics that occur in the $\mathcal{E}$ manifold. Interestingly, it turned out that the $g \rightarrow c_1$ is strongly coupled along some of the left out modes. A strategy to account for the left out modes via analytical line-shape functions (while still accounting for the 15 selected modes via the WPO approach) is described in Ref. 4.
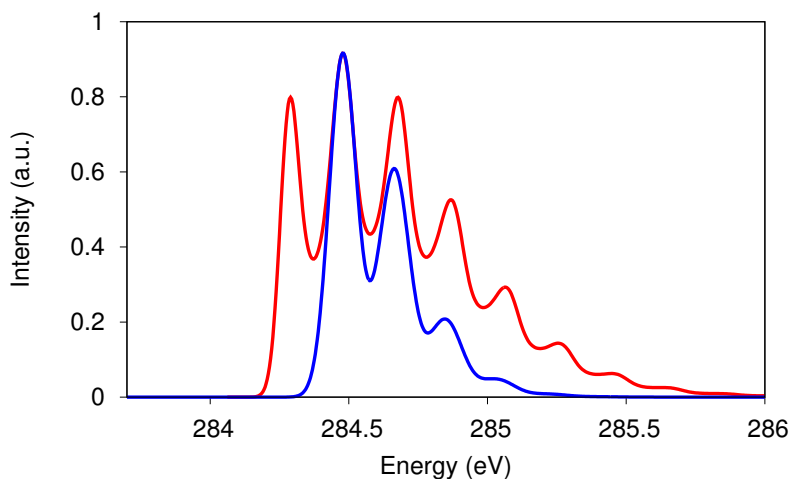

Figure s12: Comparison of the $g \rightarrow c_1$ linear absorption spectrum for the 49 modes pyrene model (red curve) and the 15 modes pyrene model (blue curve). The linear absorption spectra has been normalized to a common value.

S11 Response function Fourier transform: technical details

The valence-to-core transitions cover a range of a few eV, say around 10 eV. This is problematic when performing the FT. In fact, we have a discrete time step of 0.5 fs, that is going to produce aliasing of the spectra every $\sim$10 eV. But now we are precisely in this situation (one does not encounter such a problem in the UV/Vis window, as in that case the spectrum is typically only a few eV wide). This requires special attention when performing the FT.

One might reduce the time-step of the quantum dynamics, at the same time increasing the computational cost of such calculations. We have solved the problem in a different (and more efficient) way: every transition has its own central frequency. We shift this to 2.5 eV. Then we make the FT in between 0 and 5 eV (i.e. around the center of the transition, assuming that 5 eV is a large enough range to cover the full bandwidth of the given transition). After completing the FT, we recenter and add this chunk of spectrum to the total spectrum in the right place, i.e. shifting the center from 2.5 eV to the actual central frequency. Since everything is done in discrete grids, one should take care of the alignment of the grids for different transitions.

References

(1) Segatta, F.; Ruiz, D. A.; Aleotti, F.; Yaghoubi, M.; Mukamel, S.; Garavelli, M.; Santoro, F.; Nenov, A. Nonlinear Molecular Electronic Spectroscopy via MCTDH Quantum Dynamics: From Exact to Approximate Expressions. *Journal of Chemical Theory and Computation* **2023**,

(2) Mukamel, S. *Principles of Nonlinear Optical Spectroscopy*; Oxford University Press: New York, 1995.

(3) Abramavicius, D.; Palmieri, B.; Voronine, D. V.; Šanda, F.; Mukamel, S. Coherent Multidimensional Optical Spectroscopy of Excitons in Molecular Aggregates Quasiparticle versus Supermolecule Perspectives. *Chemical Reviews* **2009**, *109*, 2350–2408.

(4) Montorsi, F.; Aranda, D.; Garavelli, M.; Santoro, F.; Segatta, F. Spectroscopy from quantum dynamics: a mixed wave function/analytical line shape functions approach. *Theoretical Chemistry Accounts* **2023**, *142*.