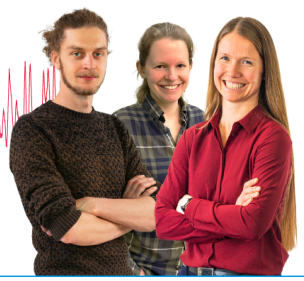
The Journal of Chemical Physics

RESEARCH ARTICLE | FEBRUARY 03 2023

Electron and ion spectroscopy of azobenzene in the valence and core shells

L. Carlini; F. Montorsi; Y. Wu ; P. Bolognesi ; R. Borrego-Varillas ; A. R. Casavola ; M. C. Castrovilli; J. Chiarinelli; D. Mocci; F. Vismarra ; M. Lucchini ; M. Nisoli ; S. Mukamel ; M. Garavelli ; R. Richter ; A. Nenov ; L. Avaldi

J. Chem. Phys. 158, 054201 (2023)
https://doi.org/10.1063/5.0133824
CHORUS



AIP Publishing

The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Electron and ion spectroscopy of azobenzene in the valence and core shells

Cite as: J. Chem. Phys. 158, 054201 (2023); doi: 10.1063/5.0133824
Submitted: 6 November 2022 • Accepted: 16 January 2023 •
Published Online: 3 February 2023

L. Carlini,[1] F. Montorsi,[2] Y. Wu,[3,4] P. Bolognesi,[1,a)] R. Borrego-Varillas,[4] A. R. Casavola,[1] M. C. Castrovilli,[1] J. Chiarinelli,[1] D. Mocci,[3,4] F. Vismarra,[3,4] M. Lucchini,[3,4] M. Nisoli,[3,4] S. Mukamel,[5] M. Garavelli,[2] R. Richter,[6] A. Nenov,[2,a)] and L. Avaldi[1]

AFFILIATIONS

[1] CNR-Istituto di Struttura Della Materia, CNR-ISM, Area Della Ricerca di Roma 1, Monterotondo, Italy
[2] Dipartimento di Chimica Industriale, Università Degli Studi di Bologna, Bologna, Italy
[3] Dipartimento di Fisica, Politecnico di Milano, Piazza Leonardo da Vinci 32, Milano, Italy
[4] CNR-Istituto di Fotonica e Nanotecnologie, CNR-IFN, Piazza Leonardo da Vinci 32, Milano, Italy
[5] Department of Chemistry and Department of Physics and Astronomy, University of California, Irvine, California 92697, USA
[6] Elettra Sincrotrone Trieste, Area Science Park, Basovizza, Italy

[a)] **Authors to whom correspondence should be addressed:** paola.bolognesi@cnr.it and artur.nenov@unibo.it

ABSTRACT

Azobenzene is a prototype and a building block of a class of molecules of extreme technological interest as molecular photo-switches. We present a joint experimental and theoretical study of its response to irradiation with light across the UV to x-ray spectrum. The study of valence and inner shell photo-ionization and excitation processes combined with measurement of valence photoelectron-photoion coincidence and mass spectra across the core thresholds provides a detailed insight into the site- and state-selected photo-induced processes. Photo-ionization and excitation measurements are interpreted via the multi-configurational restricted active space self-consistent field method corrected by second order perturbation theory. Using static modeling, we demonstrate that the carbon and nitrogen K edges of azobenzene are suitable candidates for exploring its photoinduced dynamics thanks to the transient signals appearing in background-free regions of the NEXAFS and XPS.

I. INTRODUCTION

Molecules with structural, optical, or electronic properties that can be tuned and controlled reversibly by light flashes are known as molecular photoswitches and hold a high potential for future nanometer-scale devices.[1] Azobenzene (AB, Fig. 1) and its derivatives are often considered prototypes of fast (picosecond) molecular electronic switches[2] and are among the most frequently studied class of optically switchable substances. The thermodynamically stable *trans*-isomer of AB can be efficiently converted into the *cis*-isomer by irradiation with UV light of about 365 nm, whereas the cis form can be isomerized back to trans with visible light (~435 nm).[3] This property has been widely exploited for molecular motors,[3] triggers for peptide folding,[4] or optical storing applications,[5] to name only a few examples.

These applications call for a complete understanding of the photo-physical properties of AB and its derivatives, and in particular of the electronic and structural changes after light absorption, in the UV, VUV, and soft x-ray regimes, where valence and core ionizations drive new charge redistribution, relaxation, and molecular fragmentation processes. While recent studies have unveiled the primary steps of the AB isomerization reaction in the condensed phase,[6] much less is known about the isolated molecule in the gas phase. High resolution synchrotron radiation spectroscopies can provide the most detailed understanding of the electronic and structural properties of the isolated stable *trans*-isomer when compared with accurate theoretical modeling of time dependent process,[7,8] setting the ground for time-resolved investigations.

X-ray molecular spectroscopy offers numerous unique features, such as high temporal and spatial resolution, atom specificity, and sensitivity to chemical environments, which are currently drawing intensive experimental and theoretical activity. On the experimental

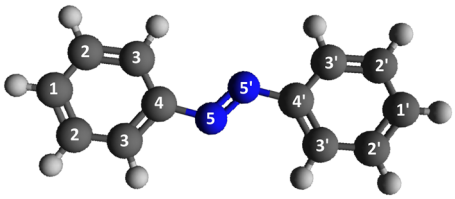


FIG. 1. Azobenzene, AB (CAS: 103-33-3), m = 182 amu, molecular formula $C_6H_5N = NC_6H_5$. The four non-equivalent C atoms on each benzene ring have been labeled 1(1′) to 4(4′).

side, seminal XPS experiments on the ESCA molecule (ethyl trifluoroacetate) have demonstrated the sensitivity of core-spectroscopies, such as x-ray photoemission spectroscopy (XPS)[9] and near edge x-ray absorption fine structure (NEXAFS),[10] to the immediate chemical environment, resulting in several eV chemical shifts. NEXAFS studies of small volatile organic molecules have accessed the K-edges of second row atoms, thereby highlighting the pre-edge window rich in absorption features manifesting the electronic structure of the valence shell.[11–17] On the theory side, a number of electronic structure theory methods have been adapted to study core-excitations and core-ionization.[10,18–30] The restricted active space self-consistent field (RASSCF) method corrected by multi-reference second order theory (RASPT2; i.e., the RASSCF/RASPT2 protocol) from the family of multiconfiguration wave function-based methods has emerged as one of the most accurate methods due to its sub-eV accuracy and spectrum completeness.[31]

Thanks to the promising results obtained for small organic compounds, there is a current drive toward applying the x-ray techniques to larger biologically[32–35] and technologically relevant systems[36,37] as well as to time resolved studies.[20,38–40] These require understanding the sensitivity of core spectroscopy toward changes in the electronic and geometrical structures on the molecular scale by probing it on the atomic scale as well as toward small variations in the chemical environment. AB is an excellent candidate to address these questions and to calibrate experiments and theories. It presents four chemically distinct carbons (labeled 1–4 in Fig. 1) and a nitrogen, two valence electronic states in the UV/Vis—an $n\pi^*$ and a $\pi\pi^*$—with markedly different electronic structures and spectroscopic properties, as well as ultrafast photoinduced dynamics characterized by pronounced structural changes. Time-resolved NEXAFS simulations at the TP-DFT/ΔSCF[20] and RASSCF/RASPT2[41] levels have identified a pre-edge background-free signal at the nitrogen K-edge as a fingerprint of the optically dark $n\pi^*$ state. The existence of this feature is yet to be demonstrated experimentally. Thus, putting the core electronic structure of AB on a firm footing is of paramount importance for future time-resolved studies.

Even though the core spectroscopies have the invaluable advantage of site selectivity thanks to the localized nature of the core electrons, the frontier orbitals are the ones that rule the "chemistry" of a molecule, including its binding properties and interaction with the surrounding environment. Photoelectron spectroscopy[42] directly probes the occupied molecular orbitals, providing ionization energies and electronic symmetries, which are key factors in understanding molecular reactivity and can be used to study the effects of chemical substitution, such as changes of ligands or functional groups. Furthermore, the analysis of the time resolved PES spectra[43] as well as the time coincidence of the photoelectron with the resulting fragment ion from the cation dissociation in photoelectron—ion coincidence (PEPICO) experiments provide the most complete information about the fate of the photoionized molecule.[44]

In all spectroscopic characterizations, the close interplay between theory and experiments is crucial, as the experiments provide electronically sensitive benchmarks against which different computational methods may be tested. Quantum chemistry is invaluable in the assignment and interpretation of spectra.

In this work, we present XPS, NEXAFS, photoelectron-photoion coincidence (PEPICO) measurements, and mass spectra (MS) across the C and N K-edges in the *trans*-isomer of AB. The results are compared to the photoelectron (PES) and photofragmentation spectra of the valence shell and interpreted by electronic structure computations at the RASSCF/RASPT2 level, thus providing a comprehensive characterization of valence and inner shell states. Finally, the suitability of XPS and NEXAFS to study the photoinduced dynamics in AB is assessed computationally by simulation of time-resolved spectra.

II. METHODS

A. Experimental

All experiments were performed at the Gas Phase Photoemission (GAPH) beamline[45] of the Elettra synchrotron radiation facility in Trieste, Italy. This beamline provides tunable, monochromatic, and fully linearly polarized radiation from about 13.5 to 900 eV, by means of five interchangeable gratings. The experimental station is equipped with a hemispherical electron analyzer (VG 220i) and a time-of-flight (TOF) ion mass spectrometer mounted opposite to each other at the magic angle, i.e., ~55° with respect to the polarization axis of the radiation, in order to cancel out possible angular effects on the photoelectrons emitted from the different molecular orbitals. The electron and ion spectrometers have been used either independently of each other for photoelectron spectroscopy (PES and XPS) and photoion mass spectrometry (MS and NEXAFS) or "in conjunction" for photoelectron-photoion coincidence (PEPICO) experiments. All the details of the acquisition and data analysis procedures have been discussed in Refs. 46 and 47 and will be only briefly summarized here.

For the photoelectron measurements, the TOF is switched off in order to have the interaction region free of penetrating fields from the high voltage of the drift tube. The electron analyzer hosts a 2D position-sensitive detector built by Sincrotrone Trieste.[48,49] This detector allows the simultaneous acquisition of electrons over an energy window of about 10% of the pass energy in the analyzer with an energy resolution of about 2% of the pass energy. The PES spectrum was measured at 60 eV photon energy, while the XPS spectra of the C(1s) and N(1s) shells were measured at about 90 eV above their respective ionization thresholds, i.e., at 382 and 495 eV photon energies, respectively, with pass energies of 5 and 30 eV in the C and N cases, respectively. A higher pass energy, and, therefore, a reduced energy resolution, has been used in the case of N(1s) XPS, where only one contribution is expected. The chosen photon energies are sufficiently far from the ionization thresholds so that post-collision interaction effects can be neglected. The experimental

TABLE I. The experimental and calculated binding energies of the valence and core orbitals of AB. For the numbering of the C atoms, see Fig. 1. For orbitals, see Fig. 3. Only cation states with Dyson norms higher than 1×10^{-1} are reported for the PES. A complete list of the valence cation states can be found in Tables SIII–SVIII in the supplementary material.

	Theory (eV)	Exp (eV)	Configuration (weight %)
Valence	8.01	8.43 ± 0.14	$\pi_7^{[1]}$ (74%)
Valence	8.23		$n_2^{[1]}$ (72%)
Valence	8.91	9.30 ± 0.12	$\pi_3^{[1]}$ (73%)
Valence	8.91		$\pi_6^{[1]}$ (73%)
Valence	9.29	9.76 ± 0.21	$\pi_2^{[1]}$ (69%)
Valence	10.98	11–13	$\pi_5^{[1]}$ (47%)
Valence	11.33		$\sigma_6^{[1]}$ (48%)
			$\sigma_8^{[1]}$ (40%)
Valence	11.36		$\sigma_2^{[1]}$ (68%)
Valence	11.54		$\pi_1^{[1]}$ (49%)
Valence	12.17		$n_1^{[1]}$ (43%)
Valence	12.41		$\pi_4^{[1]}$ (20%)
Valence	12.43		$\sigma_1^{[1]}$ (36%)
			$\sigma_4^{[1]}$ (29%)
C1(1s)	290.232	290.26 ± 0.02	$1s_{C1}^{[1]}$ (74%)
C2(1s)	290.344	290.43 ± 0.01	$1s_{C2}^{[1]}$ (74%)
C3(1s)	290.106	290.16 ± 0.01	$1s_{C3}^{[1]}$ (74%)
C4(1s)	290.889	290.88 ± 0.01	$1s_{C4}^{[1]}$ (74%)
N(1s)	405.18	405.53 ± 0.01	$1s_N^{[1]}$ (68%)
N(1s)	407.246	⋯	$1s^{[1]}\pi_7^{[1]}\pi_1^{*[1]}$ (46%)
N(1s)	408.582	409.00 ± 0.04	$1s^{[1]}n_2^{[1]}\pi_1^{*[1]}$ (35%)

binding energy values reported in Table I are obtained by a fit of the XPS data with Gaussian functions, upon the subtraction of a Shirley background. In the C(1s) case, the relative intensity has been constrained to the one determined by the number of non-equivalent C atoms (i.e., 1:2:2:1 for C1:C2:C3:C4, see Fig. 1 for the numbering), and the full width at half maximum of the four states has been constrained to be the same. In the N(1s) XPS case, the FWHM was left as a free parameter and optimized by the fitting procedure to 710 meV.

In the TOF spectrometer, the repeller and extractor electrodes are polarized with anti-symmetric voltages to produce a field up to 370 V/cm. For NEXAFS measurements, the TOF is operated in the total ion yield mode with continuous extraction. The ion yields are measured at different photon energies across the C and N(1s) near-edge regions and are normalized to the photon beam intensity variation read by a photodiode placed at the end of the beamline. For the measurement of the ion flight time, the TOF extraction is operated in pulsed mode triggered by either (i) the signal of a random pulse generator operated at 1-kHz frequency for MS measurements or by (ii) the signal of either a photoelectron or a pulse generator running at 100 Hz for PEPICO measurements. The latter "random extraction pulse" is used for the measurement of the "random mass spectrum," due to ions residing in the interaction region, but uncorrelated with the detected photoelectron. All electrons and ions that hit the detectors are recorded by the acquisition software; however, the PEPICO data analysis selects only electron-single ion events. The "random mass spectrum," after a proper normalization to the number of triggering events, is subtracted from the measured electron-ion coincidence spectrum.[46] MS were measured at several photon energies across the C and N(1s) edge regions, with a step size of 0.5 eV and a typical acquisition time of 1500 s/point. In the data analysis, the background was subtracted from each raw mass spectrum, and then, the intensity of each fragment, identified by a specific mass-over-charge (m/z) ratio, was evaluated by a Lorentzian fit to resolve the contribution of each peak to the total mass spectra.

The PEPICO spectra were measured at a fixed photon energy of 60 eV and an overall energy resolution of 0.4 eV, scanning the electron kinetic energy in steps of 2.7 eV in order to cover the binding energy region 7.5–26 eV with sufficient overlap between adjacent acquisition windows. An Al filter was inserted along the photon beam pathway to filter out second order radiation. The spectra have been measured in several scans that allow us to check the stability of the measurements, for a total acquisition time of about 18 h. The first onset of the main fragments has been extracted from a linear fit of their respective PEPICO branching ratios, and these values have been compared with the appearance energy (AE) from literature[50] in Table II. These two quantities do not necessarily coincide due to the fact that at variance with a "conventional" AE measurement where the electron/photon energy is scanned, the PEPICO experiments are performed at fixed photon energy so that the involvement of several resonant/autoionizing states is not possible, and this can alter both the ionization and fragmentation processes.

The XPS and NEXAFS spectra were calibrated according to well-known Refs. 51–54 in CO_2: $C(1s)^{-1}$ at 297.6 eV, $O(1s)^{-1}$ at 541.3 eV, $C(1s) \rightarrow \pi^*$ at 290.77 eV, and $O(1s) \rightarrow \pi^*$ at 535.4 eV and Refs. 55 and 56 in N_2: $N(1s)^{-1}$ at 409.9 eV and $N(1s) \rightarrow \pi^*$ at 400.87 eV. Calibration gases were inserted as diffuse gases in the end-station hosting the electron and ion spectrometers for the calibration of the XPS measurements or in a separate vacuum chamber equipped with a channeltron detector for total ion yield measurements and positioned along the beam path for the calibration of the NEXAFS spectra. In both cases, the calibration gas and AB spectra were measured simultaneously. The PEPICO spectra were calibrated

TABLE II. The main fragments observed in the PEPICO measurement are shown in Fig. 2(b), along with their proposed assignments that are consistent with Refs. 50 and 74, the AE value from literature data, and the first onset from the PEPICO branching ratio, BR, of Fig. 2(b) (estimated uncertainty are around 0.2–0.3 eV). (⋆) Ref. 50; (◇) Ref. 74; and (•) not observed/discussed in Refs. 50 and 74.

m/z	Proposed ⋆,◇ assignment	PEPICO BR onset (eV)	AE ⋆ (eV)
182^+	$C_{12}H_{10}N_2^+$	7.8	8.5 ± 0.05
154^+	$C_{12}H_{10}^+$	9.9	11.6 ± 0.2
153^+	$C_{12}H_9^+$	11.9	12.4 ± 0.1
152^+	$C_{12}H_8^+$	13.2	14.7 ± 0.05
105^+	$C_6H_5N_2^+$	9.9	9.8 ± 0.1
77^+	$C_6H_5^+$	10.2	11.9 ± 0.1
51^+	$C_4H_3^+$ (•)	17.7	⋯

against the ground state of the water photoelectron spectrum at 12.62 eV,[57] which is clearly identifiable in the PEPICO spectrum of fragments m/z 18.

The *trans*-azobenzene, AB molecule, Fig. 1, is commercially available with 99% purity and appears as orange-red crystals at ambient temperature and pressure. It has been introduced to the vacuum chamber in a crucible and sublimated at room temperature, with a residual pressure of 7×10^{-7} mbar in the vacuum chamber. To maximize the gas density of the sample in the interaction region, the exit channel of the crucible is equipped with a needle that fits into the extraction region of the TOF. The tip of the needle has been positioned a few mm below the photon beam.

B. Theoretical methods

All quantum chemical computations were performed using the QM software OpenMolcas[58] starting from an MP2 optimized geometry. Dyson norms and transition dipole moments were calculated with the RAS sate interaction routine (RASSI) implemented in OpenMolcas.[59]

Electronic structure computations were performed at the RASSCF/RASPT2 level of theory.[60,61] This theoretical methodology, belonging to the family of multiconfigurational wave-function based methods, is a generalized form of the Complete Active Space SCF (CASSCF) method that allows a restricted number of holes and electrons in certain subspaces (i.e., RAS1 and RAS3) in contrast to the complete set of permutations of the electrons in the active orbitals of CASSCF adopted for the RAS2 subspace.

As PES, XPS, and NEXAFS probe different manifolds of the electronic structure, each experiment was simulated with an independent calculation using different parameters in the RASSCF/RASPT2 protocol. Throughout, the C_{2h} molecular point group with irreducible representations (IrRep) A_g, B_g, A_u, and B_u is employed, and the construction of the active space (AS) was made in accordance with the four IrRep.

1. Photoelectron spectroscopy (PES)

The simulation of valence shell photo-ionization spectroscopy requires a calculation of the manifold of doublet cationic states and the intensity of ionization with respect to the neutral electronic ground state. The manifold of cation states was obtained by imposing a positive charge and setting the spin quantum number to 2. Thereby, two different AS compositions were used:

- An AS consisting of 17 electrons in 16 orbitals partitioned in the following way: RAS1 was left empty; RAS2 includes the two nitrogen's lone pairs (n) belonging to A_g and B_u IrRep and seven occupied π orbitals belonging to B_g and A_u IrRep; and RAS3 includes seven virtual π^* orbitals belonging to B_g and A_u IrRep. Up to 4 electrons were allowed to be excited from RAS2 in the RAS3 subspace. We denote this active space as RAS(17,0,4;0,9,7) where the first three indices identify: the number of electrons, the number of holes in RAS1 and of excitations in RAS3, respectively, while the last three indices denote the number of orbitals in each subspace. In order to cover a broad energy window, 33 cation states of different symmetry were computed. In particular, seven states of A_g symmetry, nine states of B_g symmetry, ten states of A_u symmetry, and seven states of B_u symmetry have been computed in four independent calculations with the state-average version of RASSCF (i.e., SA-RASSCF). This is the minimal number of states required to capture all nine singly ionized states (of n^+ and π^+ characters). The remaining 24 states represent shake-up features in which the ionization is accompanied by a valence electronic transition.
- An AS consisting of ten occupied σ orbitals of A_g and B_u IrRep in RAS2 and ten virtual σ^* orbitals of A_g and B_u IrRep in RAS3. Up to 4 electrons were allowed to be excited from RAS2 in the RAS3 subspace. We denote this active space as RAS(20,0,4;0,10,10). 5 of A_g symmetries and 5 B_u symmetries have been computed in two independent calculations at the SA-RASSCF level. This allowed us to describe ten singly ionized σ^+ states.

Perturbation correction to the RASSCF solutions was obtained using the multi state (MS) flavor of RASPT2 applying an imaginary shift of 0.5 and setting the ionization potential electron affinity IPEA to 0.0.

In a separate calculation, the neutral ground electronic state was calculated with the AS compositions used to obtain the doublet manifolds in a state-specific fashion (i.e., SS-RASSCF), and its energy was corrected by a second order perturbation correction.

Finally, the ionization intensities from the neutral ground state were evaluated within the sudden approximation using the Dyson orbitals formalism. This constitutes a very efficient approximation that has been shown to be comparable with a higher level of theory.[62–64]

2. X-ray photoelectron spectroscopy (XPS)

The simulation of the core photo-ionization spectroscopy requires a calculation of the manifold of doublet cation states exhibiting an unpaired electron in a carbon or nitrogen core-orbital and the intensity of ionization with respect to the neutral electronic ground state. To that aim, the RAS(17,0,4;0,9,7) used to simulate valence shell PES was augmented to include, one at a time, a core orbital of interest, which is placed inside RAS1 and kept frozen, giving rise to a RAS(19,1,4;1,9,7).

The manifold of core cation states was obtained within the core-valence separation (CVS) framework through a projection technique named HEXS[65] (highly excited states), which sets the CI coefficients for those configuration state functions with a maximum occupation from a given subspace to zero, thus effectively projecting them out of the wave-function. In particular, by applying this procedure to the RAS1 subspace, the core-excited states for the nitrogen and the four chemically non-equivalent carbon atoms could be computed. We note that a thorough benchmarking of the protocol has demonstrated a sub-eV accuracy of the predicted absolute transition energies.[31] For the nitrogen K-edge state, averaging over 15 states is employed in order to describe both the main signal and the shake-up satellites observed experimentally, while for the carbon K-edge, single state calculation for each non-equivalent carbon turned out to be enough to describe the experimental spectra.

Benchmark calculations were performed to compare the delocalized (canonical orbitals) and localized (obtained through Cholesky localization[66] of the canonical orbitals) descriptions of the core orbitals. The outcomes of these tests are reported in the supplementary material and show that better results are obtained with

localized core orbitals. We note that the molecular symmetry is lost upon localization, and, as a consequence, calculations had to be performed with the C_1 point group.

Perturbation correction to the RASSCF solutions was done using the single state (SS) flavor of RASPT2 with an imaginary shift of 0.5 and IPEA of 0.0. The number of frozen orbitals was set to zero.

In a separate calculation, the neutral ground electronic state was calculated with AS composition used to obtain the doublet manifold in a state-specific fashion, and its energy was corrected by a second order perturbation correction.

The XPS peak intensity was evaluated within the aforementioned sudden approximation.[62–64]

3. Near edge x-ray absorption spectroscopy (NEXAFS)

Simulation of the core-excited spectroscopy requires a calculation of the manifold of singlet neutral states exhibiting an unpaired electron in a carbon or nitrogen core-orbital and the oscillator strength with respect to the neutral electronic ground state. The AS for these simulations was constructed following the same considerations discussed in the XPS section (i.e., one 1s orbital at a time in RAS1, an occupied valence orbital in RAS2, and virtual orbitals in RAS3), and it is denoted RAS(20,1,4;1,9,7).

For the simulation of the nitrogen K-edge, an averaging over ten states is sufficient to capture both the main signal and the shake-up satellites (i.e., mixed double excitations in which a core-excitation is accompanied by a valence electronic transition), while for each non-equivalent carbon K-edge, 15 states had to be considered.

Second order perturbation theory is then applied on top of the RASSCF wavefunction in a multi-state fashion for nitrogen while an extended multi-state flavor of CASPT2 (XMS-CASPT) was applied for carbon.

In a separate calculation, the neutral ground electronic state was calculated with RAS(20,1,4;1,9,7) in a state-specific fashion, and its energy was corrected by a second order perturbation correction.

The transition intensity between the ground and the core-excited states was modeled as proportional to the squared transition dipole moments (TDMs).

4. Technical details of the XPS and NEXAFS simulations

Scalar relativistic effects were taken into account via a second order Douglass–Kroll–Hess Hamiltonian in combination with a relativistic atomic natural orbital basis set[67,68] (ANO-RCC) with contractions 4s3p2d on carbon, 4s3p2d1f on nitrogen, and 2s1p on hydrogen. A density-fitting approximation, known as Cholesky decomposition,[69] of the electron repulsion integrals has been used to speed up the calculation of two-electron integrals.

Additional details, such as the selected active spaces and the leading configuration state functions of each excited/ionized state, are reported in the supplementary material.

5. Fragmentation

To investigate the photo-fragmentation process quantum chemical calculations have been performed with Density Functional Theory (DFT). The geometries were optimized using the Becke, three-parameter, Lee-Yang-Parr (B3LYP) functional with the 6-311++G** basis set, while single-point energy calculations were performed at CCSD/6-311++G** level. All calculations were performed using the GAUSSIAN 09 suite of programs.[70]

III. RESULTS AND DISCUSSION

A. Valence shell photo-ionization and photo-fragmentation

The PES spectrum reported in Fig. 2(a) displays a distinct structure with at least two experimental bands in the binding energy, BE, region 8–10 eV, which are separated by a small but well defined gap from several broader bands (BE > 12 eV). The lowest energy band of the PES spectrum, centered at around 8.43 ± 0.14 eV in the present measurements, is assigned to the ionization of two molecular orbitals (Fig. 3): a superposition of the lone pair atomic orbitals of the two nitrogen atoms (denoted n_2) and the highest occupied π-orbital delocalized over the entire molecule (denoted π_7). As the $n_2\pi^*$ electronic state of the neutral system is energetically lower than the $\pi_7\pi^*$ state (thus suggesting a smaller n_2–π^* energy gap compared to π–π^*), it would be natural to assume that the n_2 lone pair is easier to ionize compared to the π-orbital. In fact, this argument

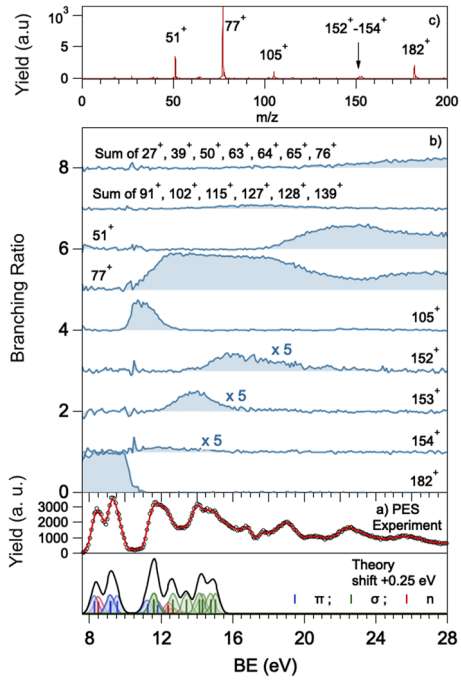


FIG. 2. (a) The AB PES spectrum measured at 60 eV photon energy and 20 eV pass energy in the electron analyzer, with an overall energy resolution of about 0.4 eV. Present theoretical results are reported as vertical sticks and convoluted Gaussian functions in the bottom part of the panel (a). (b) Branching ratios of the main fragments from PEPICO spectra. The plots of the different fragments have been shifted up by 1 unit for clarity in the display. PEPICO fragments with intensity <5% are not reported individually but have been added up according to the BE range where they provided an appreciable contribution. These two groups are reported on the top of the panel, where the contribution fragments have been indicated (b). (c) Sum of all PEPICO mass spectra from the ionization threshold to 26 eV binding energy. The narrow positive signal at 10.2 eV is a spike due to the calculation of the BR in a region of vanishing signal.

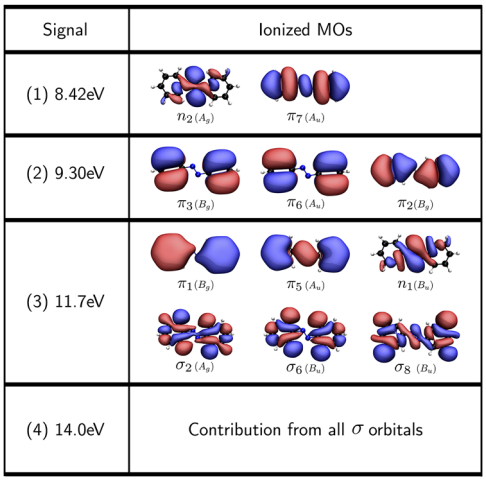


FIG. 3. Valence molecular orbitals that are ionized and give rise to the first four bands in the photoelectron spectrum in Fig. 2(a).

has been used previously by others reporting the PES of AB.[71,72] Our calculations demonstrate, however, that the π_7^+ state (8.01 eV, Table I) is ~0.2 eV lower in energy compared to the n_2^+ state (8.23 eV, see Table I). This can be rationalized with the delocalization of the positive charge over the entire molecule upon ionization from the π-orbital in contrast to its localization on the nitrogen atoms in the case of ionization from the lone pair. The lowest band in the PES spectrum has been shown to be very sensitive to functional groups.[73] We attribute this observation to the interaction of electron-donating and -withdrawing substituents with the aromatic system expected to affect the π_7^+ rather than the n_2^+ state.

The second band centered around 9.3 ± 0.12 eV with a shoulder at 9.76 ± 0.21 eV in the present experiments is assigned as the ionization from several π orbitals. In particular, the maximum of the band is determined by two degenerate ionized states (at 8.91 eV, Table I) involving orbitals π_3 and π_6 localized over the benzene rings (Fig. 3), whereas the shoulder arises from the ionization from π_2 (9.29 eV, Table I) delocalized over the entire molecule.

A number of ionized states contribute to the binding energy region above 10 eV (listed in Tables SIII–SVIII in the supplementary material). The band between 11 and 13 eV is due to the ionization of lower lying π-orbitals and the higher lying σ-orbitals (see Fig. 3 and Table I). It also exhibits several less intense satellite features arising through valence $\pi \rightarrow \pi^*$ excitations accompanying the ionization. Ionization from a second, energetically lower lying, superposition of nitrogen lone pairs (denoted n_1) also falls in this region contributing to the shoulder at 12.5 eV. The fourth band in the PES spectrum spreads between 13 and 16 eV and is dominated by ionization from multiple σ-orbitals (listed in SVII–SVIII in the supplementary material).

The AB photo-fragmentation mass spectrum reported in Fig. 2(c) is the sum of all the PEPICO mass spectra measured in the binding energy region 8–26 eV at a photon energy of 60 eV. Even though the contribution of double ionization fragmentation processes cannot be ruled out, especially in the high binding energy region, the spectrum is expected to be dominated by single ionization events, and no clear evidence of doubly charged ions has been detected. Further than the parent ion, which is clearly visible at m/z 182 and also indicated with the notation 182^+, the spectrum is dominated by a relatively small number of fragments with branching ratios >5%, namely 77^+ and 51^+, and among the most intense, 105^+ and the group 154^+–152^+. The proposed assignment is reported in Table II, together with a previous determination of appearance energy by electron impact ionization experiments[50] and the present estimate of the first onset in the PEPICO branching ration, see Fig. 2(b).

According to the PEPICO measurements, the parent ion (182^+) is characterized by high stability over the entire binding energy range of 8–10 eV, where it is the only observed charged species. The cis and trans structures of both the neutral and charged parent molecule have been analyzed. In both cases, we observe a stabilization of the trans form, which is 0.5 and 0.28 eV more stable than the cis form in 182 and 182^+, respectively. Exploring the energy potential due to torsion of dihedral angle C4–N5–N5′–C6 going from *trans* to *cis*, we identified the transition state of this reaction at an energy of 8.57 eV in 182^+. Therefore, based on these results, we can safely assume that while the neutral gaseous species, desorbed at room temperature, is in a pure trans conformation, the *trans*–cis isomerization is energetically open in the cation at about 600 meV above the ionization threshold.

In the binding energy region >10 eV, the most intense fragments are 105^+ and 77^+, with a very tiny contribution from 154^+. The first two, lighter, fragments can be assigned to the two complementary fragments $C_6H_5N_2^+$ and $C_6H_5^+$, respectively, corresponding to the detachment of a phenyl ring from the molecule and the charge on either of the two moieties. Their first onsets in the PEPICO experiments are very close (at least within the present experimental uncertainty), but they display a very different behavior vs binding energy, with 105^+ being more state-selective, i.e., existing in quite narrow binding energy region, while 77^+ can be found over the entire binding energy region investigated above its formation. This suggests that 77^+ is a highly stable species, requiring a large amount of energy for subsequent fragmentation. It is interesting to explore the energy and formation pathways of these fragments along the potential energy surface. Starting from the trans-configuration, 182^+ can produce 105^+ + 77 with AE(105^+) = 9.53 eV, or 105 + 77^+ with AE(77^+) = 12.04 eV, a value much higher than the experimental one. Considering the 105^+ fragmentation into 77^+ + 28, we calculate AE(77^+) = 10.91 eV, in better agreement with the experimental value (seeFig. 4). Thus, the calculations suggest that, according to the minimum energy path, 77^+ derives from the 105^+ fragment and not directly from the 182^+. This is consistent with MIKE experiments[74] where, at least in the threshold energy region, these two channels are strictly related by subsequent losses: $C_{12}H_{10}N_2^+(182^+) \rightarrow C_6H_5N_2^+(105^+) + C_6H_5$ and $C_6H_5N_2^+(105^+) \rightarrow C_6H_5^+(77^+) + N_2$, rather than being independent channels.

The group of fragments 154^+–152^+, of much smaller intensities with respect to 105^+ and 77^+, is very intriguing. Being separated from each other by one mass unit and about 2 eV in the energy onset, they are clearly related by successive[74] hydrogen losses from 154^+. This fragment from both our calculations and Refs. 50 and 74 is attributed to $C_{12}H_{10}^+$, i.e., to the loss of N_2 from the parent

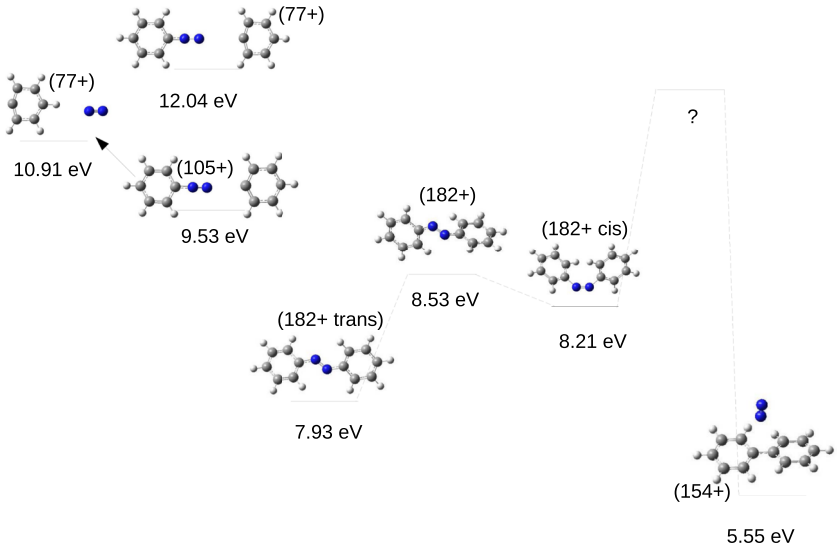


FIG. 4. The most relevant features in the potential energy surface of AB are discussed in the text. On the right side, parent ion conversion between trans and cis form is considered a plausible preliminary step for N_2 elimination. The existence of an energy barrier in this latter process, whose transition state has not been identified, is hypothetical. On the left side, the fragmentation of the trans-isomer leads to 105^+ and 77^+ formation, with no barrier processes.

cation. However, its mechanism of formation, with N_2 extraction, and structure remain controversial. Natalis and Franklin[50] considered the two possible structures: acenaphthene, a polycyclic aromatic hydrocarbon consisting of naphthalene with an ethylene bridge, and biphenyl, an aromatic hydrocarbon consisting of two connected phenyl rings. Based on a thermochemical analysis, they considered biphenyl as the most likely structure but do not provide any information on the possible mechanisms of N_2 extraction with recombination of the phenyl rings. According to our calculations, from an energetic point of view, the fragmentation of the parent ion into $154^+ + N_2$ is very favorable because the formation energy of 154^+ without taking into account any energy barrier is 5.55 eV, indicating an extremely stable fragment that would most likely require overtaking a significant energy barrier for further fragmentation. However, the low intensity of fragment 154^+ in the PEPICO spectra suggests that dynamical considerations may play a crucial role in turning this process into an improbable one. It is, therefore, likely that the N_2 elimination involves a multi step process with energy barriers to be overcome. As a plausible mechanism, we propose that the N_2 elimination may be due to a preliminary trans to cis conversion of 182^+ whose transition state at 8.53 eV is energetically achievable below the AE of 154^+. Then from *cis*-182^+, either a sequential or concerted reaction, favored by the close proximity of the two phenyl rings, may lead to the elimination of N_2 with phenyl ring recombination. For the sequential case, the transition state for the N–C bond breaking has been calculated at 9.14 eV. However, the following steps of ring recombination could not be investigated with the present theoretical methodology.

Fragment 51^+ has not been discussed before. Its first onset at 17.7 eV in the PEPICO spectra indicates that its origin is related to relatively high energy processes, which could see the involvement of inner-valence molecular orbitals and, at higher energies, the fragmentation of doubly charged ion states, which are estimated[75] to be in the range 20–21 eV. As for the fragmentation path, the subsequent dissociation of fragments 152^+ and/or 77^+, which, in the valence region,[74] are the evolution of the initial steps of fragmentation from

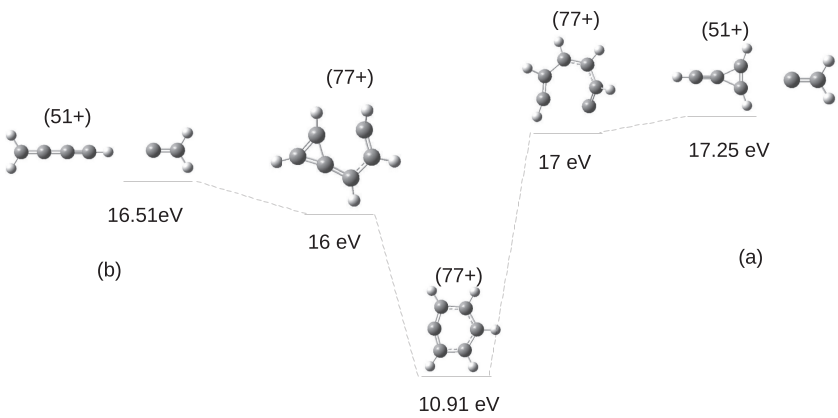


FIG. 5. Production of 51^+ from (a) the breaking of C3–C4 bonds and (b) the breaking of C1–C2 from 77^+.

154^+ and 105^+, respectively, can be considered. In any case, the formation of fragment 51^+ requires the fragmentation of the phenyl ring and some rearrangement, most likely favored by an extensive hydrogen scrambling[74] that could lead to the formation of cyclobutadiene ($C_4H_3^+$). The PEPICO BR of these fragments, Fig. 2(b), may provide some hints on the formation of 51^+. Indeed, the BR of both 152^+ and 77^+ decrease in the same binding energy region that sees the rise of 51^+, i.e., 18–22 eV. However, while this decrease coincides with the complete disappearance of the low intensity fragment 152^+ from the PEPICO spectra, 77^+ can be seen rising again in the higher binding energy region, most likely via a double ionization process, as suggested by the broadening of its line shape in the PEPICO spectra measured at BE > 19.5 eV. Then, the PEPICO experiments suggest that both precursor ions (152^+ and 77^+) could be considered in the formation pathway of 51^+.

Considering the low BR value of fragment 152^+ and the negligible contribution it may give to the BR of 51^+, we theoretically investigated the 51^+ formation starting from 77^+ and leading to $51^+ + 28(C_2H_2)$. In the final form, we obtain for 51^+ either a three-member ring structure with a formation energy of 14.70 eV or a linear structure with a formation energy of 15.44 eV, see right hand side of Fig. 5. Considering the energy barrier for the ring breaking, we studied two possible channels: (a) the breaking of the C3–C4 bond, with a transition state of 17 eV, and (b) the breaking of C1–C2, with a transition state of 16.0 eV. It is interesting to observe that in both cases, either directly or via rearrangement in the case of C3–C4 breaking, it is the CH_2C fragment to be produced rather than C_2H_2. Taking into account this new product, we obtain a formation energy of 51^+ starting from the breaking of C3–C4 of 17.25 eV, while the formation energy starting from the breaking of C1–C2 is 16.51 eV. Both these channels are possible, as consistent with the experiments.

B. Core-ionization and photo-fragmentation

The experimental and theoretical XPS spectra for C(1s) and N(1s) are shown in Figs. 6(a) and 6(b), respectively, while the corresponding peak assignments are reported in Table I.

Depending on their position on the ring, the 1s binding energies of the carbon atoms present positive/negative shifts compared to the benzene molecule, where the adiabatic ionization energy is 290.26 eV.[76] The larger shift is expected in the substitution site, i.e., C4 in the present notation (Fig. 1). The negative inductive effect (-I) exerted by the nitrogen lowers the electron density around C4, resulting in higher binding energy. As a consequence, the C4 band appears most blue-shifted at 290.88 eV in the XPS spectrum (cyan peak in Fig. 6). While the effect of the electronegativity of the substituent atom decreases as the distance increases, the inductive effect alone cannot explain the ordering of the binding energies of the four carbon atoms. In fact, our calculations show that the binding energy decreases in the order C2 > C1 > C3. This ordering can be rationalized by taking into consideration the nitrogen positive mesomeric effect (+M), which restores electron density back in the phenyl ring through the π-orbitals. The + M effect decreases in the order ortho (C3) > para (C1) > meta (C2) with respect to the substituted atom, thus resulting in the highest electron density and, conversely, lowest binding energy for C3 (red peak in Fig. 6). A similar effect has been reported for halopyrimidines.[77]

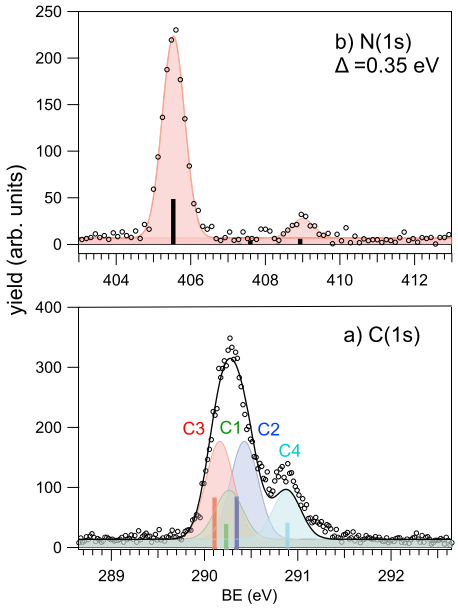


FIG. 6. (a) and (b) The nitrogen (top) and carbon (bottom) 1s XPS spectra of AB (open dots) measured at 382 and 495 eV photon energy, respectively, experimental fits (full areas), and theoretical predictions (vertical bars). A shift of 0.35 eV has been applied to the N(1s) theoretical results to match the experimental features.

In the N(1s) XPS spectrum, Fig. 6(b), the main band is found at 405.53 eV, in good agreement with the theoretically predicted ionization energy of 405.18 eV. A second less intense feature appears ~3.5 eV above the main band, i.e., at 408.58 eV. XPS measurements on AB layers adsorbed on layered crystals have assigned a similar feature to a satellite structure corresponding to a shake-up $\pi \rightarrow \pi^*$ transition.[78,79] Our calculations demonstrate that the corresponding shake-up $\pi \rightarrow \pi^*$ feature should give rise to a less intense satellite ~2.0 eV above the main band (407.25 eV), which remains experimentally unresolved. Instead, the satellite at 408.58 eV is due to a shake-up $n_2 \rightarrow \pi^*$ transition accompanying the 1s ionization. The 3.5 eV energy required for this shake-up with respect to the bare N(1s) ionization is about 1 eV higher compared to the energy required to excite the same transition in the neutral system ($n_2\pi^*$ has a maximum in the visible spectrum at ~2.7 eV[80]). It is energetically more demanding to excite an electron from the lone pair into the aromatic system and, thus, to displace electron density away from the nitrogen atoms when the density is simultaneously depleted through ionization.

C. The core near-edge absorption

The experimental and theoretical NEXAFS spectra for C(1s) and N(1s) are reported in Figs. 7 and 8, respectively, while the corresponding peak assignments are reported in Table III. The carbon spectrum exhibits a complex multi-peak pre-edge structure with three well separated bands peaking at 285.3, 287.5, and 289.5 eV superimposed on a linearly increasing ionization continuum. Each of these three peaks arises due to excitations from multiple carbon

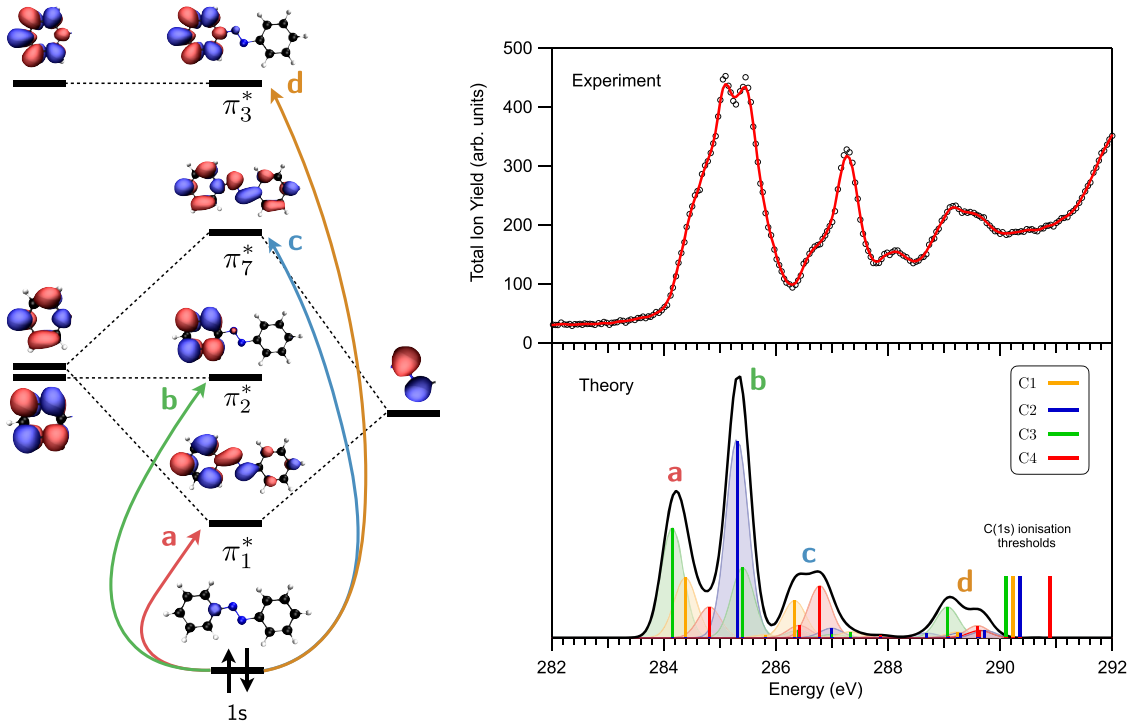


FIG. 7. Molecular orbitals diagram (left) and carbon K-edge NEXAFS spectra (right) for AB. The experimental data are reported in the top panel as black dots with a superimposed three-point smooth line in red, while the theoretical results, convoluted by Gaussian functions of FWHM = 0.7 eV, are depicted in the bottom panel. Individual contributions from different carbon 1s orbitals are also shown with different colors. Labels a, b, c, and d indicate signals associated with transitions to different virtual orbitals. The C(1s) ionization threshold is also highlighted with vertical lines.

centers distinguishable by the calculations. In particular, the lowest band at 285.3 eV exhibits a fine splitting and a shoulder on the low energy side at 284.5 eV. The shoulder forms due to transitions from C1, C3, and C4 to the lowest virtual π_1^* orbital (Fig. 7).[81] Interestingly, the transition energy for the sites increases in the order C3 > C1 > C4, which matches with the order of the XPS binding energies [Fig. 6(a)]. Transitions to the second virtual π_2^* orbital give rise to the main peak. Only C2 and C3 contribute to this peak due to the shape of the virtual orbital having no coefficient on C1 and C4 (Fig. 7). Notably, the transitions into the two π^* orbitals appear as distinct peaks in the calculations while they merge into a single band in the experiment. The second band at 287.5 eV, which also exhibits a shoulder on the low energy side, arises due to transitions from all four carbon core orbitals to the virtual orbital labeled π_7^*. In fact, orbitals π_1^* and π_7^* are formed through bonding (π_1^*) and anti-bonding (π_7^*) interactions of the phenyl rings with the central N_2 unit. The third band at 289.5 eV is dominated by shake-up features due to transitions into orbital π_3^*, i.e., the completely anti-bonding phenyl orbital (Tables SIX–SXII in the supplementary material). In agreement with the above analysis, the peaks at 285 and 289.5 eV arising due to core transitions into the weakly perturbed phenyl orbitals π_2^* and π_3^* have been reported for bare benzene chemisorbed on Cu(110).[82]

Overall, the calculations show that the core orbitals of the four nonequivalent carbon atoms lie energetically close, as can be also inferred from the XPS spectrum (Fig. 6). Hence, the multi-peak nature of the NEXAFS spectrum reflects rather the energetic structure of the virtual orbital manifold [Fig. 7(a)].

The nitrogen spectrum is characterized by one dominant pre-edge band at 398.57 eV assigned to the 1s→ π_1^* transition. Notably, no further intense transitions into the π-system are observed, and as among the virtual π^*-orbitals, only π_1^* has a significant coefficient on the nitrogens (Fig. S1 in the supplementary material), which is a requirement for a finite dipolar coupling.[31] A less intense feature between 400.5 and 402.5 eV is due to a shake-up feature characterized by the simultaneous excitation of a 1s and a valence π electrons in the virtual orbitals.

In the K-edge regions around the C and N thresholds whose main NEXAFS features have been discussed above, several MS spectra have been acquired at regular steps. These background-subtracted mass spectra are shown in Fig. 9 where they are compared to the mass spectra recorded at 60 eV (black curve), i.e., the photon energy used for the PEPICO measurements.

In both cases where the mass spectra have been measured above 1s ionization thresholds, the contributions of the parent ion and the m/z > 100 fragments are vanishing. These spectra are dominated by the $C_6H_5^+$ (m/z 77) and $C_4H_3^+$ (m/z 51) observed also in the mass spectrum measured at 60 eV and other fragments ($C_5H_4^+$, $C_3H_2^+$, and $C_2H_3^+$) with a relative contribution enhanced with respect

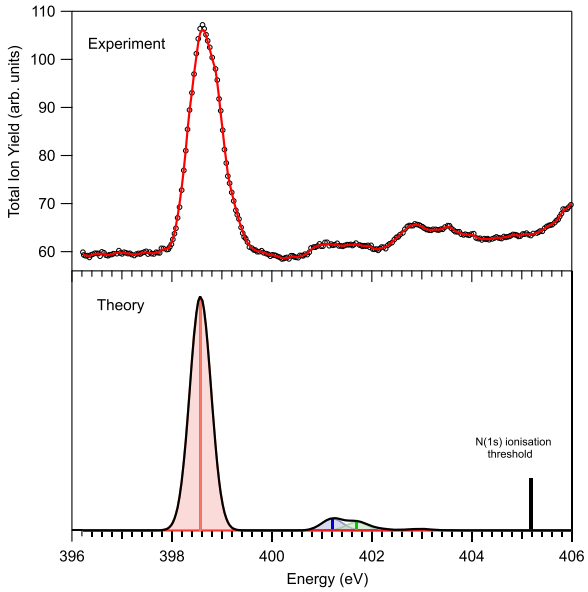


FIG. 8. The nitrogen K-edge NEXAFS spectrum of AB. The experimental data are reported in the top panel as black dots with a superimposed three-point smooth line in red, while the theoretical results convoluted by Gaussian functions of FWHM = 0.7 eV are depicted in the bottom panel. The N(1s) ionization threshold is also reported.

to the spectrum at 60 eV. All fragments are broader and accompanied by other minor contributions corresponding to H losses. These observations, together with the low cross section for direct photo-ionization of the valence shell, suggest that these fragments are produced in the fragmentation of AB dication formed by the radiationless Auger–Meitner decay following core ionization.

In Fig. 10, the branching ratio of the most representative fragments is reported vs photon energy in the region of the N(1s) excitation. A similar representation for all fragments is reported in the supplementary material.

In the N(1s) near threshold region, all main fragments display a clear enhancement at the 1s→ π_1^* transition and 51^+, which is characterized by the largest BR, also in the case of the other excitations. Consistently with the observation in the 1s continuum, 51^+ and 77^+ have the largest BR. This observation can be rationalized in the following way. Inner shell excitation can decay to cation states via resonant Auger. In the case of participator decay, the excited electron is involved in the process, and the final states correspond to one-hole states, as the ones more likely populated in direct ionization at lower photon energies. Vice versa, in the case of spectator decay, the excited electron does not participate in the process, and the final state is a singly charge ion with two holes and an excited electron. These states are the ones that converge to double ionization. Thus, participator decays are expected to produce mass spectra similar to the one obtained in the valence ionization, i.e., with a dominant contribution of 182^+, 152–154^+, 105^+, while spectator decays will produce mass spectra similar to the ones dominated by dications fragmentation. The 77^+ represents a peculiar case because, as also observed in the PEPICO experiments (Fig. 2), it can be produced via the fragmentation of either cations or dications. Looking at the variation of the BR in the resonance and at the width of the peak, one can state that in this energy region, the contribution of direct ionization is negligible. According to the different BRs, one can clearly state that the radiationless decay of the N 1s→ π_1^* is dominated by spectator decays.

TABLE III. The experimental and calculated core excited states of AB. For the numbering of the C atoms, see Fig. 1. For orbitals, see Figs. S1 and S3 of the supplementary material. Only core excited states with $|\mathrm{TDM}|^2$ higher than 1×10^{-4} are reported. A complete list of the core excited states, which give rise to the NEXAFS spectra, can be found in Tables SII and SIX–SXII in the supplementary material. The experimental positions refer to the features distinguishable in the spectra reported in Figs. 7 and 8.

	Theory (eV)	Exp (eV)	Configuration (weight %)
C3(1s)	284.15	285.72	$1s^{[1]}\pi_1^{*[1]}$ (76%)
C1(1s)	284.40		$1s^{[1]}\pi_1^{*[1]}$ (76%)
C4(1s)	284.80		$1s^{[1]}\pi_1^{*[1]}$ (76%)
C2(1s)	285.31	285.12	$1s^{[1]}\pi_2^{*[1]}$ (75%)
C3(1s)	285.41	285.48	$1s^{[1]}\pi_2^{*[1]}$ (62%)
C1(1s)	286.33	286.63	$1s^{[1]}\pi_7^{*[1]}$ (32%)
			$1s^{[1]}\pi_2^{*[1]}$ (30%)
C3(1s)	286.40	287.27	$1s^{[1]}\pi_7^{*[1]}$ (34%)
C4(1s)	286.78		$1s^{[1]}\pi_7^{*[1]}$ (48%)
C2(1s)	286.99		$1s^{[1]}\pi_7^{*[1]}$ (24%)
C3(1s)	289.06	288.11	$1s^{[1]}n_2^{[1]}\pi_1^{*[2]}$ (17%)
			$1s^{[1]}\pi_3^{*[1]}$ (16%)
			$1s^{[1]}\pi_4^{*[1]}\pi_1^{*[1]}\pi_1^{*[2]}$ (11%)
C4(1s)	289.59	289.07	$1s^{[1]}\pi_3^{*[1]}$ (12%)
			$1s^{[1]}\pi_4^{*[1]}\pi_1^{*[1]}\pi_1^{*[2]}$ (10%)
C4(1s)	289.63		ex.m.
N(1s)	398.57	398.67	$1s^{[1]}\pi_1^{*[1]}$ (78%)
N(1s)	401.21	401.13	$1s^{[1]}\pi_4^{[1]}\pi_1^{*[2]}$ (27%)
N(1s)	401.70	401.75	$1s^{[1]}\pi_7^{[1]}\pi_1^{*[1]}\pi_5^{*[1]}$ (19%)
			$1s^{[1]}\pi_6^{[1]}\pi_1^{*[1]}\pi_7^{*[1]}$ (19%)

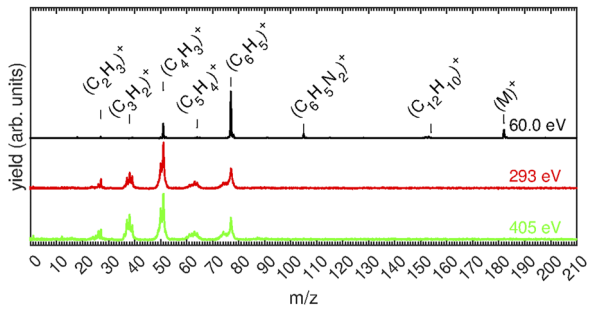


FIG. 9. Mass spectra of AB measured at photon energies of 60 eV (black), 293 eV (red), and 405 eV (green).

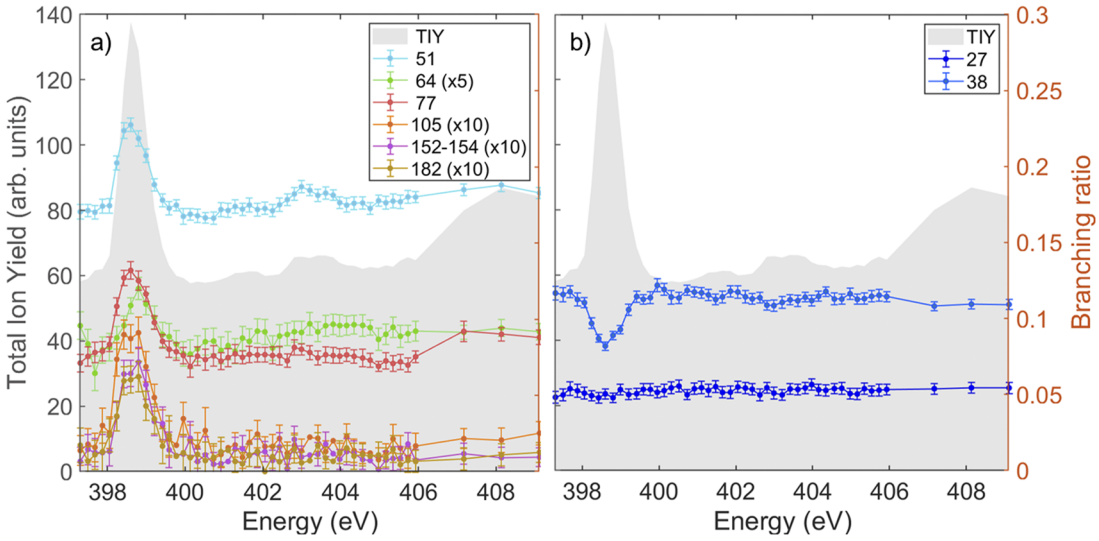


FIG. 10. Total ion yield (TIY, gray shaded area) and branching ratios of the most representative fragments at the N(1s) excitation/ionization region (colored lines). Panel (a) reports examples of fragments with a branching ratio increasing at the π^* resonance. Panel (b) displays examples of fragments either decreasing or unaffected by the core resonant excitation. The entire set of results is reported in Fig. IV of the supplementary material. The error bars represent the confidence of the fit.

As for the total ion yield at the C(1s) edge (see Fig. S4 in supplementary material), it is clearly observed that on resonance, the H-loss process is reduced in all channels, stays almost constant outside the resonance, and decreases in the continuum. Qualitatively, one might ascribe this observation to the rapid dissociation of the molecular ion, formed in highly excited states in the spectator decay, that prevents the energy redistribution, leading to further H losses. In the C(1s) near threshold region, the presence of excitations of the four non-equivalent C atoms partially masks the effects clearly detected in the case of the N(1s) strong isolated resonance. However, a similar trend in the region of the resonances, with increasing BR dominated by fragments 77^+ and 51^+, is observed.

D. The excited state core ionization and near-edge absorption

Well established valence-shell spectroscopy techniques probe changes in electron density on a molecular scale by interrogating transitions among generally delocalized frontier orbitals. Transient UV/Vis and PES have been successfully applied to study the ultrafast non-adiabatic dynamics in AB.[6,83] Core–shell spectroscopy, on the other hand, addresses changes in electron density on the atomic scale and acts as a sensor for the involvement of individual atoms in the electronic structure of an excited state. Having demonstrated the accuracy of our simulation protocol in the reproduction of ground state NEXAFS and XPS features, we conclude by reporting simulations of the corresponding spectra out of the $n\pi^*$ and $\pi\pi^*$ excited states (Fig. 11) and discuss the characteristic features, which would facilitate studying photoinduced dynamics in AB by means of transient core–shell spectroscopy techniques. Thereby, we omit to couple between electronic and vibrational degrees of freedom and discuss only the FC spectra (potentially obtainable with ultrashort pulses at very short pump-probe delays).

Electronic transitions create vacancies in the occupied valence orbitals, which could give rise to characteristic red-shifted absorption features in transient NEXAFS if the provided core and valence orbitals exhibit finite overlap. In the case of AB, the excited state NEXAFS shows peaks that have been red-shifted by several eV to a background-free spectral region [Fig. 11(a) between 281 and 282 eV in the $\pi\pi^*$ carbon K edge and Fig. 11(c) at 396 eV in the $n\pi^*$ nitrogen K edge]. They are associated with core-to-valence C(1s) $\rightarrow \pi$ and N(1s) $\rightarrow n$ transitions. The complementary signals C(1s) $\rightarrow n$ and N(1s) $\rightarrow \pi$ are weak due to the vanishing molecular orbital overlap.

Excitations are often associated with partial charge transfer throughout the system. Depletion of electron density (detachment) around an atom hampers, whereas accumulation (attachment) facilitates core ionization/excitation. In the case of AB, $\pi \rightarrow \pi^*$ excitation leads to a depletion in the aromatic rings accompanied by density increase in the central NN fragment (Fig. 11, central panel). Density flow in the opposite direction, albeit to a less extent, is observed in the $n\pi^*$ state. The consequences of this partial charge transfer are best seen in the nitrogen NEXAFS and XPS [Figs. 11(c) and 11(d)] where 1s $\rightarrow \pi^*$ transition (NEXAFS) and 1s ionization (XPS) exhibit up to 1 eV red-shift in the $\pi\pi^*$ state (due to the electron attachment) or blue-shift in the $n\pi^*$ state (due to the electron detachment). The complementary detachment/attachment in the aromatic ring depends on the position of the carbon (ortho, meta, para) with respect to the NN fragment and gives rise to more complex patterns, e.g., the broad absorption band between 289 and 291 eV in the carbon XPS out of the $\pi\pi^*$ state [Fig. 11(b)].

Finally, the nitrogen and carbon XPS [Figs. 11(b) and 11(d)] out of the $\pi\pi^*$ state show a weak signal red-shifted by about 3 eV associated with a shake-down feature, i.e., a core-ionization accompanied by a $\pi \leftarrow \pi^*$ de-excitation. These features disappear following the decay to the $n\pi^*$.

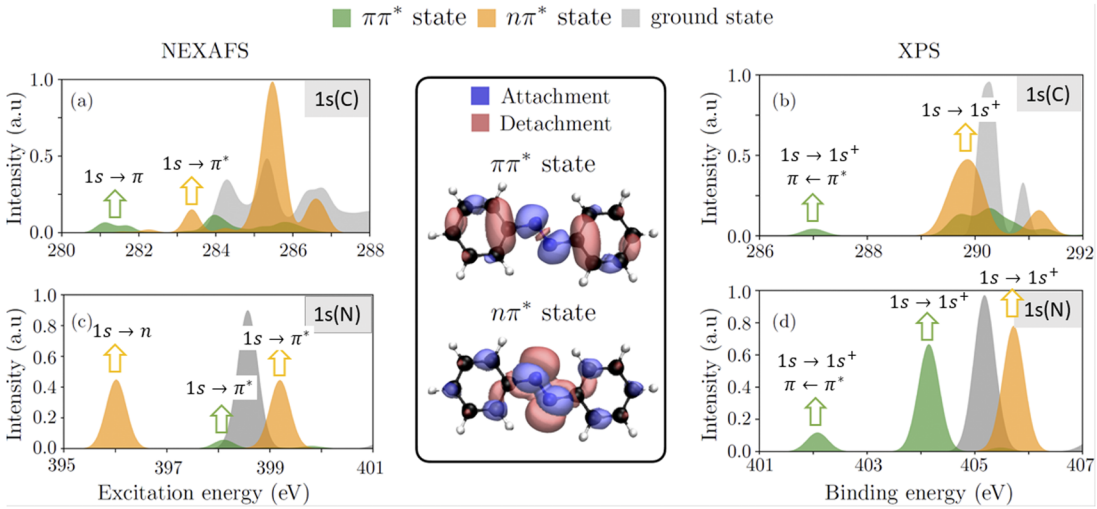


FIG. 11. Carbon 1s [(a) and (b)] and nitrogen 1s [(c) and (d)] NEXAFS [(a) and (c)] and XPS [(b) and (d)] signals from the $\pi\pi^*$ (green), $n\pi^*$ (yellow), and ground (gray) electronic states calculated at the FC point. The central panel shows detachment (red)/attachment (blue) electron densities for the excited states, thus highlighting how the electronic density is redistributed upon excitation.

In summary, both the carbon and the nitrogen K edges can be utilized to follow the $\pi\pi^* \rightarrow n\pi^* \rightarrow GS$ internal conversion by means of transient NEXAFS and/or XPS owing to their pronounced sensitivity to the electronic structure.

IV. CONCLUSIONS

The photo-ionization and photo-fragmentation of the valence and core shells of azobenzene as well as of the NEXAFS in the carbon and nitrogen K edge regions have been studied by a combination of several experimental and theoretical methods, providing a comprehensive characterization of the electronic structure of this prototype molecule for a class of molecular photo-switches.

In the valence shell, the lowest energy band has been attributed to the π_7^+ state delocalized over the aromatic ring, thus explaining the well known high sensitivity of the PES to the functional group in the class of AB molecules as due to competing electron-donating and -withdrawing role of the substituent. Then, guided by the theoretical simulations, the main features in the PES spectrum have also been assigned, showing a dominance of π and lone pair orbitals in the binding energy region below 11 eV and a dominance of σ orbitals above that. The study of the fragmentation paths has allowed shedding light on the subsequent fate of these valence and core shell excited/ionized states of the molecule.

The valence shell photo-fragmentation of AB relies on two primary and competing processes, namely the C_6H_5 loss and N_2 extraction from the parent ion, with the branching ratio of the first process dominating over the second. In both cases, the phenyl rings are preserved, thanks to their high stability, and the first bond to break in the AB molecular structure is(are) the C–N bond(s) in the binding energy region >10 eV. Theoretical calculations propose that fragment 77^+, attributed to $C_6H_5^+$, is more easily produced by a two-step mechanism from $C_6H_5N_2^+$ (105^+) followed by N_2 loss rather than a single C–N bond breakage with a charge on $C_6H_5^+$ rather than on $C_6H_5N_2^+$. The proposed mechanism for the formation of fragment 154^+, attributed to $C_{12}H_{10}^+$, may involve a *trans*–cis isomerization of the molecular ion, which brings phenyl rings in closer proximity, thus favoring the formation of a biphenyl structure with the expulsion of the stable N_2. Even though with the present theoretical approach, it has not been possible to prove the "condensation" of the two phenyl rings following *trans*–cis isomerization, and the relatively low energy barrier calculated for this isomerization in the molecular ion places this transition well below the calculated AE of 154^+ as well as below its first onset in the PEPICO measurements, making isomerization at least energetically possible as a preliminary rearrangement preceding N_2 loss.

The theoretical analysis of 154^+ shows how the existence of this fragment is extremely favorable from a purely energetic/structural point of view, suggesting that ones formed, it should be stable. However, this is clearly in contrast with the PEPICO observation, where this fragment contributes to just a few percent of the branching ratio. Therefore, the most likely explanation for the low intensity of fragment 154^+ has to be found in dynamical aspects where from the *cis*-state of the molecular ion, the breakage of two C–N bonds with the rebound of the two phenyl rings makes this a quite unlikely process to happen.

In the core region, apart from a rigid energy shift, an overall excellent agreement has been observed between theory and experiment. The NEXAFS exhibit features that reflect the energetic structure of the manifold of virtual orbitals. In particular, we attributed each individual peak in the carbon K edge to a core transition to a specific π^* orbital. The NEXAFS, however, cannot resolve the four types of carbon atoms present in AB as the chemical shifts are smaller than the bandwidth. Comparison between the carbon and nitrogen spectra highlights another trait of NEXAFS, namely that virtual orbitals must exhibit contribution over the atoms whose

core orbitals are interrogated to give rise to signals. In that regard, the single-peak structure of the nitrogen K edge opposite to the multi-peak structure of the carbon K edge is a consequence of the localization of most π^* orbitals on the aromatic rings.

The XPS exhibit generally less features as each non-equivalent 1s orbital contributes only with one intense peak. The N XPS resolves a couple of satellites, one confirmed by the experiment. The C XPS exhibits a non-trivial ordering of the binding energies of the four non equivalent C atoms, a consequence of the interplay of inductive and resonant effects in the aromatic ring. The chemical shifts allow distinguishing spectroscopically of the carbon directly bonded to the nitrogen.

Finally, we demonstrated that the carbon and nitrogen K edges can be utilized to study the photoinduced dynamics in AB owing to transient signals—core-transitions to vacant occupied orbitals and shake-down features—appearing in background-free regions of the NEXAFS and XPS. Moreover, core-excitation and ionization are shown to be very sensitive to the unique distribution of electron density in each excited state, which makes them excellent internal observers of the participation of certain atoms in the excited state deactivation.

SUPPLEMENTARY MATERIAL

See the supplementary material for figures of the full active spaces used in the RASSCF/RASPT2 calculations of the PES, carbon and nitrogen XPS and NEXAFS, as well as Tables reporting energies, leading configurations, and transition intensities of the computed PES, carbon and nitrogen XPS, and NEXAFS.

ACKNOWLEDGMENTS

We acknowledge Elettra Sincrotrone Trieste for providing access to its synchrotron radiation facilities and for financial support, and the Gas Phase staff for providing technical and scientific support. Support from the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Chemical Sciences, Geosciences and Biosciences Division under Award No. DE-SC0019484 (F.M., M.G., S.M., and A.N.) is gratefully acknowledged. RB-V acknowledges support from the Fondazione Cariplo through the project DINAMO (Grant No. 2020-4380). This work was performed in the framework of the PRIN 20173B72NB research project "Predicting and controlling the fate of biomolecules driven by extreme-ultraviolet radiation."

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

PB, AN, RB-V, LA designed the project, interpreted the results and wrote the manuscript. PB, LC, DM, YW, RR performed the experimental measurements. LC and YW and analysed the data and LC, PB, LA, RB-V, FV, ML and MN contributed to the discussion of the experimental results. FM, AN and ARC performed the numerical simulations. All authors have carefully read the manuscript. LC, FM and YW contributed equally.

L. Carlini: Data curation (equal); Formal analysis (equal); Investigation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **F. Montorsi**: Data curation (equal); Formal analysis (equal); Investigation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Y. Wu**: Data curation (equal); Formal analysis (equal); Investigation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **P. Bolognesi**: Conceptualization (equal); Methodology (equal); Project administration (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal). **R. Borrego-Varillas**: Conceptualization (equal); Funding acquisition (equal); Methodology (equal); Project administration (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal). **A. R. Casavola**: Investigation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **M. C. Castrovilli**: Writing – review & editing (equal). **J. Chiarinelli**: Writing – review & editing (equal). **D. Mocci**: Data curation (equal); Investigation (equal); Writing – review & editing (equal). **F. Vismarra**: Investigation (equal); Writing – review & editing (equal). **M. Lucchini**: Investigation (equal); Resources (equal); Supervision (equal); Writing – review & editing (equal). **M. Nisoli**: Funding acquisition (equal); Investigation (equal); Resources (equal); Supervision (equal); Writing – review & editing (equal). **S. Mukamel**: Investigation (equal); Writing – review & editing (equal). **M. Garavelli**: Funding acquisition (equal); Resources (equal); Writing – review & editing (equal). **R. Richter**: Data curation (equal); Investigation (equal); Resources (equal); Writing – review & editing (equal). **A. Nenov**: Conceptualization (equal); Methodology (equal); Project administration (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal). **L. Avaldi**: Conceptualization (equal); Funding acquisition (equal); Resources (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

REFERENCES

[1] J. D. Harris, M. J. Moran, and I. Aprahamian, "New molecular switch architectures," Proc. Natl. Acad. Sci. U. S. A. **115**, 9414–9422 (2018).
[2] A. Stolow, "Femtosecond time-resolved photoelectron spectroscopy of polyatomic molecules," Annu. Rev. Phys. Chem. **54**, 89–119 (2003).
[3] H. M. D. Bandara and S. C. Burdette, "Photoisomerization in different classes of azobenzene," Chem. Soc. Rev. **41**, 1809–1825 (2012).
[4] C. Renner and L. Moroder, "Azobenzene as conformational switch in model peptides," ChemBioChem **7**, 868–878 (2006).
[5] T. Ikeda and O. Tsutsumi, "Optical switching and image storage by means of azobenzene liquid-crystal films," Science **268**, 1873–1875 (1995).
[6] A. Nenov, R. Borrego-Varillas, A. Oriana, L. Ganzer, F. Segatta, I. Conti, J. Segarra-Marti, J. Omachi, M. Dapor, S. Taioli *et al.*, "Uv-light-induced vibrational coherences: The key to understand kasha rule violation in trans-azobenzene," J. Phys. Chem. Lett. **9**, 1534–1541 (2018).

[7]D. Keefer, F. Aleotti, J. R. Rouxel, F. Segatta, B. Gu, A. Nenov, M. Garavelli, and S. Mukamel, "Imaging conical intersection dynamics during azobenzene photoisomerization by ultrafast x-ray diffraction," Proc. Natl. Acad. Sci. U. S. A. **118**, e2022037118 (2021).
[8]J. R. Rouxel, D. Keefer, F. Aleotti, A. Nenov, M. Garavelli, and S. Mukamel, "Coupled electronic and nuclear motions during azobenzene photoisomerization monitored by ultrafast electron diffraction," J. Chem. Theory Comput. **18**, 605–613 (2022).
[9]F. A. Delesma, M. Van den Bossche, H. Grönbeck, P. Calaminici, A. M. Köster, and L. G. M. Pettersson, "A chemical view on x-ray photoelectron spectroscopy: The ESCA molecule and surface-to-bulk XPS shifts," ChemPhysChem **19**, 169–174 (2018).
[10]S. L. Sorensen, X. Zheng, S. H. Southworth, M. Patanen, E. Kokkonen, B. Oostenrijk, O. Travnikova, T. Marchenko, M. Simon, C. Bostedt *et al.*, "From synchrotrons for XFELs: The soft x-ray near-edge spectrum of the ESCA molecule," J. Phys. B: At., Mol. Opt. Phys. **53**, 244011 (2020).
[11]C. Kolczewski, R. Püttner, M. Martins, A. S. Schlachter, G. Snell, M. M. Sant'Anna, K. Hermann, and G. Kaindl, "Spectroscopic analysis of small organic molecules: A comprehensive near-edge x-ray-absorption fine-structure study of C_6-ring-containing molecules," J. Chem. Phys. **124**, 034302 (2006).
[12]D. Solomon, J. Lehmann, J. Kinyangi, B. Liang, K. Heymann, L. Dathe, K. Hanley, S. Wirick, and C. Jacobsen, "Carbon (1s) NEXAFS spectroscopy of biogeochemically relevant reference organic compounds," Soil Sci. Soc. Am. J. **73**, 1817–1830 (2009).
[13]A. Ponzi, E. Bernes, D. Toffoli, G. Fronzoni, C. Callegari, A. Ciavardini, M. Di Fraia, R. Richter, K. C. Prince, H. Sa'adeh *et al.*, "Carbon and nitrogen K-edge NEXAFS spectra of indole, 2, 3-dihydro-7-azaindole, and 3-formylindole," J. Phys. Chem. A **125**, 4160–4172 (2021).
[14]T. J. Wasowicz, I. Ljubić, A. Kivimäki, and R. Richter, "Core–shell excitation of isoxazole at the C, N, and O K-edges–an experimental NEXAFS and theoretical TD-DFT study," Phys. Chem. Chem. Phys. **24**, 19302–19313 (2022).
[15]G. G. B. de Souza and J. C. Gonzalez, "Near edge x-ray absorption fine structure (NEXAFS) spectroscopy of molecules of biological interest: From chemically pure to complex samples," in *Radiation in Bioanalysis* (Springer, 2019), pp. 287–312.
[16]V. Feyer, O. Plekan, R. Richter, M. Coreno, K. C. Prince, and V. Carravetta, "Photoemission and photoabsorption spectroscopy of glycyl-glycine in the gas phase," J. Phys. Chem. A **113**, 10726–10733 (2009).
[17]C. Li, P. Salén, V. Yatsyna, L. Schio, R. Feifel, R. Squibb, M. Kamińska, M. Larsson, R. Richter, M. Alagia *et al.*, "Experimental and theoretical XPS and NEXAFS studies of n-methylacetamide and n-methyltrifluoroacetamide," Phys. Chem. Chem. Phys. **18**, 2210–2218 (2016).
[18]S. Coriani and H. Koch, "Communication: X-Ray absorption spectra and core-ionization potentials within a core-valence separated coupled cluster framework," J. Chem. Phys. **143**, 181103 (2015).
[19]P. Norman and A. Dreuw, "Simulating x-ray spectroscopies and calculating core-excited states of molecules," Chem. Rev. **118**, 7208–7248 (2018).
[20]C. Ehlert, M. Gühr, and P. Saalfrank, "An efficient first principles method for molecular pump-probe NEXAFS spectra: Application to thymine and azobenzene," J. Chem. Phys. **149**, 144112 (2018).
[21]M. L. Vidal, X. Feng, E. Epifanovsky, A. I. Krylov, and S. Coriani, "New and efficient equation-of-motion coupled-cluster framework for core-excited and core-ionized states," J. Chem. Theory Comput. **15**, 3117–3133 (2019).
[22]M. L. Vidal, A. I. Krylov, and S. Coriani, "Dyson orbitals within the fc-CVS-EOM-CCSD framework: Theory and application to x-ray photoelectron spectroscopy of ground and excited states," Phys. Chem. Chem. Phys. **22**, 2693–2703 (2020).
[23]M. L. Vidal, A. I. Krylov, and S. Coriani, "Correction: Dyson orbitals within the fc-CVS-EOM-CCSD framework: Theory and application to x-ray photoelectron spectroscopy of ground and excited states," Phys. Chem. Chem. Phys. **22**, 3744–3747 (2020).
[24]N. A. Besley, "Density functional theory based methods for the calculation of x-ray spectroscopy," Acc. Chem. Res. **53**, 1306–1315 (2020).
[25]N. A. Besley, "Modeling of the spectroscopy of core electrons with density functional theory," Wiley Interdiscip. Rev.: Comput. Mol. Sci. **11**, e1527 (2021).
[26]C. D. Rankine and T. J. Penfold, "Progress in the theory of x-ray spectroscopy: From quantum chemistry to machine learning and ultrafast dynamics," J. Phys. Chem. A **125**, 4276–4293 (2021).
[27]B. Helmich-Paris, "Simulating x-ray absorption spectra with complete active space self-consistent field linear response methods," Int. J. Quantum Chem. **121**, e26559 (2021).
[28]A. C. Paul, R. H. Myhre, and H. Koch, "New and efficient implementation of CC3," J. Chem. Theory Comput. **17**, 117–126 (2021).
[29]Y. Yao, D. Golze, P. Rinke, V. Blum, and Y. Kanai, "All-electron BSE@GW method for k-edge core electron excitation energies," J. Chem. Theory Comput. **18**, 1569–1583 (2022).
[30]W. Park, M. Alías-Rodríguez, D. Cho, S. Lee, M. Huix-Rotllant, and C. H. Choi, "Mixed-reference spin-flip time-dependent density functional theory for accurate x-ray absorption spectroscopy," J. Chem. Theory Comput. **18**, 6240–6250 (2022).
[31]F. Montorsi, F. Segatta, A. Nenov, S. Mukamel, and M. Garavelli, "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. **18**, 1003–1016 (2022).
[32]O. González-Magaña, G. Reitsma, M. Tiemens, L. Boschman, R. Hoekstra, and T. Schlathölter, "Near-edge x-ray absorption mass spectrometry of a gas-phase peptide," J. Phys. Chem. A **116**, 10745–10751 (2012).
[33]S. Dörner, L. Schwob, K. Atak, K. Schubert, R. Boll, T. Schlathölter, M. Timm, C. Bülow, V. Zamudio-Bayer, B. von Issendorff *et al.*, "Probing structural information of gas-phase peptides by near-edge x-ray absorption mass spectrometry," J. Am. Soc. Mass Spectrom. **32**, 670–684 (2021).
[34]A. R. Milosavljević, K. Jänkälä, M. L. Ranković, F. Canon, J. Bozek, C. Nicolas, and A. Giuliani, "Oxygen K-shell spectroscopy of isolated progressively solvated peptide," Phys. Chem. Chem. Phys. **22**, 12909–12917 (2020).
[35]A. R. Milosavljević, C. Nicolas, M. L. Ranković, F. Canon, C. Miron, and A. Giuliani, "K-shell excitation and ionization of a gas-phase protein: Interplay between electronic structure and protein folding," J. Phys. Chem. Lett. **6**, 3132–3138 (2015).
[36]G. Polzonetti and C. Battocchio, "Functional and nanostructured materials investigated by XPS and NEXAFS spectroscopies," in *Advances in Macromolecules* (Springer, 2010), pp. 165–217.
[37]A. Tyagi, V. Mišeikis, L. Martini, S. Forti, N. Mishra, Z. M. Gebeyehu, M. A. Giambra, J. Zribi, M. Frégnaux, D. Aureau *et al.*, "Ultra-clean high-mobility graphene on technologically relevant substrates," Nanoscale **14**, 2167–2176 (2022).
[38]T. J. A. Wolf and M. Gühr, "Photochemical pathways in nucleobases measured with an x-ray fel," Philos. Trans. R. Soc., A **377**, 20170473 (2019).
[39]T. J. A. Wolf, R. H. Myhre, J. P. Cryan, S. Coriani, R. J. Squibb, A. Battistoni, N. Berrah, C. Bostedt, P. Bucksbaum, G. Coslovich *et al.*, "Probing ultrafast $\pi\pi^*/n\pi^*$ internal conversion in organic chromophores via K-edge resonant absorption," Nat. Commun. **8**, 29 (2017).
[40]F. Lever, D. Mayer, J. Metje, S. Alisauskas, F. Calegari, S. Düsterer, R. Feifel, M. Niebuhr, B. Manschwetus, M. Kuhlmann *et al.*, "Core-level spectroscopy of 2-thiouracil at the sulfur L_1- and $L_{2,3}$-edges utilizing a SASE free-electron laser," Molecules **26**, 6469 (2021).
[41]F. Segatta, A. Nenov, S. Orlandi, A. Arcioni, S. Mukamel, and M. Garavelli, "Exploring the capabilities of optical pump x-ray probe NEXAFS spectroscopy to track photo-induced dynamics mediated by conical intersections," Faraday Discuss. **221**, 245–264 (2019).
[42]I. Fischer and S. T. Pratt, "Photoelectron spectroscopy in molecular physical chemistry," Phys. Chem. Chem. Phys. **24**, 1944–1959 (2022).
[43]M. S. Schuurman and V. Blanchet, "Time-resolved photoelectron spectroscopy: The continuing evolution of a mature technique," Phys. Chem. Chem. Phys. **24**, 20012–20024 (2022).
[44]P. Bolognesi and L. Avaldi, "Photoelectron–photoion(s) coincidence studies of molecules of biological interest," Phys. Chem. Chem. Phys. **24**, 22356–22370 (2022).
[45]R. R. Blyth, R. Delaunay, M. Zitnik, J. Krempasky, R. Krempaska, J. Slezak, K. C. Prince, R. Richter, M. Vondracek, R. Camilloni *et al.*, "The high resolution gas phase photoemission beamline, Elettra," J. Electron Spectrosc. Relat. Phenom. **101-103**, 959 (1999).
[46]J. Chiarinelli, D. Barreiro-Lage, P. Bolognesi, R. Richter, H. Zettergren, M. H. Stockett, S. Díaz-Tendero, and L. Avaldi, "Electron and ion spectroscopy of

the cyclo-alanine–alanine dipeptide," Phys. Chem. Chem. Phys. **24**, 5855–5867 (2022).
[47]P. Bolognesi, V. Carravetta, L. Sementa, G. Barcaro, S. Monti, P. Manjari Mishra, A. Cartoni, M. C. Castrovilli, J. Chiarinelli, S. Tosic *et al.*, "Core shell investigation of 2-nitroimidazole," Front. Chem. **7**, 151 (2019).
[48]G. Cautero, R. Sergo, L. Stebel, P. Lacovig, P. Pittana, M. Predonzani, and S. Carrato, "A two-dimensional detector for pump-and-probe and time resolved experiments," Nucl. Instrum. Methods Phys. Res., Sect. A **595**, 447–459 (2008).
[49]R. H. Menk, M. Antonelli, G. Brajnik, J. Bufon, C. Dri, D. Giuressi, A. Gubertini, C. Nichetti, G. Pinaroli, P. Pittana *et al.*, "Detectors for present and future light sources at Elettra," AIP Conf. Proc. **2054**, 060071 (2019).
[50]P. Natalis and J. L. Franklin, "Ionization and dissociation of diphenyl and condensed-ring aromatics by electron impact. iii. azobenzene," Int. J. Mass Spectrom. **40**, 35–42 (1981).
[51]G. Wright and C. Brion, "K-shell energy loss spectra of 2.5 keV electrons in CO_2 and N_2O," J. Electron Spectrosc. Relat. Phenom. **3**, 191–205 (1974).
[52]M. Tronc, G. C. King, and F. H. Read, "Carbon k-shell excitation in small molecules by high-resolution electron impact," J. Phys. B: At., Mol. Opt. Phys. **12**, 137 (1979).
[53]M. Tronc, G. C. King, and F. H. Read, "Nitrogen k-shell excitation in N_2, NO and N_2O by high-resolution electron energy-loss spectroscopy," J. Phys. B: At., Mol. Opt. Phys. **13**, 999 (1980).
[54]T. Hatamoto, M. Matsumoto, X.-J. Liu, K. Ueda, M. Hoshino, K. Nakagawa, T. Tanaka, H. Tanaka, M. Ehara, R. Tamaki *et al.*, "Vibrationally resolved C and O 1s photoelectron spectra of carbon dioxide," J. Electron Spectrosc. Relat. Phenom. **155**, 54–57 (2007).
[55]T. D. Thomas and R. W. Shaw, Jr, "Accurate core ionization potentials and photoelectron kinetic energies for light elements," J. Electron Spectrosc. Relat. Phenom. **5**, 1081 (1974).
[56]R. N. S. Sodhi and C. E. Brion, "Reference energies for inner shell electron energy-loss spectroscopy," J. Electron Spectrosc. Relat. Phenom. **34**, 363–372 (1984).
[57]K. Kimura, *Handbook of HeI Photoelectron of Fundamental Organic Molecules* (Halsted Press, 1981).
[58]F. Aquilante, J. Autschbach, A. Baiardi, S. Battaglia, V. A. Borin, L. F. Chibotaru, I. Conti, L. De Vico, M. Delcey, I. Fdez Galván, N. Ferré, L. Freitag, M. Garavelli, X. Gong, S. Knecht, E. D. Larsson, R. Lindh, M. Lundberg, P. Å. Malmqvist, A. Nenov, J. Norell, M. Odelius, M. Olivucci, T. B. Pedersen, L. Pedraza-González, Q. M. Phung, K. Pierloot, M. Reiher, I. Schapiro, J. Segarra-Martí, F. Segatta, L. Seijo, S. Sen, D.-C. Sergentu, C. J. Stein, L. Ungur, M. Vacher, A. Valentini, and V. Veryazov, "Modern quantum chemistry with [open]molcas," J. Chem. Phys. **152**, 214117 (2020).
[59]B. N. C. Tenorio, A. Ponzi, S. Coriani, and P. Decleva, "Photoionization observables from multi-reference dyson orbitals coupled to B-spline DFT and TD-DFT continuum," Molecules **27**, 1203 (2022).
[60]P. A. Malmqvist, A. Rendell, and B. O. Roos, "The restricted active space self-consistent-field method, implemented with a split graph unitary group approach," J. Phys. Chem. **94**, 5477–5482 (1990).
[61]P. Å. Malmqvist, K. Pierloot, A. R. M. Shahi, C. J. Cramer, and L. Gagliardi, "The restricted active space followed by second-order perturbation theory method: Theory and application to the study of CuO_2 and Cu_2O_2 systems," J. Chem. Phys. **128**, 204109 (2008).
[62]R. Klooster, R. Broer, and M. Filatov, "Calculation of x-ray photoelectron spectra with the use of the normalized elimination of the small component method," Chem. Phys. **395**, 122–127 (2012).
[63]G. Grell, S. I. Bokarev, B. Winter, R. Seidel, E. F. Aziz, S. G. Aziz, and O. Kühn, "Erratum: "multi-reference approach to the calculation of photoelectron spectra including spin-orbit coupling" [J. Chem. Phys. **143**, 074104 (2015)]," J. Chem. Phys. **145**, 089901 (2016).
[64]G. Grell, S. I. Bokarev, B. Winter, R. Seidel, E. F. Aziz, S. G. Aziz, and O. Kühn, "Multi-reference approach to the calculation of photoelectron spectra including spin-orbit coupling," J. Chem. Phys. **143**, 074104 (2015).
[65]M. G. Delcey, L. K. Sørensen, M. Vacher, R. C. Couto, and M. Lundberg, "Efficient calculations of a large number of highly excited states for multiconfigurational wavefunctions," J. Comput. Chem. **40**, 1789–1799 (2019).
[66]F. Aquilante, T. Bondo Pedersen, A. Sánchez de Merás, and H. Koch, "Fast noniterative orbital localization for large molecules," J. Chem. Phys. **125**, 174101 (2006).
[67]J. Almlöf and P. R. Taylor, "General contraction of Gaussian basis sets. I. Atomic natural orbitals for first- and second-row atoms," J. Chem. Phys. **86**, 4070–4077 (1987).
[68]B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov, and P.-O. Widmark, "New relativistic ANO basis sets for transition metal atoms," J. Phys. Chem. A **109**, 6575–6579 (2005).
[69]T. B. Pedersen, F. Aquilante, and R. Lindh, "Density fitting with auxiliary basis sets from Cholesky decompositions," Theor. Chem. Acc. **124**, 1–10 (2009).
[70]M. Frisch, G. Trucks, H. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson *et al.*, GAUSSIAN 09, Revision D.01, Gaussian, Wallingford, CT, 2013.
[71]T. Kobayashi, K. Yokota, and S. Nagakura, "Photoelectron spectra of trans-azobenzene and benzalaniline," J. Electron Spectrosc. Relat. Phenom. **6**, 167–170 (1975).
[72]J. r. Kroner, W. Schneid, N. Wiberg, B. Wrackmeyer, and G. Ziegleder, "13 C, 14 N and 29 Si nuclear magnetic resonance and photoelectron spectra of azo compounds," J. Chem. Soc., Faraday Trans. **74**(74), 1909–1919 (1978).
[73]S. Millefiori, A. Millefiori, S. Pignataro, G. Distefano, and F. P. Colonna, "Gas phase ups investigation of trans-azobenzenes," Z. Naturforsch., A **34**, 1496–1499 (1979).
[74]D. Srzic, M. Žinić, Z. Meić, G. Czira, and J. Tainás, "Mechanistic study of the decomposition reactions of azobenzene," Org. Mass Spectrom. **27**, 1305–1310 (1992).
[75]R. D. Molloy, A. Danielsson, L. Karlsson, and J. H. D. Eland, "Double photoionisation spectra of small molecules and a new empirical rule for double ionisation energies," Chem. Phys. **335**, 49–54 (2007).
[76]V. Myrseth, K. J. Børve, K. Wiesner, M. Bässler, S. Svensson, and L. J. Sæthre, "Vibrational structure and vibronic coupling in the carbon 1s photoelectron spectra of benzene and deuterobenzene," Phys. Chem. Chem. Phys. **4**, 5937–5943 (2002).
[77]P. Bolognesi, G. Mattioli, P. O'Keeffe, V. Feyer, O. Plekan, Y. Ovcharenko, K. C. Prince, M. Coreno, A. Amore Bonapasta, and L. Avaldi, "Investigation of halogenated pyrimidines by x-ray photoemission spectroscopy and theoretical DFT methods," J. Phys. Chem. A **113**, 13593–13600 (2009).
[78]E. Ludwig, T. Strunskus, S. Hellmann, A. Nefedov, C. Wöll, L. Kipp, and K. Rossnagel, "Electronic structure, adsorption geometry, and photoswitchability of azobenzene layers adsorbed on layered crystals," Phys. Chem. Chem. Phys. **15**, 20272–20280 (2013).
[79]C. Bronner, M. Schulze, S. Hagen, and P. Tegeder, "The influence of the electronic structure of adsorbate–substrate complexes on photoisomerization ability," New J. Phys. **14**, 043023 (2012).
[80]M. Quick, A. L. Dobryakov, M. Gerecke, C. Richter, F. Berndt, I. N. Ioffe, A. A. Granovsky, R. Mahrwald, N. P. Ernsting, and S. A. Kovalenko, "Photoisomerization dynamics and pathways of trans- and cis-azobenzene in solution from broadband femtosecond spectroscopies and calculations," J. Phys. Chem. B **118**, 8756–8771 (2014).
[81]The C2 → π_1^* is dark since this Orbital has no coefficient on C2.
[82]A. Nilsson, N. Wassdahl, M. Weinelt, O. Karis, T. Wiell, P. Bennich, J. Hasselström, A. Föhlisch, J. Stöhr, and M. Samant, "Local probing of the surface chemical bond using x-ray emission spectroscopy," Appl. Phys. A **65**, 147–154 (1997).
[83]H. Satzger, C. Root, and M. Braun, "Excited-state dynamics of trans-and cis-azobenzene after UV excitation in the π π^* band," J. Phys. Chem. A **108**, 6265–6271 (2004).

Supplemental Material: Electron and ion spectroscopy of Azobenzene in the valence and core shells

L. Carlini,[1] F. Montorsi,[2] Y. Wu,[3,4] P. Bolognesi*,[1] R. Borrego-Varillas,[4] A. R. Casavola,[1] M. C. Castrovilli,[1] J. Chiarinelli,[1] D. Mocci,[3,4] F. Vismarra,[3,4] M. Lucchini,[3,4] M. Nisoli,[3,4] S. Mukamel,[5] M. Garavelli,[2] R. Richter,[6] A. Nenov*,[2] and L. Avaldi[1]

[1)]*CNR-Istituto di Struttura della Materia, CNR-ISM, Area della Ricerca di Roma 1, Monterotondo, Italy*

[2)]*Dipartimento di Chimica Industriale, Università degli Studi di Bologna, Bologna, Italy*

[3)]*Dipartimento di Fisica, Politecnico di Milano, Piazza Leonardo da Vinci 32, Milano, Italy*

[4)]*CNR-Istituto di Fotonica e Nanotecnologie, CNR-IFN, Piazza Leonardo da Vinci 32, Milano, Italy*

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

[6)]*Elettra Sincrotrone Trieste, Area Science Park, Basovizza, Italy*

(*Electronic mail: artur.nenov@unibo.it)

(*Electronic mail: paola.bolognesi@cnr.it)

(Dated: 13 January 2023)

I. NITROGEN XPS AND NEXAFS

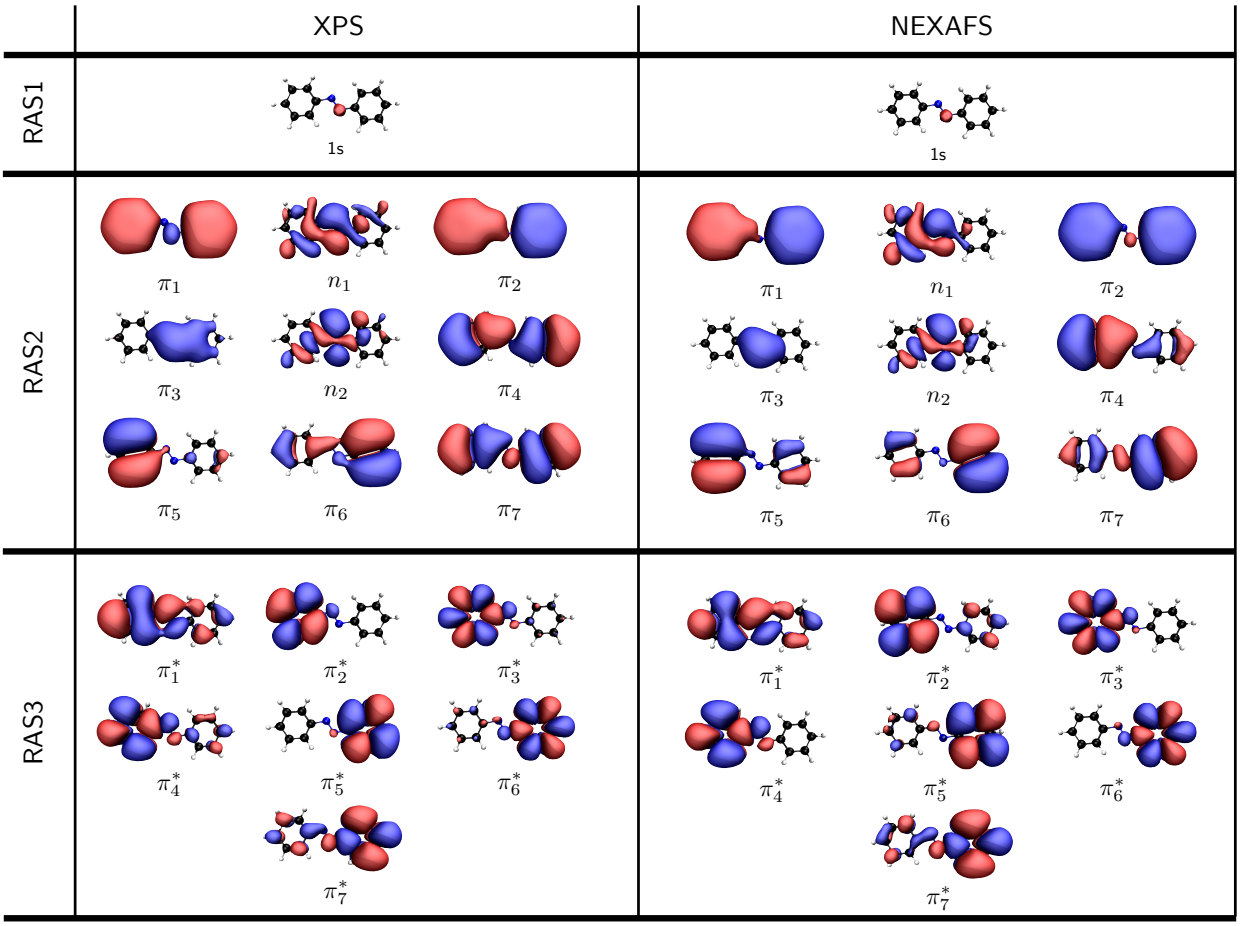


FIG. 1: Active spaces used for the computation of the N K-edge XPS (left) and NEXAFS (right).

Root	BE (eV)	Dyson (a.u.)	CSF (weight)
1	405.180	0.487E+00	$1s^{[1]}(\uparrow)$ (68%)
2	406.953	0.507E-02	$1s^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$ (56%)
3	407.246	0.397E-01	$1s^{[1]}(\uparrow)\pi_7^{[1]}(\uparrow)\pi_1^{*[1]}(\downarrow)$ (46%)
9	407.985	0.514E-02	$1s^{[1]}(\uparrow)n_2^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$ (57%)
11	408.582	0.593E-01	$1s^{[1]}(\uparrow)n_2^{[1]}(\uparrow)\pi_1^{*[1]}(\downarrow)$ (35%)

TABLE I: Main core ionized states, with their leading configurations, that give rise to the XPS spectra reported in Fig. 6.a of the main text. Up and Down arrows are used to highlight the electronic spin. The orbitals involved in the ionization process are highlighted in the left side of fig. 1.

Root	ΔE (eV)	$\lvert TDM\rvert^2$	CSF	Weight
1	398.57	4.51e-03	$1s^{[1]}\pi_1^{*[1]}$	0.78
2	401.21	2.25e-04	$1s^{[1]}\pi_4^{*[1]}$	0.14
			$1s^{[1]}\pi_4{}^{[1]}\pi_1^{*[2]}$	0.27
			$1s^{[1]}\pi_4{}^{[1]}\pi_1^{*[1]}\pi_4^{*[1]}$	0.11
			$1s^{[1]}\pi_4{}^{[1]}\pi_1^{*[1]}\pi_4^{*[1]}$	0.05
7	401.70	1.60e-04	$1s^{[1]}\pi_7{}^{[1]}\pi_1^{*[1]}\pi_5^{*[1]}$	0.18
			$1s^{[1]}\pi_7{}^{[1]}\pi_1^{*[1]}\pi_5^{*[1]}$	0.19
			$1s^{[1]}\pi_6{}^{[1]}\pi_1^{*[2]}$	0.07
			$1s^{[1]}\pi_6{}^{[1]}\pi_1^{*[1]}\pi_7^{*[1]}$	0.19

TABLE II: Main core excited states, with their leading configurations, that give rise to the NEXAFS spectra. Only roots with $|\mathrm{TDM}|^2$ grater than 10^{-4} a.u. are reported in the table. The orbitals involved in the excitation process are highlighted in the right side of figure 1.

II. PES

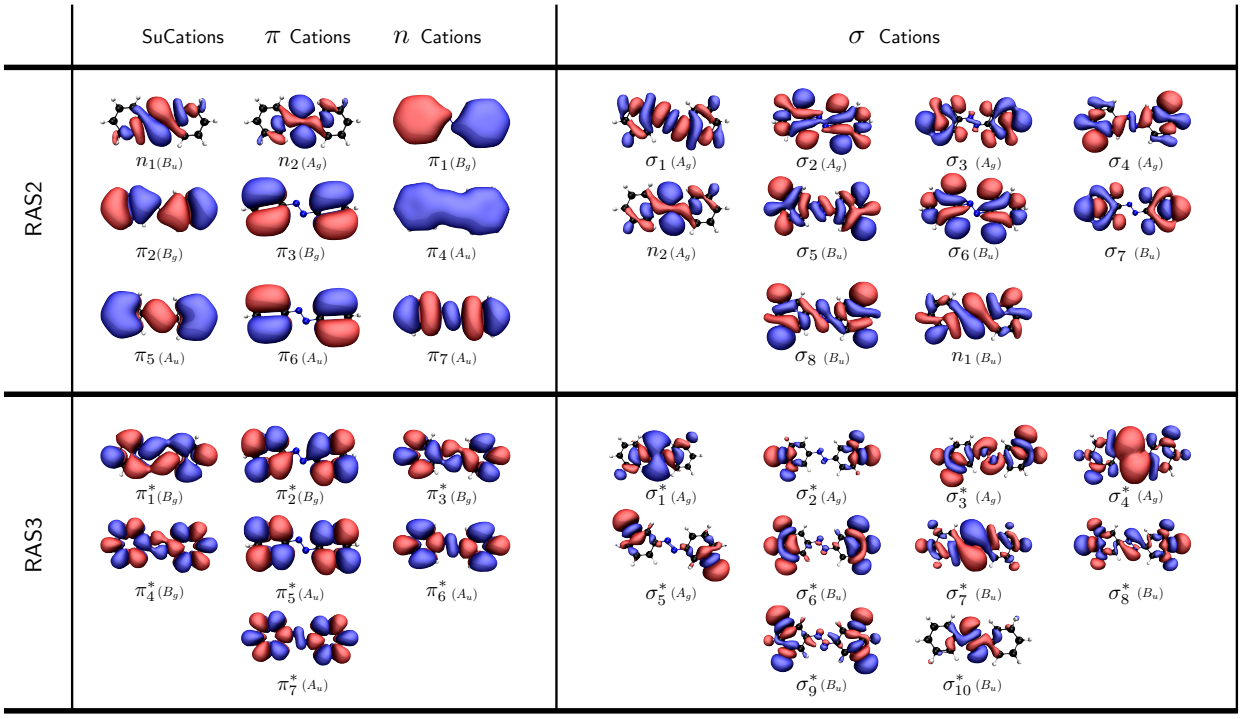


FIG. 2: Active spaces used for the computation of the π&n ionized states (left) and σ ionized states (right).

Root	ΔE (eV)	Dyson (a.u)	Configuration	Weight
1	8.23	7.83e-01	$n_2^{[1]}(\uparrow)$	0.72
2	10.79	5.27e-04	$n_2^{[1]}(\uparrow)\pi_3^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.58
			$n_2^{[1]}(\uparrow)\pi_3^{[1]}(\uparrow)\pi_1^{*[1]}(\downarrow)$	0.05
4	11.53	2.18e-02	$n_2^{[1]}(\uparrow)\pi_2^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.56
5	11.84	2.98e-02	$n_2^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_6^{*[1]}(\uparrow)$	0.07
			$n_2^{[1]}(\uparrow)\pi_2^{[1]}(\uparrow)\pi_1^{*[1]}(\downarrow)$	0.44
			$n_2^{[1]}(\uparrow)\pi_7^{[0]}\pi_1^{*[2]}$	0.10
6	12.46	1.13e-04	$n_2^{[1]}(\uparrow)\pi_3^{[1]}(\downarrow)\pi_2^{*[1]}(\uparrow)$	0.22
			$n_2^{[1]}(\uparrow)\pi_6^{[1]}(\downarrow)\pi_5^{*[1]}(\uparrow)$	0.24
			$n_2^{[1]}(\uparrow)\pi_3^{[1]}(\downarrow)\pi_2^{*[1]}(\uparrow)$	0.07
			$n_2^{[1]}(\uparrow)\pi_6^{[1]}(\uparrow)\pi_5^{*[1]}(\downarrow)$	0.08
7	13.02	1.44e-02	$n_2^{[1]}(\uparrow)\pi_7^{[1]}(\uparrow)\pi_5^{*[1]}(\downarrow)$	0.05
			$n_2^{[1]}(\uparrow)\pi_7^{[0]}\pi_1^{*[2]}$	0.23

TABLE III: Main valence ionized states of A_g symmetry, with their leading configurations, that give rise to PES spectra reported in Fig. 2.a of the main text. The orbitals involved in the ionization process are highlighted in the left side of figure 2. Up and Down arrows are used to highlight the electronic spin.

Root	ΔE (eV)	Dyson (a.u.)	Configuration	Weight
1	8.91	8.34e-01	$\pi_3^{[1]}(\uparrow)$	0.71
2	9.29	8.09e-01	$\pi_2^{[1]}(\uparrow)$	0.69
3	10.74	5.91e-02	$\pi_7^{[0]}\pi_1^{*[1]}(\uparrow)$	0.58
			$\pi_2^{[0]}\pi_1^{*[1]}(\uparrow)$	0.06
4	11.30	8.18e-03	$\pi_6^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.36
			$\pi_2^{[1]}(\uparrow)\pi_3^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.12
			$\pi_6^{[1]}(\uparrow)\pi_7^{[1]}(\uparrow)\pi_1^{*[1]}(\downarrow)$	0.11
5	11.54	5.76e-01	$\pi_1^{[1]}(\uparrow)$	0.49
			$\pi_6^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_2^{*[1]}(\uparrow)$	0.05
6	11.93	1.15e-04	$\pi_6^{[0]}\pi_1^{*[1]}(\uparrow)$	0.30
			$\pi_3^{[0]}\pi_1^{*[1]}(\uparrow)$	0.29
7	12.02	3.58e-02	$\pi_7^{[0]}\pi_2^{*[1]}(\uparrow)$	0.13
			$\pi_6^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.05
			$\pi_2^{[1]}(\uparrow)\pi_3^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.06
			$\pi_6^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_1^{*[1]}(\downarrow)$	0.22
			$\pi_2^{[1]}(\uparrow)\pi_3^{[1]}(\uparrow)\pi_1^{*[1]}(\downarrow)$	0.08
			$\pi_3^{[1]}(\uparrow)\pi_7^{[1]}(\uparrow)\pi_6^{*[1]}(\downarrow)$	0.06
8	12.66	1.75e-02	$\pi_6^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_2^{*[1]}(\uparrow)$	0.15
			$\pi_2^{[1]}(\uparrow)\pi_3^{[1]}(\downarrow)\pi_2^{*[1]}(\uparrow)$	0.06
			$\pi_3^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_5^{*[1]}(\uparrow)$	0.14
			$\pi_2^{[1]}(\uparrow)\pi_6^{[1]}(\downarrow)\pi_5^{*[1]}(\uparrow)$	0.05
			$\pi_6^{[1]}(\uparrow)\pi_7^{[1]}(\uparrow)\pi_2^{*[1]}(\downarrow)$	0.07
			$\pi_3^{[1]}(\uparrow)\pi_7^{[1]}(\uparrow)\pi_5^{*[1]}(\downarrow)$	0.06
9	12.92	9.02e-03	$\pi_7^{[0]}\pi_3^{*[1]}(\uparrow)$	0.07
			$\pi_5^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.20
			$\pi_2^{[0]}\pi_1^{*[1]}(\uparrow)$	0.19
			$\pi_2^{[1]}(\uparrow)\pi_7^{[1]}(\uparrow)\pi_6^{*[1]}(\downarrow)$	0.07

TABLE IV: Main valence ionized states of B_g symmetry, with their leading configurations, that give rise to PES spectra reported in Fig. 2.a of the main text. The orbitals involved in the ionization process are highlighted in the left side of figure 2. Up and Down arrows are used to highlight the electronic spin.

Root	ΔE (eV)	Dyson (a.u.)	Configuration	Weight
1	8.01	8.20e-01	$\pi_7^{[1]}(\uparrow)$	0.74
2	8.91	8.34e-01	$\pi_6^{[1]}(\uparrow)$	0.71
3	10.98	6.07e-01	$\pi_5^{[1]}(\uparrow)$	0.47
			$\pi_4^{[1]}(\uparrow)$	0.05
			$\pi_2^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.06
4	11.30	2.57e-03	$\pi_3^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.38
			$\pi_2^{[1]}(\uparrow)\pi_6^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.13
			$\pi_3^{[1]}(\uparrow)\pi_7^{[1]}(\uparrow)\pi_1^{*[1]}(\downarrow)$	0.12
5	11.99	3.38e-02	$\pi_7^{[0]}\pi_5^{*[1]}(\uparrow)$	0.08
			$\pi_7^{[0]}\pi_6^{*[1]}(\uparrow)$	0.10
			$\pi_3^{[1]}(\uparrow)\pi_7^{[1]}(\uparrow)\pi_1^{*[1]}(\downarrow)$	0.12
			$\pi_2^{[1]}(\uparrow)\pi_7^{[1]}(\uparrow)\pi_1^{*[1]}(\downarrow)$	0.14
6	12.12	4.82e-02	$\pi_7^{[0]}\pi_5^{*[1]}(\uparrow)$	0.06
			$\pi_7^{[0]}\pi_6^{*[1]}(\uparrow)$	0.11
			$\pi_3^{[1]}(\uparrow)\pi_7^{[1]}(\uparrow)\pi_1^{*[1]}(\downarrow)$	0.09
			$\pi_2^{[1]}(\uparrow)\pi_7^{[1]}(\uparrow)\pi_1^{*[1]}(\downarrow)$	0.22
7	12.13	7.07e-03	$\pi_3^{[1]}(\uparrow)\pi_6^{[1]}(\uparrow)\pi_1^{*[1]}(\downarrow)$	0.52
8	12.41	2.41e-01	$\pi_4^{[1]}(\uparrow)$	0.20
			$\pi_3^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_2^{*[1]}(\uparrow)$	0.08
			$\pi_2^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.11
			$\pi_2^{[1]}(\uparrow)\pi_5^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.06
			$\pi_6^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_5^{*[1]}(\uparrow)$	0.10
9	12.89	8.09e-02	$\pi_4^{[1]}(\uparrow)$	0.06
			$\pi_3^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_2^{*[1]}(\uparrow)$	0.07
			$\pi_2^{[1]}(\uparrow)\pi_6^{[1]}(\downarrow)\pi_2^{*[1]}(\uparrow)$	0.09
			$\pi_6^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_5^{*[1]}(\uparrow)$	0.07
			$\pi_2^{[1]}(\uparrow)\pi_3^{[1]}(\downarrow)\pi_5^{*[1]}(\uparrow)$	0.09
10	13.04	1.84e-03	$\pi_6^{[0]}\pi_5^{*[1]}(\uparrow)$	0.21
			$\pi_3^{[0]}\pi_5^{*[1]}(\uparrow)$	0.13
			$\pi_3^{[1]}(\uparrow)\pi_6^{[1]}(\uparrow)\pi_2^{*[1]}(\downarrow)$	0.13

TABLE V: Main valence ionized states of A_u symmetry, with their leading configurations, that give rise to PES spectra reported in Fig. 2.a of the main text. The orbitals involved in the ionization process are highlighted in the left side of figure 2. Up and Down arrows are used to highlight the electronic spin.

Root	ΔE (eV)	Dyson (a.u)	Configuration	Weight
1	10.08	2.59e-02	$n_2^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.66
2	10.80	7.42e-03	$n_2^{[1]}(\uparrow)\pi_6^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.49
			$n_2^{[1]}(\uparrow)\pi_7^{[1]}(\uparrow)\pi_1^{*[1]}(\downarrow)$	0.07
3	10.84	3.21e-02	$n_2^{[1]}(\uparrow)\pi_6^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.07
			$n_2^{[1]}(\uparrow)\pi_7^{[1]}(\uparrow)\pi_1^{*[1]}(\downarrow)$	0.53
4	11.27	3.31e-04	$n_2^{[1]}(\uparrow)\pi_6^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.06
			$n_2^{[1]}(\uparrow)\pi_6^{[1]}(\uparrow)\pi_1^{*[1]}(\downarrow)$	0.51
5	12.17	5.17e-01	$n_1^{[1]}(\uparrow)$	0.43
			$n_2^{[1]}(\uparrow)\pi_5^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.11
6	12.44	2.46e-02	$n_2^{[1]}(\uparrow)\pi_6^{[1]}(\downarrow)\pi_2^{*[1]}(\uparrow)$	0.23
			$n_2^{[1]}(\uparrow)\pi_3^{[1]}(\downarrow)\pi_5^{*[1]}(\uparrow)$	0.22
			$n_2^{[1]}(\uparrow)\pi_6^{[1]}(\uparrow)\pi_2^{*[1]}(\downarrow)$	0.07
			$n_2^{[1]}(\uparrow)\pi_3^{[1]}(\uparrow)\pi_5^{*[1]}(\downarrow)$	0.07
7	13.01	2.34e-02	$n_2^{[1]}(\uparrow)\pi_7^{[1]}(\downarrow)\pi_3^{*[1]}(\uparrow)$	0.14
			$n_2^{[1]}(\uparrow)\pi_5^{[1]}(\downarrow)\pi_1^{*[1]}(\uparrow)$	0.16
			$n_2^{[1]}(\uparrow)\pi_2^{[1]}(\downarrow)\pi_6^{*[1]}(\uparrow)$	0.08
			$n_2^{[1]}(\uparrow)\pi_5^{[1]}(\uparrow)\pi_1^{*[1]}(\downarrow)$	0.08

TABLE VI: Main valence ionized states of B_u symmetry, with their leading configurations, that give rise to PES spectra reported in Fig. 2.a of the main text. The orbitals involved in the ionization process are highlighted in the left side of figure 2. Up and Down arrows are used to highlight the electronic spin.

Root	ΔE (eV)	Dyson (a.u)	Configuration	Weight
2	11.36	9.10e-01	$\sigma_2^{[1]}(\uparrow)$	0.68
			$\sigma_1^{[1]}(\uparrow)$	0.11
3	12.43	9.02e-01	$\sigma_1^{[1]}(\uparrow)$	0.36
			$\sigma_4^{[1]}(\uparrow)$	0.29
			$\sigma_3^{[1]}(\uparrow)$	0.22
4	14.04	9.00e-01	$\sigma_3^{[1]}(\uparrow)$	0.58
			$\sigma_1^{[1]}(\uparrow)$	0.34
5	14.52	9.05e-01	$\sigma_4^{[1]}(\uparrow)$	0.57
			$\sigma_2^{[1]}(\uparrow)$	0.19
			$\sigma_1^{[1]}(\uparrow)$	0.11

TABLE VII: Additional σ ionized states belonging to A_g symmetry, with their leading configurations, that give rise to PES spectra reported in Fig. 2.a of the main text. The orbitals involved in the ionization process are highlighted in the right side of figure 2 and the up arrow are used to highlight the electronic spin.

Root	ΔE (eV)	Dyson (a.u)	Configuration	Weight
1	11.33	9.18e-01	$\sigma_6^{[1]}(\uparrow)$	0.48
			$\sigma_8^{[1]}(\uparrow)$	0.40
3	13.17	9.07e-01	$\sigma_7^{[1]}(\uparrow)$	0.41
			$n_1^{[1]}(\uparrow)$	0.40
4	13.89	9.17e-01	$\sigma_8^{[1]}(\uparrow)$	0.41
			$\sigma_6^{[1]}(\uparrow)$	0.29
			$n_1^{[1]}(\uparrow)$	0.11
5	14.80	9.16e-01	$\sigma_5^{[1]}(\uparrow)$	0.46
			$\sigma_7^{[1]}(\uparrow)$	0.31

TABLE VIII: Additional σ ionized states belonging to B_u symmetry, with their leading configurations, that give rise to PES spectra reported in Fig. 2.a of the main text. The orbitals involved in the ionization process are highlighted in the right side of figure 2 and the up arrow are used to highlight the electronic spin.

III. CARBON XPS AND NEXAFS

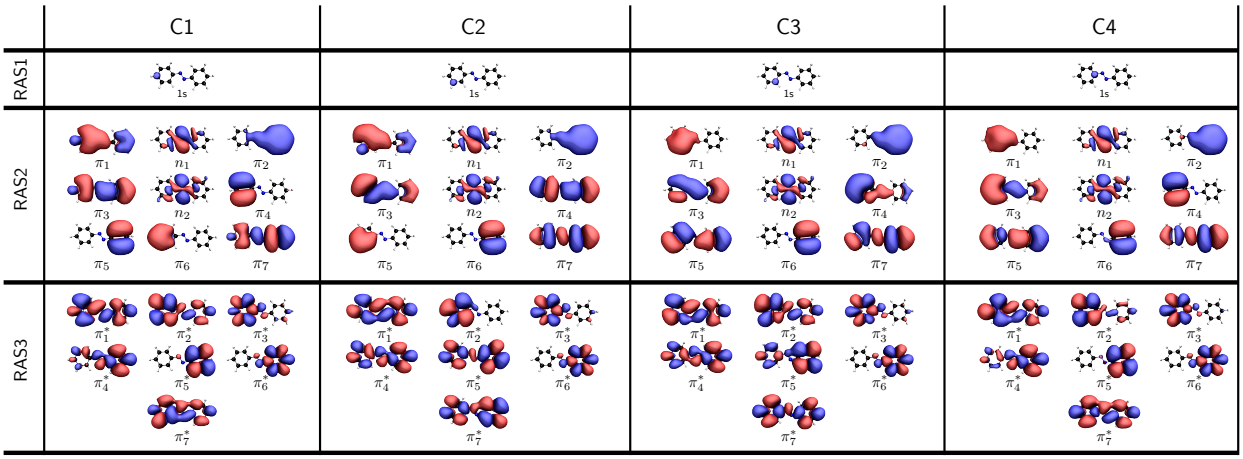


FIG. 3: Active spaces used for the computation of the C K-edge NEXAFS.

Root	ΔE (eV)	$\|TDM\|^2$ (a.u)	Configuration	Weight
1	284.385	3.8799e-03	$1s^{[1]}\pi_1^{*[1]}$	76
2	285.815	1.8243e-04	$1s^{[1]}\pi_2^{*[1]}$	0.44
			$1s^{[1]}\pi_7^{*[1]}$	0.30
4	286.332	2.3916e-03	$1s^{[1]}\pi_7^{*[1]}$	0.32
			$1s^{[1]}\pi_2^{*[1]}$	0.30
			$1s^{[1]}\pi_7^{[1]}\pi_1^{*[2]}$	0.11
5	287.334	3.8315e-04	$1s^{[1]}\pi_7^{[1]}\pi_1^{*[2]}$	0.37
14	289.228	3.6046e-04	$1s^{[1]}\pi_3^{*[1]}$	0.15
			$1s^{[1]}\pi_7^{[1]}\pi_1^{*[1]}\pi_2^{*[1]}$	0.12

TABLE IX: Main core excited states, with their leading configurations, for C1. Only roots with $|TDM|^2$ grater than 10^{-4} a.u. are reported in the table. The orbitals involved in the excitation process are highlighted in figure 3

Root	ΔE (eV)	$\|TDM\|^2$ (a.u)	Configuration	Weight
2	285.308	6.4060e-03	$1s^{[1]}\pi_2^{*[1]}$	0.75
5	286.988	3.0434e-04	$1s^{[1]}\pi_7^{*[1]}$	0.24
			$1s^{[1]}\pi_5^{*[1]}$	0.18
10	288.685	1.4805e-04	$1s^{[1]}\pi_6^{[1]}\pi_1^{*[2]}$	0.13
13	289.299	1.5833e-04	$1s^{[1]}\pi_4^{[1]}\pi_2^{*[1]}\pi_5^{*[1]}$	0.16
			$1s^{[1]}\pi_4^{[1]}\pi_2^{*[1]}\pi_7^{*[1]}$	0.11
15	289.720	2.3497e-04	$1s^{[1]}\pi_7^{[1]}\pi_2^{*[1]}\pi_4^{*[1]}$	0.16

TABLE X: Main core excited states, with their leading configurations, for C2. Only roots with $|TDM|^2$ grater than 10^{-4} a.u. are reported in the table. The orbitals involved in the excitation process are highlighted in figure 3.

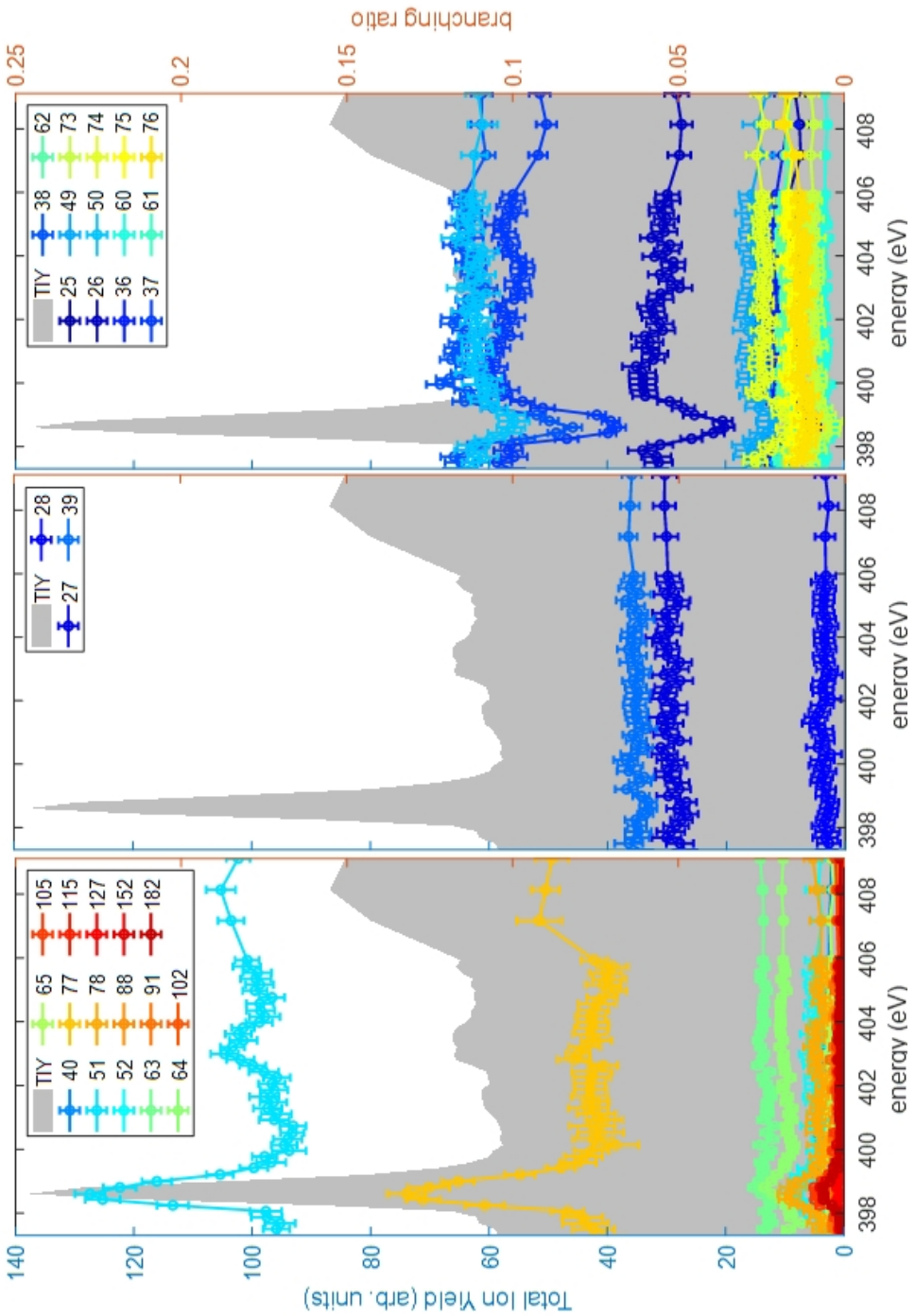
Root	ΔE (eV)	$\|TDM\|^2$ (a.u)	Configuration	Weight
1	284.151	3.5406e-03	$1s^{[1]}\pi_1^{*[1]}$	0.76
2	285.407	2.2593e-03	$1s^{[1]}\pi_2^{*[1]}$	0.62
4	286.397	2.4017e-04	$1s^{[1]}\pi_7^{*[1]}$	0.34
			$1s^{[1]}\pi_4^{[1]}\pi_1^{*[2]}$	0.15
5	287.008	1.0395e-04	$1s^{[1]}\pi_4^{[1]}\pi_1^{*[2]}$	0.35
14	289.059	9.9427e-04	$1s^{[1]}n_2^{[1]}\pi_1^{*[2]}$	0.17
			$1s^{[1]}\pi_3^{*[1]}$	0.16
			$1s^{[1]}\pi_4^{[1]}\pi_1^{*[1]}\pi_2^{*[1]}$	0.11

TABLE XI: Main core excited states, with their leading configurations, for C3. Only roots with $|TDM|^2$ grater than 10^{-4} a.u. are reported in the table. The orbitals involved in the excitation process are highlighted in figure 3.

Root	ΔE (eV)	$\|TDM\|^2$ (a.u)	Configuration	Weight
1	284.813	1.9960e-03	$1s^{[1]}\pi_1^{*[1]}$	0.76
2	286.415	7.9207e-04	$1s^{[1]}\pi_2^{*[1]}$	0.57
			$1s^{[1]}\pi_7^{*[1]}$	0.16
3	286.781	3.3496e-03	$1s^{[1]}\pi_7^{*[1]}$	0.48
			$1s^{[1]}\pi_2^{*[1]}$	0.16
5	287.870	1.6187e-04	$1s^{[1]}\pi_7^{[1]}\pi_1^{*[2]}$	0.34
10	289.079	1.0812e-04	$1s^{[1]}\pi_5^{[1]}\pi_1^{*[2]}$	0.20
13	289.593	7.7559e-04	$1s^{[1]}\pi_3^{*[1]}$	0.12
			$1s^{[1]}\pi_4^{[1]}\pi_1^{*[1]}\pi_2^{*[1]}$	0.10
14	289.633	4.6756e-04	ex.m.	

TABLE XII: Main core excited states, with their leading configurations, for C4. Only roots with $|TDM|^2$ grater than 10^{-4} a.u. are reported in the table. The orbitals involved in the excitation process are highlighted in figure 3.

IV. TOTAL ION YIELD AND BRANCHING RATIOS OF ALL FRAGMENTS AT THE N(1S) AND THE C(1S) EXCITATION/IONIZATION REGIONS



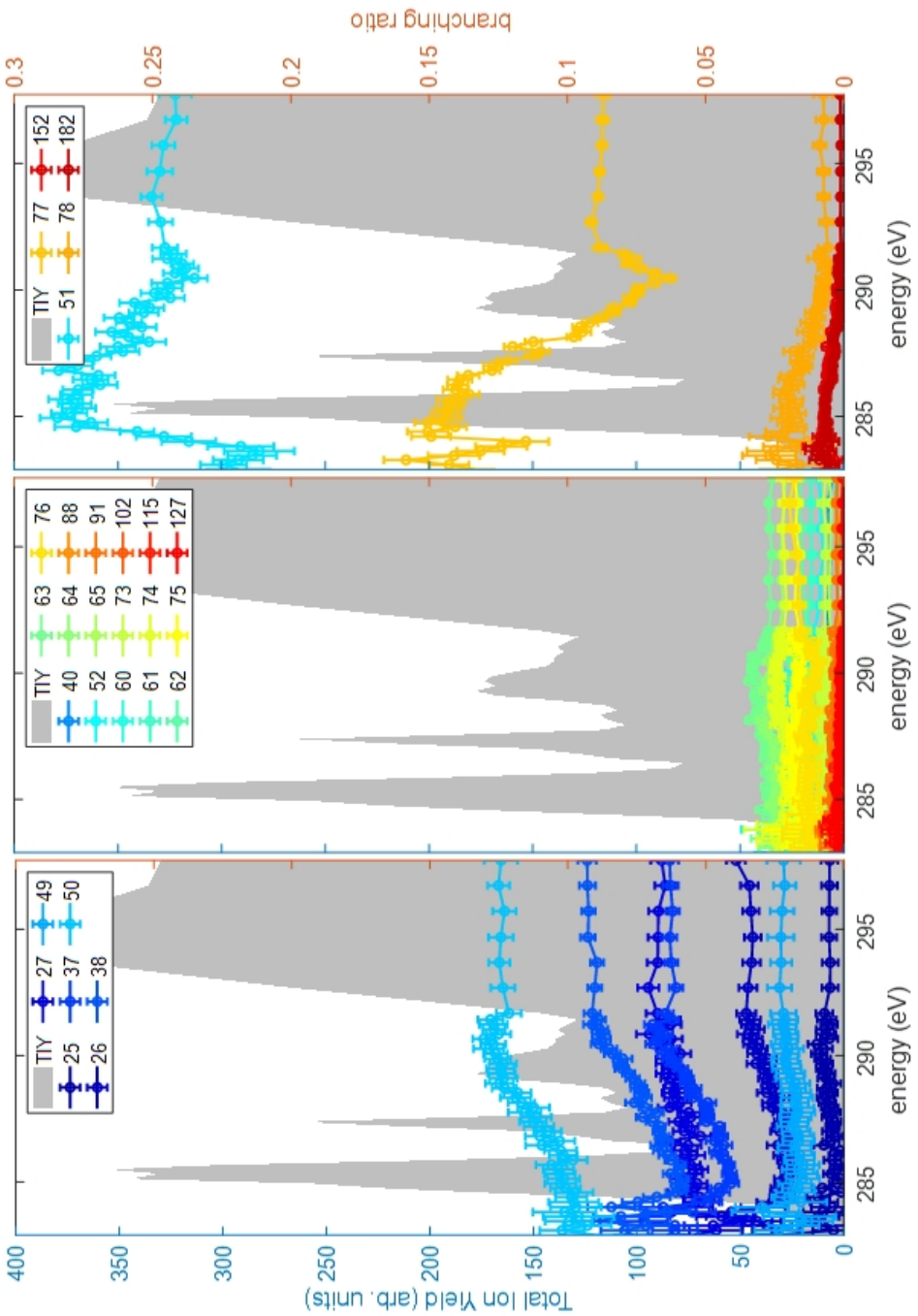


FIG. 4: Total ion yield (TYI, grey shaded area) and branching ratios of all fragments at the N(1s) and C(1s) excitation/ionization regions (coloured lines). Fragments that show a similar behaviour versus photon energy have been displayed in different panels. The error bars represent the confidence of the fit.