

UC Berkeley

UC Berkeley Previously Published Works

Title

Simulation of X-ray transient absorption for following vibrations in coherently ionized F2 molecules

Permalink

<https://escholarship.org/uc/item/5fp1v9nw>

Authors

Dutoi, Anthony D
Leone, Stephen R

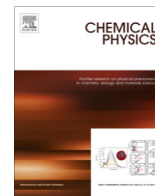
Publication Date

2017

DOI

10.1016/j.chemphys.2016.10.006

Peer reviewed



Simulation of X-ray transient absorption for following vibrations in coherently ionized F₂ molecules



Anthony D. Dutoi^{a,*}, Stephen R. Leone^{b,c}

^a Department of Chemistry, University of the Pacific, Stockton, CA 95211, USA

^b Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

^c Departments of Chemistry and Physics, University of California, Berkeley, CA 94720, USA

ARTICLE INFO

Article history:

Available online 6 October 2016

ABSTRACT

Femtosecond and attosecond X-ray transient absorption experiments are becoming increasingly sophisticated tools for probing nuclear dynamics. In this work, we explore and develop theoretical tools needed for interpretation of such spectra, in order to characterize the vibrational coherences that result from ionizing a molecule in a strong IR field. *Ab initio* data for F₂ is combined with simulations of nuclear dynamics, in order to simulate time-resolved X-ray absorption spectra for vibrational wavepackets after coherent ionization at 0 K and at finite temperature. Dihalogens pose rather difficult electronic structure problems, and the issues encountered in this work will be reflective of those encountered with any core-valence excitation simulation when a bond is breaking. The simulations reveal a strong dependence of the X-ray absorption maximum on the locations of the vibrational wave packets. A Fourier transform of the simulated signal shows features at the overtone frequencies of both the neutral and the cation, which reflect spatial interferences of the vibrational eigenstates. This provides a direct path for implementing ultrafast X-ray spectroscopic methods to visualize coherent nuclear dynamics.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The ability to produce extremely short (<1–10 fs) soft X-ray pulses using the coherent high-harmonic generation (HHG) process [1–3] is allowing the study of molecules at previously unthinkable time scales. Although the apparatuses to perform experiments using such pulses currently exist in only a few laboratories, the fact that these systems are of table-top scale opens the possibility that they will become commonplace in chemistry and physics departments as the necessary expertise advances and spreads. HHG experimental set-ups have been used to probe motions of nuclei in systems with up to five atoms [4], and even the motions of the electrons themselves in a single atomic ion [5].

There is a pressing need to assemble the theoretical tools needed to simulate and understand the time dependence of the X-ray spectra of molecules as they undergo nuclear motion on different electronic surfaces. The essential ingredients for this are (1) the energies and oscillator strengths of core-valence electronic excitations as a function of nuclear coordinates, (2) simulation of the relevant nuclear dynamics, and (3) a framework for modeling an experimental spectrum from this information. Each of these

presents its own challenges, both generally and in the context of a given system of interest. Additionally, each of these topics contains a number of sub-topics. For example, concerning electronic structure, many levels of approximation may be applied, and there are multiple approaches to treating the core-hole lifetimes that will affect precise line shapes. Concerning nuclear dynamics, these might be qualitatively modeled, or they may be predicted by detailed simulation of some physical process. The manner in which these data are incorporated into a simulated spectrum will, of course, depend on the targeted experimental set-up.

In this report we present work on each of these three primary aspects, ultimately producing an X-ray transient absorption spectrum (XTAS) for ionized F₂, in which both the ion and the remaining neutral are undergoing a model dynamic qualitatively consistent with a coherent ionization process. Importantly, we explore the nature of core-valence excitations under bond-breaking conditions, and we develop a straightforward approach to building a spectrum, projecting nuclear wavepackets distributed along spatial coordinates onto the corresponding excitation-energy distribution. This work should be thought of as a starting point, which can be expanded upon in a number of complex directions, most importantly, extension to other molecules.

The study undertaken here is most directly related to recent experiments collecting the XTAS of Br₂ molecules, as a function

* Corresponding author.

E-mail addresses: adutoi@pacific.edu (A.D. Dutoi), srleone@lbl.gov (S.R. Leone).

of delay time after these molecules have been ionized by a strong IR field [6]. The predominant character of these spectra is that the energies of the absorption maxima and minima oscillate in time. As analyzed, this is due to a coherent nuclear dynamic that was initiated during the ionization process. Since the core–valence excitation energy depends on the nuclear separation, the energy of the absorption maximum/minimum oscillates in time with the nuclei. However, the nuclear wavepackets are evolving on two different potential energy surfaces; the production of a coherent wavepacket on the electronic ground-state surface of the cation leaves behind a coherently depleted wavefunction on the ground-state surface of the neutral, which also oscillates. In addition, both of these surfaces are anharmonic, leading to a potentially quite complicated nuclear dance.

Concerning the action of the extreme ultraviolet/soft X-ray probe, the cation has available a new core→valence- π_g^* resonance, leading to an oscillatory enhancement of absorption at lower energy. Meanwhile, assuming that the core→valence- σ_u^* resonance has a different energy for the neutral and the cation (more generally, a different dependence of this energy on internuclear separation), then this drives additional enhancement features, as well as depletion features, which oscillate with either the cation or neutral vibrational wavepackets, respectively. The core–valence excitation energy should not be expected to differ very dramatically between the cation and the neutral. Disentangling these dynamics, which therefore occur in overlapping regions of the spectrum, is vital to learning what kind of vibrational coherences are produced by the strong IR pulse, and therefore gaining insight into the impact of a strong field on a chemical bond, an important topic of several recent experimental investigations [7–12,6].

In the publication of the aforementioned Br₂ experiments, calculations had been done for core–valence excitations of the neutral at different internuclear separations. This already provided convincing evidence that the dominant core–valence excited state is steeply repulsive in the neighborhood of the neutral ground-state equilibrium bond distance, indeed admitting the explanation that the energy of strongest X-ray absorption would be a function of bond distance, and therefore have a significant dependence on delay time. However, this only qualitatively explains depletion features of the spectrum. Furthermore, those simulations had not yet incorporated any information about oscillator strengths. The quality of the calculations is also still uncertain, as they were based on a variant of linear-response density functional theory (LR-DFT), which has yet to be well validated at X-ray energies. Dihalogens and their molecular cations have particularly difficult electronic structures [13,14], with F₂ being highly diradicaloid (bond-broken) already at equilibrium bond distances. Accurate treatment of bond breaking is a known issue for DFT [15,16]. For all these reasons it is important to revisit the XTAS of coherently ionized dihalogens, particularly starting with F₂, with a more systematic and calibrated approach.

2. Electronic structure calculations

All calculations were done with an unmodified test (beta) version of a recent release of the Q-Chem program package [17]. The cc-pVTZ basis is used for all calculations, and an increment of 0.02 Å along the F–F coordinate was used for all electronic surfaces.

2.1. Neutral and cation ground states

A number of high-quality computational [13,14,18–21] and experimental [22–25] studies have been undertaken on F₂ and its molecular cation F₂⁺ (see also references contained in citations). Because of its pathological electronic structure, in which the

mean-field (Hartree–Fock, HF) solution is spin-polarized already at the equilibrium geometry, F₂ has provided an excellent test case for *ab initio* method development. We will not recount all of the findings of these studies, but, given the central role played by the particulars of this system, we reproduce and discuss the most important electronic structure aspects here.

For each level of theory reported, for the neutral and the cation, the position (r_e) and energetic depth (D_e) of the bonding minimum, as well as the harmonic vibrational frequency (ν_e), were determined by a quadratic fit to the lowest three computed points. This is sufficiently accurate for comparison to experiments, in which the vibrational wavefunction of F₂ anyway has a full width at half maximum (FWHM) of about 10× the grid spacing (~ 0.2 Å, as obtained from the experimental ν_e [26]). These fits should provide roughly three decimal places for r_e , and even greater precision for D_e (energy derivative vanishes). Comparison to a quadratic fit with an increment of twice the grid spacing quickly verifies this convergence and establishes that the finite-difference ν_e values are also converged to three significant figures.

We will be satisfied with a level of theory in which the ground-state wavefunctions for the neutral and cation have the correct qualitative structure. It should then be expected that a response theory for the core–valence excitations built on top of these should provide, at a minimum, qualitatively accurate insight.

2.1.1. Neutral

For neutral F₂, coupled-cluster with single and double substitutions (CCSD) is a qualitatively complete wavefunction. On account of the two-particle correlations explicitly in the wavefunction, CCSD is exact (within a basis) for two-electron systems, such as H₂, regardless of the bond distance. CCSD is therefore a good choice of wavefunction for a system whose dominant characteristic is the breaking of a two-electron (single) bond. Furthermore, the “spectator” electrons are correlated at a high enough level that CCSD is known to give quantitatively useful thermochemical results. Fig. 1 compares CCSD against a few other options, including such disastrous results as restricted and unrestricted HF (RHF and UHF, respectively) and second-order Møller–Plesset perturbation theory (MP2).

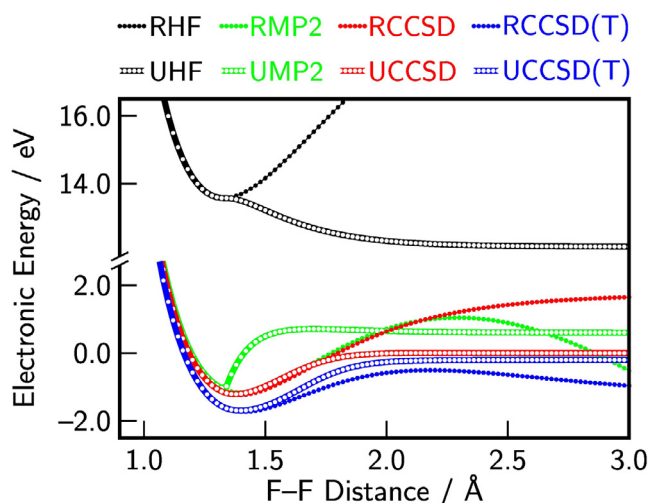


Fig. 1. Ground-state dissociation curves of F₂ at several levels of theory, using both spin restricted and spin unrestricted reference wavefunctions. The only acceptable curves are UCCSD and UCCSD(T). UCCSD is the method chosen for all subsequent calibrations, on account of expense and lack of an associated response theory for UCCSD(T). The points on the curves indicate the bond distances at which calculations were performed. All energies are relative to the UCCSD asymptote.

RHF is overbound because the bond is never allowed to break; the lack of electron correlation admits ionic ($F^+ - F^-$) configurations into the wavefunction even at infinite distance, raising the asymptote. UHF allows the electrons to unpair and mimic correlation (if one of the bonding electrons is on one atom, then the other is on the other atom). A homolytically cleaved bond is eventually obtained, giving an energy that is asymptotically much lower than RHF, but at the expense of having a symmetry-broken, spin-polarized molecule. The necessity of spin polarization in order to break the bond would be represented in a configuration-interaction (CI) language as having forced an admixture of triplet spin, which, by definition is completely diradicaloid at all distances. This destroys some of the interatomic electronic resonance that accounts for the bond, and constitutes an unnatural energetic penalty to unpairing (relative to a pure-singlet CI). This keeps the electrons paired well into the region where the bonding minimum should be, raising the energy there, after which they suddenly unpair. This ultimately prevents the formation of a bonding minimum, as has been famously noted some time ago [13]. At the root of this problem is the fact that electron unpairing occurs relatively early along the internuclear coordinate for F_2 , near where the energetic minimum should be. This is on account of high ionization energy of fluorine; even at short distance, neither F atom is “willing” to be found as F^+ next to an F^- , which arises in a description of independently resonating electrons. This unpairing would occur early in the CI treatment as well, but without the triplet contamination that erases the bonding minimum.

MP2 accounts for electron correlation perturbatively, but bond breaking is not a small perturbation. At smaller nuclear separation, before the bond breaks in earnest, the total magnitude of the MP2 correlation correction agrees very well with the higher-level CCSD calculations, whether using either a restricted or unrestricted reference wavefunction (RMP2 and UMP2, respectively). In the case of RMP2, at long bond stretches, the attempt to account for the large-amplitude change inherent to the bond-breaking wavefunction without self-consistent, iterative optimization of correlation leads to a break-down of the non-variational projective MP2 energy expression, giving a divergent result. For UMP2 the bond breaking is taken care of by localization of half-occupied orbitals. These change character suddenly at about 1.3 Å, when the molecule spin polarizes, and the perturbative correction to the energy suddenly changes at that point. The perturbative correction fails to reintroduce the bonding resonance, and the energy arrives quickly at the asymptote. It is worth noting here, for the sake of discussion to come, that the orbitals themselves are not discontinuous functions of the stretch coordinate; only the manner in which they change (the derivative) is discontinuous. Therefore, the UHF and UMP2 energies are continuous, but have discontinuous derivatives.

CCSD indeed provides normal looking dissociation curves. In the vicinity of the minimum, there is little dependence of the binding potential on the choice between an RHF or UHF reference wavefunction (RCCSD or UCCSD, respectively). This is related to the fact that CCSD is exact for a two-electron problem, so that the choice of orbital representation does not have a major impact on the description of breaking a single bond. However, F_2 is not really a two electron problem, and the difference between these reference wavefunctions affects the predicted dissociation energy. At long distances, the spin-polarized reference is a better description of the dominant character of the wavefunction and, therefore, a better starting point from which to describe the weaker atomic-like correlations. This results in a much lower UCCSD asymptote.

The perturbative inclusion of explicit three-electron correlations [triple substitutions in CCSD(T)] shifts the RCCSD(T) and UCCSD(T) minima down together. The fully self-consistent inclusion of three-electron correlations (CCSDT) provides a yet smaller

Table 1
Comparison of selected computational results to experiment for F_2 .

	r_e (Å)	ν_e/c (cm^{-1})	D_e (eV)	D_e (kJ/mol)
UCCSD	1.377	1070	1.21	117
UCCSD(T)	1.398	968	1.49	144
Experiment [26]	1.412	917	1.66	160

correction [14], suggesting that there are unlikely to be significant correlations beyond triples. For RCCSD(T), as with the RMP2 result, the use of an RHF reference is deleterious in a perturbative calculation, giving unphysical behavior at long bond lengths. We therefore focus on UCCSD(T). Inclusion of triples does not shift the UCCSD asymptote as much as it shifts the minimum, indicating that such correlations are much more important in the molecule than for the individual atoms. Their inclusion affects the depth of the binding well and brings UCCSD(T) into good agreement with experiment, whereby the majority of the remaining error in the bond energy is due to the limited basis set size [14].

The UCCSD and UCCSD(T) bond lengths, vibrational frequencies, and dissociation energies are compared to experiment in Table 1. The F_2 bond is one of the weakest “single bonds” known, having a well depth of less than 2 eV; this and its associated high reactivity are well explained by the forgoing discussion that the bond is substantially broken even at the equilibrium geometry. Therefore, even the good agreement of UCCSD(T) for D_e on an absolute scale (16 kJ/mol \approx 4 kcal/mol) for this subtle and difficult problem does leave something to be desired in terms of the percentage of the well depth recovered, but all of the important electronic physics have clearly been accounted for. The smaller percent error in the vibrational frequency is evidence of the quality of the potential in the neighborhood of the minimum, which is more important in this work. The seemingly odd fact that the shallower UCCSD potential has a larger vibrational frequency is explained by the shorter equilibrium distance, since it is closer to the steeply repulsive part.

We will be satisfied that, since a perturbation treatment of triples on top of UCCSD can provide semi-quantitative results, that UCCSD itself is a qualitatively acceptable wavefunction, especially in the region of the bonding minimum. In addition to the increased computational cost of even perturbative inclusion of triple substitutions, CCSD(T) does not yet have a well-validated associated response theory to allow for the computation of core-valence excitations. We therefore focus any further attention on UCCSD alone.

2.1.2. Cation

For a comparable treatment of F_2^+ , the UCCSD dissociation curve for the cation is shown along with that of the neutral in Fig. 2; the vertical distance between the curves at any point is the ionization energy at that bond stretch.

We should not necessarily expect UCCSD to do as well for the cation as it does for the neutral. As the bond stretches, the original (neutral-like) single bond is still being broken in the σ -symmetry space, and, in addition, a completely separate π -symmetry half anti-bond is being broken. The overall bond order of F_2^+ is 3/2, and it is effectively a strongly interacting three-body problem (two electrons and one hole), dressed by the smaller correlations of the other electrons.

Recently, quite good experimental data have become available for F_2^+ [25], and our computed values of r_e , ν_e , and D_e are compared to these in Table 2. The qualitative result for both computation and experiment is that the bond strengthens and shortens relative to F_2 , since the loss of an antibonding electron increases the bond order. In both computation and experiment, the bond shortens by ~ 0.1 Å, though the computational result is too short by a couple

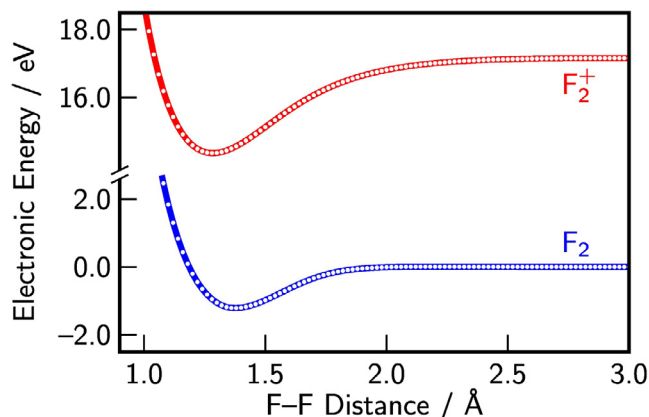


Fig. 2. Ground-state dissociation curves of F_2 and F_2^+ at the UCCSD level of theory. Upon ionization from the antibonding $2\pi_g^*$ orbital, the bond strengthens and shortens. All energies are relative to the neutral asymptote.

Table 2

Comparison of selected computational results to experiment for F_2^+ . The experimental value of D_e is computed as $D_0 + h\nu_e/2$ (from values given in Ref. [25], where D_0 is the measured dissociation energy).

	r_e (Å)	ν_e/c (cm $^{-1}$)	D_e (eV)	D_e (kJ/mol)
UCCSD	1.284	1260	2.80	270
Experiment [25]	1.312	1090	3.40	328

hundredths of an Ångstrom for both F_2 and F_2^+ . The quality of the UCCSD well depth for F_2^+ is either better or worse than that for F_2 , depending on whether judged by ratio (73% \rightarrow 82%) or absolute deviation (0.45 eV \rightarrow 0.60 eV). Similarly, though the absolute error in ν_e is worse for the cation, the percent error is roughly the same for both systems (\sim 16%), illustrating again that the shape of the potentials are indeed most reliable in the neighborhood of the minimum.

Another figure of merit is the adiabatic F_2 ionization energy of 15.69 eV [25], though this will play no role in our simulations, since the ionization process is not explicitly addressed. Given the foregoing comparisons of D_e and ν_e for both F_2 and F_2^+ , the only remaining variable is how well the computed difference of 17.16 eV between the F_2 and the F_2^+ dissociation asymptotes matches the experimental F-atom ionization energy of 17.42 eV [27]. The quite precise match to experiment here (0.26 eV error), in light of lesser agreement for the molecules, is a reflection on the complexity of the F_2 bond, relative to the physics of atomic electrons. This result, and the relatively reliable shape of the potential in the neighborhood of the bond, indicates that the greatest shortcoming in our electronic structure model (UCCSD) is in representing relative energy changes as the bond breakage nears completion (also called the spin-recoupling region). Fortunately, our wavepackets will largely be confined to the regions of the bonding minima of both F_2 and F_2^+ .

2.2. Neutral and cation core–valence states

Core-to-valence excitations are actually electronic resonances that decay on the femtosecond time scale. Nevertheless, they may be preliminarily modeled as stationary states in a finite basis, which will locate their central energies. If no special measures are taken, such as the application of complex absorbing potentials [28], the finite basis prevents the decay of the excited state via autoionization channels. This is certainly acceptable when the spacings between bright resonances are on the order of an eV or more, since

a 1 fs lifetime corresponds to a line-width of about 1 eV, and it is still qualitatively useful even as this limitation is approached.

Such core–valence calculations are far from trivial. A typical electronic structure calculation for the ground state, or the lowest excited states, of a molecule relies on the variational theorem; the easiest eigenvalues to obtain for any matrix are those at the upper or lower boundary of the spectrum. This is made again easier for the lower and ground states of most electronic problems because the largest energy spacings are at the low-energy part of the spectrum. For high-lying states, one must focus on a densely populated part of the spectrum, riddled with dark states of lesser interest. Furthermore, it is necessary to have a relatively precise estimate of the energy of the state of interest, lest the computational work increase dramatically with the size of the energy window considered. Given the complexity of this task, algorithms for this are not generally robust, in particular for automated calculation of surfaces as a function of geometric distortions. Furthermore, at the coupled-cluster level of theory, due to the non-Hermitian projective nature of the energy expression and response theory, once the eigenvalues and eigenstates are obtained, a subsequent iterative calculation is needed to converge any property of each state, such as oscillator strength.

Given the foregoing, it has not presently been feasible to use linear-response theory on top of the UCCSD wavefunctions discussed above (LR-UCCSD), in order to compute the full core–valence potential curves and associated oscillator strengths. However, given the exacting demands of this system on the level of electronic structure model applied, it is certainly necessary to calibrate any proposed model against it. We begin with a discussion of the approximations used in lieu of LR-UCCSD for F_2 and then proceed to F_2^+ , which is further complicated by the breaking of charge symmetry along with spin-space symmetry.

In the molecular-orbital (MO) picture of F_2 , the only unoccupied valence orbital is $2\sigma_u^{(p)}$. [The (p) superscript denotes a linear combination of the $2p_z$ atomic orbitals aligned with the molecular axis.] Though we understand from the above that the MO picture is a largely inadequate concept of F_2 bonding, the most notable bright excitation in this system at equilibrium bond length will still be mostly $1\sigma_g \rightarrow 2\sigma_u^{(p)}$ in character. In F_2^+ , at equilibrium, the vacancy from the ionized electron admits also a bright excitation that is mostly $1\sigma_u^* \rightarrow 2\pi_g^*$ in character. Asymptotically, however, for both systems, the relevant excitations are best thought of as either local excitations on the atomic fragments or charge-transfer (CT) excitations between the fragments.

2.2.1. Neutral

Fig. 3 shows the spectrum of excited states obtained from LR-UCCSD calculations on neutral F_2 . These calculations were done by first converging single eigenstates while restricting their orbital character. The energy of the lowest state that vacates one of the lowest two (core) orbitals is then used to provide an energetic shift to the Hamiltonian necessary to focus on the correct spectral window. A subsequent pseudo-variational calculation captures a fixed number of nearby excited states with smallest shifted energy. Only the spectrum from the second calculation is shown, of which, a subset of these points are the same as from the first calculation (convergence details notwithstanding).

As is evidenced from this plot, LR-UCCSD convergence is spotty, but clear core–valence resonance surfaces are present that suffice for calibration purposes. At short internuclear separation, there are two states that are close in energy, but one begins to rise rapidly just before 1.5 Å (after the point at which the UHF reference for the ground state is known to have localized spins). At intermediate distances of 1.5–2.0 Å, this upper curve has CT character, while the lower curve has the character of local atomic excitations.

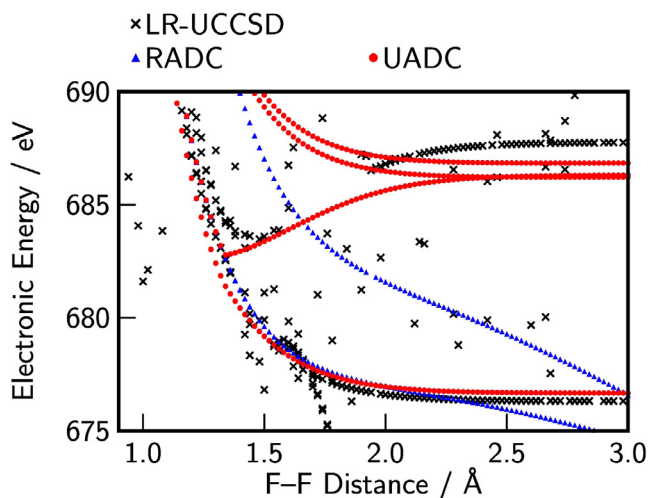


Fig. 3. Excited states in the core–valence energy window for F_2 at three levels of theory. LR-UCCSD is the excited-state analogue of UCCSD, but its convergence is irregular, and it is computationally expensive. RADC reproduces the shape of the lower state (the only bright one) reasonably well near the equilibrium bond distance. UADC provides a better overall representation of this lower curve, especially asymptotically, because the reference is allowed to dissociate to radical fragments; however, it is too low at short distances, due to spin contamination. All energies are relative to the UCCSD neutral ground-state asymptote.

The local character of the lower state persists asymptotically, whereas the CT state mixes with higher states, raising the remainder of its Coulombic asymptote off the plot. The atom-local asymptotic state connects to the core $\rightarrow 2\sigma_u^{*(p)}$ molecular excitation at short distance, and the CT state connects at short distance to a state best described as having multiple excited electrons. A full discussion of the character of these states is found in appendix B. The word “core” is used in the forgoing to be intentionally ambiguous. Since the $1\sigma_g$ and $1\sigma_u$ levels are degenerate, there are two degenerate singlet excitations to the $2\sigma_u^{*(p)}$ orbital, just as there are two spin-allowed atom-local excitations asymptotically. Just as only the $1\sigma_g \rightarrow 2\sigma_u^{*(p)}$ excitation is bright at short distance, asymptotically, only the in-phase superposition of the two atomic excitations will be spectroscopically bright. This bright asymptotic superposition transforms into the $1\sigma_g \rightarrow 2\sigma_u^{*(p)}$ excitation at (very) short distance.

In order to compute smooth surfaces, including associated oscillator strengths, we use the algebraic diagrammatic construction [29] (ADC) for the core–valence states, specifically the extended second-order variant (ADC2x) [30]. ADC has a long history of being applied to highly energetic electronic states and resonances [30–54]. In particular, we employ the recently implemented core–valence separated approximation of ADC2x (CVS-ADC2x) [49,50,53], in which excitations that do not have one index referring to the annihilation of one of the target core orbitals are neglected entirely, this being justified by the large energetic separation of this group of resonances from other states (at least, from those to which there are nonzero Hamiltonian couplings). As with coupled-cluster theory, ADC may be applied on top of a spin-restricted reference (RADC) or an unrestricted reference (UADC). Since all ADC calculations in this work are CVS-ADC2x, we henceforth simply use the abbreviations RADC and UADC.

Both RADC and UADC results for F_2 are included in Fig. 3. State energies are obtained by adding the RADC and UADC excitation energies to the RMP2 or UMP2 ground-state energies, respectively. One can consider MP2 to be the ground-state level of theory associated with ADC excitation energies; the exact same correlation amplitudes that build the MP2 correlation correction are used to dress the excitation energy in ADC.

RADC gives a better representation of the energy of the LR-UCCSD result for core $\rightarrow 2\sigma_u^{*(p)}$ excitations for separations around 1.5 Å, but it starts to perform very poorly a little after 2 Å, consistent with the breakdown of the RMP2 energy for the ground state at this distance (see Fig. 1). The CT state is not at all included with RADC; it is intuitive that the mere existence of a CT state is intimately connected with the breaking of the bond, which is not represented in an RHF reference. The nature of the local and CT states is better reflected in the UADC calculation where the bond is already broken (albeit spin-asymmetric) in the reference; however, there is a clear artifact for the CT-like state at about 1.3 Å, when the UHF reference spin polarizes. It is interesting, though somewhat beyond scope, that the combination of UMP2 + UADC can recover a visually smooth curve for the lower state, even when UMP2 has a discontinuous slope at this point.

Although we might be satisfied with RADC for F_2 (vibrational wavepackets will anyway be confined to separations of less than about 2 Å), it is expedient to look at UADC in more detail because the F_2 σ -bond breaking is similar to that in F_2^+ , where we will soon see that RADC is not acceptable anywhere.

We begin by observing the oscillator strengths of the UADC curves in Fig. 4. Here, only the lower two curves previously mentioned are visible, because only they are bright (at some distance). We have mentioned that the lower LR-UCCSD states change character from core $\rightarrow 2\sigma_u^{*(p)}$ at short distance to atomically local excitations asymptotically, whereas the upper LR-UCCSD states are multiply excited at short distance and have CT character at large separation. This means that the upper LR-UCCSD states should always be dark (dim, technically) and one of the two lower states is always bright. However, the upper and lower states in UADC appear to be mixing and transferring oscillator strength from one to the other, in addition to the unsettling sudden change in slope of the upper states.

The strange appearance of the UADC result is due to the fact that the CT-like states are connecting to the core $\rightarrow 2\sigma_u^{*(p)}$ singlets at short distance, whereas the asymptotically local excitations are connecting to the core $\rightarrow 2\sigma_u^{*(p)}$ triplets. Therein lies the essential difficulty with unrestricted calculations; neither the ground state nor the excited states are spin pure. The asymptotic states are equal mixtures of singlets and triplets. These spin-contaminated

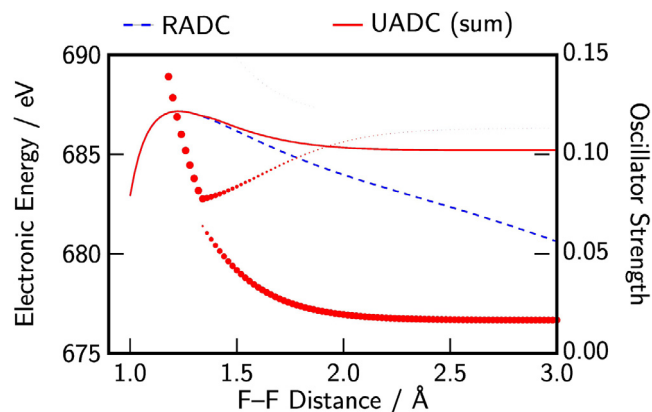


Fig. 4. The energy of the UADC core–valence states for F_2 as a function of bond distance are plotted as points (against the left-hand axis). The area of each point is proportional to the oscillator strength from the ground state. Where the UHF solution breaks symmetry, the separation of local and CT excitations is accompanied by singlet–triplet mixing, artifactually giving two bright states. Per the discussion in the text, the sum of the oscillator strengths of these two UADC states is plotted as a line (against the right-hand axis); this is compared to the oscillator strength of the lowest RADC state (the only bright one). The summed UADC oscillator strengths smoothly connects the RADC result at short distances where it is reliable to a correct flat asymptote.

states mix at short distance to allow spin-pure states when the reference symmetrizes (see appendix B). It is therefore tempting to simply “by hand” assign to the lower UADC energetic curve (which has approximately the shape of the correct bright excitation) the sum of the oscillator strengths from both curves, which each match the character of the correct lower state at one end or the other. As preliminary evidence that we will be able to justify this, Fig. 4 shows this sum compared to the oscillator strength from the RADC calculation. The sum clearly builds a smooth curve with the correct (flat) asymptotic behavior for atom-local excitations, and which agrees well with RADC in the regime where RADC is reliable.

Let us now endeavor to understand in more detail why this is reliable, in order that it will be justified when we do something similar for F_2^+ . Appendix B exposes in detail how the intuitive bright MO-basis excitation at short separation ($1\sigma_g \rightarrow 2\sigma_u^{*(p)}$) should adiabatically transition to being the bright (in-phase) linear combination of atom-local excitations that are well described by the lower UADC state asymptotically. In doing so, it is shown that the lower two UADC excitations at all separations collectively possess all of the oscillator strength inherent to the space of the lowest core–valence transitions. Knowing that all of this oscillator strength should be concentrated in a single transition in the fully symmetric picture allows the use of a limited sum rule, presented in appendix A, to approximate the oscillator strength of the properly symmetric state as the sum of the oscillator strengths of these two lowest states from the symmetry-broken calculation. This can then be assigned to the lower UADC curve, which has roughly the right shape, even though it is too low at short distance by an offset of the singlet–triplet exchange splitting.

It is worth noting before we move on to the cation that we expect excellent agreement between the oscillator strengths that should be obtained from LR-UCCSD and those that are obtained from summed UADC (and RADC below $\sim 2 \text{ \AA}$). To support this assertion, consider first that the excitation energies are in excellent agreement on an *absolute* scale (we have hitherto been focusing on the finer points of the last few eV of a $\sim 700 \text{ eV}$ excitation). It is the values of the oscillator strengths on an *absolute* scale that will enter our simulations later. In light of the forgoing arguments, and the appendices that support in detail that the correct parts of the relevant many-body space have all been accounted for, the only remaining consideration is how the different approximations affect the orbital spaces upon which the many-body spaces are built. Given the excellent agreement of absolute energies among different methods, the description of the orbitals must be sufficient to expect a quantitative comparison for the oscillator strengths.

2.2.2. Cation

For the cation, three different variants of ADC were explored, by starting from differing reference configurations. These are shown along with the spotty LR-UCCSD results in Fig. 5.

The standard “out-of-the-box” RADC variant is unacceptable, as the accuracy is severely degraded already at 1.2 \AA . The increase in complexity of the bonding interaction in going from F_2 to F_2^+ is too much for perturbation theory to correct for, given this starting point.

With UADC, a relatively accurate reflection of the energy of the asymptotically lowest LR-UCCSD state is obtained; this LR-UCCSD state adiabatically connects the core $\rightarrow 2\sigma_u^{*(p)}$ transition at short internuclear separation to an atomic excitation specifically on the neutral fragment at large separation (see appendix B). Once again, we see that the UADC state is too low at short distances because it is erroneously connecting to the quartet instead of the doublet. (To be precise, the quartet is not exactly represented, but rather, this is dominantly a quartet with some admixture of a dark doublet that has a similar exchange energy, due to orbital-extent considerations.)

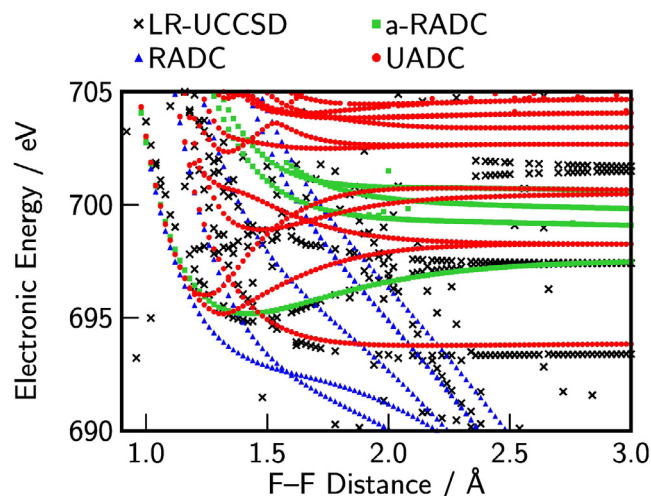


Fig. 5. Excited states in the core–valence energy window for F_2^+ at several levels of theory. The LR-UCCSD convergence is again irregular. RADC is clearly unacceptable at all bond lengths. UADC reproduces the $1s \rightarrow 2\sigma_u^{*(p)}$ excitation energy fairly well, but the UADC $1s \rightarrow 2\pi_g^*$ transition rises too quickly and bifurcates already at very short distances, due to broken symmetry. The final state of the $1s \rightarrow 2\pi_g^*$ transition can be represented well by what we call a-RADC, which uses an unconventional (but spin-restricted) choice of reference; however, the associated oscillator strength is meaningless. All energies are relative to the UCCSD neutral ground-state asymptote.

The structure of the lowest UADC asymptotes are the same as for LR-UCCSD, but higher in energy. While there should be degenerate bright and dark excitations from $1\sigma_g$ and $1\sigma_u^*$ to the empty $2\sigma_u^{*(p)}$ orbital at short distance, these should also mix (along with spin polarization of the empty σ -symmetry valence orbital) to give rise to two energetically distinct states asymptotically, a local excitation on the cation and the local excitation on the neutral just discussed. This excitation on the cation will be asymptotically degenerate with excitation to the empty π -symmetry orbital, which also localizes as an perpendicular atomic $2p$ orbital. These two cation excitations are indeed the next highest LR-UCCSD asymptote. Inward from this asymptote, the excitation to the σ space rises in energy, away from the energy window we will be interested in, whereas the LR-UCCSD state that decreases in energy and connects to the core $\rightarrow 2\pi_g^*$ excitation.

Due to the polarization of the π -symmetry singly occupied orbital already around 1.2 \AA in the UHF reference, the core $\rightarrow 2\pi_g^*$ excitation is pushed up in energy relative to LR-UCCSD at any bond distance much beyond 1.5 \AA , making it not very useful. This brings us to our third choice of reference wavefunction. By manipulating the initial orbital guess, it was possible to converge the RHF reference to an alternate local minimum on the orbital-energy landscape, moving the radical character into the σ space. Though spin-pairing of all but one electron is enforced at all bond distances, right-left symmetry does eventually break, and the reference used asymptotes to a singlet F^+ and a doublet F atom, where the double vacancy that eventually localizes on the F^+ ion keeps σ symmetry throughout. [It was attempted to also force a solution where the double vacancy keeps π symmetry, but we were not successful in doing this by employing standard user input options.] We will call this “asymmetric RADC” (a-RADC); however, this spatial-only symmetry breaking is postponed until after the end of the 3 \AA domain shown in the plots. From this odd choice of reference, excitation into the $2\sigma_u^{*(p)}$ vacancy results in the same final state as a core $\rightarrow 2\pi_g^*$ excitation from the ground-state-like configuration of the cation. From this, we obtain an energetic curve that is very close to that obtained from LR-UCCSD. However, the

oscillator strength from this calculation will be meaningless, since it is from a lower state with the wrong character.

We now have reasonable approximations for the energies of the necessary bright states. The state best characterized (at short distance) as $\text{core} \rightarrow 2\sigma_u^{*(p)}$ is well described by UADC, and our a-RADC calculation reproduces the energetic curve for the state best described (at short distance) as $\text{core} \rightarrow 2\pi_g^*$. The focus then falls on obtaining the oscillator strengths for these states along each curve. These will all be obtained from the UADC calculations, but, in each case, we will have to sum the oscillator strengths from two artifactually split and/or mixed UADC states.

Fig. 6 shows the UADC oscillator strengths. One of the characteristics that is clear on this de-cluttered plot is that each of the excitations at short distance splits into two states at about 1.2 Å. This is due to the effect on the core of the sudden polarization of the valence orbitals. The previously degenerate $1\sigma_g$ and $1\sigma_u^*$ orbitals mix with even a tiny perturbation to create completely localized $1s$ orbitals on each of the atoms. The excitations from these then differentiate in energy, due to their asymmetric electronic environment. The valence orbitals polarize more slowly, however as shown in Fig. 7.

Let us first discuss the excitations to the π -symmetry orbitals. For the $\text{core} \rightarrow 2\pi_g^*$ state, our a-RADC calculation for the energy uses a completely different transition, so the a-RADC oscillator strengths are not useful; the energy is valid because we reached the same final state by accordingly adjusting the reference. In the properly symmetric picture, the $1\sigma_u^* \rightarrow 2\pi_g^*$ state should mix with the bright $1\sigma_g \rightarrow 2\pi_u$ transition that opens up as the hole localizes, causing it to change smoothly to being an atom-local excitation on the cationic fragment asymptotically. This state possesses all of the oscillator strength in space of excitations to the π orbitals at all distances (the orthogonal state is either CT-like or multiply excited, depending on distance). As discussed in appendix B, at all distances, this sole bright state may be approximated as a linear combination of the two UADC excitations that describe transitions of site-localized core orbitals into the empty orbital that begins as $2\pi_g^*$ and becomes a site-localized p orbital (9β in Fig. 7). Therefore, by the limited sum-rule argument of appendix A, we only need to add the oscillator strengths of these two UADC states to recover the intensity of the properly symmetric transition.

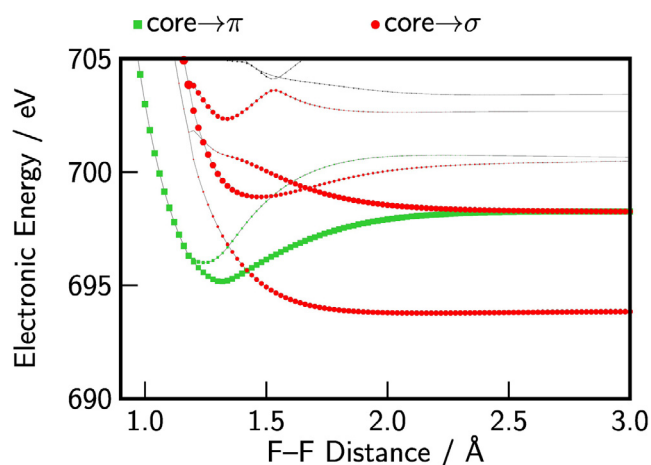


Fig. 6. The energy of the UADC core-valence states of F_2^+ as a function of bond distance. The area of each point is proportional to the oscillator strength from the ground state. The differentiation by symbol shape and color is to visually separate excitations to the π valence space from those to the σ valence space. Lines illustrate the behavior of bright states in regions where they lose their intensity. At the point when the UHF reference breaks symmetry, each state splits into two because polarization of the odd valence electron affects the core levels.

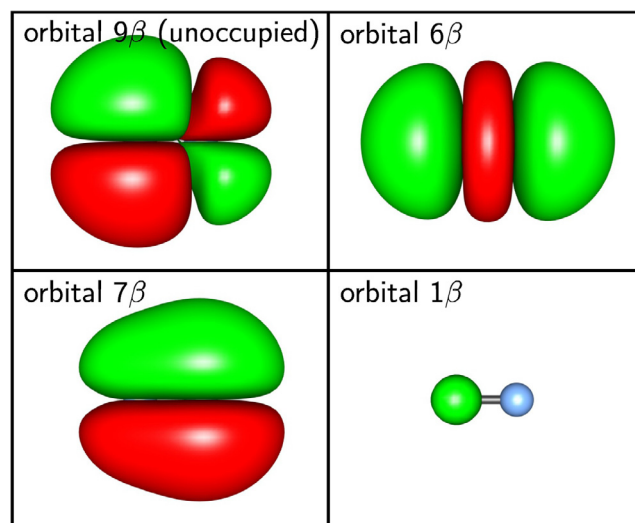


Fig. 7. Selected polarized UHF canonical orbitals of F_2^+ at 1.4 Å separation (density isosurface at 0.02 atomic units). The molecule has lost a spin-down (“ β ”) electron, and only β orbitals are shown. The nearly degenerate core orbitals (e.g., 1β) localize completely at the point of symmetry breaking (at 1.2 Å). However, the F_2^+ valence orbitals are still substantially delocalized near and just beyond the equilibrium bond length. For example, the 6β bonding orbital ($\sim 2\sigma_g^{(p)}$) is only negligibly polarized at 1.4 Å. Orbital 9β ($\sim 2\pi_g^*$) best describes from where the electron was taken, and 7β ($\sim 2\pi_u$) describes the remaining occupied part of the π -symmetry space.

The situation with the core excitations to the σ -symmetry valence orbitals appears more complicated. In addition to the splitting of each bright transition into two transitions of different energies (from distinguishable core states), we again have artificial spin-symmetry mixings. We can again fall back on the same robust abstract arguments as for the neutral. In the exact, symmetric picture, the F_2^+ bond breaking wavefunction can be described largely in the same language as the F_2 bond breaking, but where the spin-orbital state of the π -symmetry valence vacancy is entangled with the spins on the atoms (see appendix B). Asymptotically, the two atom-local excitations are energetically distinguishable because one will occur on a neutral and the other on a cation. These have the same character as the lower two asymptotically flat UADC states in Fig. 6. The state with the next highest energy asymptotically is from a CT state that transfers an electron from the neutral to the cation, such that it is also flat. The fourth state that branches off from this family of curves is a CT state that generates a dication and an anion, and this rises quickly in energy.

The two asymptotically local UADC states connect to the dark degenerate $\text{core} \rightarrow 2\sigma_u^{*(p)}$ quartets at short distance, whereas the CT states connect to the degenerate $\text{core} \rightarrow 2\sigma_u^{*(p)}$ doublets, similar to the situation with the neutral. As with the neutral, in the properly symmetric picture, at short distance, there is one bright transition, here described as a *doublet* $1\sigma_g \rightarrow 2\sigma_u^{*(p)}$ transition in the MO basis. For UADC at short separation, the oscillator strength for this state is spread over the two upper states that adiabatically connect to the CT states. As the bond breaks (in the symmetric picture), the $1\sigma_u^* \rightarrow 2\sigma_g^{(p)}$ excitation opens up; this mixes with the $1\sigma_g \rightarrow 2\sigma_u^{*(p)}$ excitation to create two bright, atomically local excitations. This is represented in UADC through the polarization of orbitals and increase of oscillator strength for the lower local excitations. Therefore, the correct oscillator strength for the lowest bright excitation to the σ -symmetry valence space in the symmetric picture is simply that of the lowest UADC state at long distance, and it is approximately the sum of the two upper UADC states at short distance (more detail in appendix B). This is easily executed by

summing the oscillator strengths of all but one of the UADC states at all distances, where the one left out is the asymptotically local cation excitation (second lowest asymptote).

2.3. *Ab initio* data for nuclear dynamics and X-ray spectra

According to the specific model of strong-field ionization that will be discussed in the coming section, the relevant nuclear dynamics will take place on the UCCSD electronic ground-state surfaces shown in Fig. 2 for neutral F_2 and its molecular cation F_2^+ . The relative displacement of the bonding minima upon ionization will contribute significantly to the character of the nuclear dynamics on these surfaces.

For the purpose of following these dynamics using X-ray spectroscopy, Fig. 8 shows the result of the preceding prescriptions for assigning reasonable oscillator strengths to the core–valence energetic curves that were identified to mimic those from LR-UCCSD for both species. Even single occupancy of the very antibonding $2\sigma_u^{*(p)}$ orbital in either the neutral or the cation creates a purely repulsive state, which has similar shape for both molecules. Excitation to the $2\pi_g^*$ orbital of the cation, however, restores the valence electronic structure of the neutral, and this core–valence state is seen to be vibrationally bound, and it is therefore much less steep in the neighborhood of 1.3 Å, around which the cation will oscillate. The steepness of these core–valence curves will dramatically impact the appearance of the spectrum.

3. Wavepacket dynamics

In contrast to the *ab initio* calculations, the nuclear dynamics of a one-dimensional oscillator are quite simple. There are two essential steps, defining the initial conditions and propagating the nuclear wavefunction. The core excitation that is probing the nuclear motion is considered to be effectively instantaneous on this time scale, and this will be considered separately in the next section.

The process that initiates the nuclear dynamics is modeled here as a sudden and completely coherent ionization, assuming that the molecule is originally prepared in a well-defined vibrational eigenstate on the neutral surface. A portion of the neutral vibrational wavefunction for vibrational quantum number ν is simply transferred to the cation surface, leaving also some fraction of the

now-distorted vibrational wavefunction on the neutral surface. The spectra obtained for each value of ν may be used to build finite-temperature simulations.

We imitate (only qualitatively) the Lochfrass process [8,9,6], in which ionization is more efficient at shorter bond distances because the antibonding orbital from which the electron comes is pushed up in energy there. To do this we define two vibrational wavepackets at the time origin by multiplying the vibrational wavefunction on the neutral surface (for given ν) by two complementary functions

$$\begin{aligned}\Psi_{\text{neutral}}(R, t = 0) &= f(R)^{1/2} \Psi_{\text{neutral}, \nu}(R) \\ \Psi_{\text{cation}}(R, t = 0) &= (1 - f(R))^{1/2} \Psi_{\text{neutral}, \nu}(R)\end{aligned}\quad (1)$$

where f is a real, positive function, whose value is bounded between 0 and 1. This leaves the nuclear density (electronic coordinates integrated over) unchanged, relative to that of the original vibrational state. The only other requirement that we will place on f is that it should be small at short distances, tend towards unity at longer distances, and pass through 1/2 at approximately the equilibrium distance of the neutral. Therefore, we choose for convenience $f(R) = \frac{1}{2}[1 + \text{erf}(20((R/\text{Å}) - 1.38))]$ because $r_e \approx 1.38$ Å for the neutral, and the multiplier of 20 was found to result in wavefunctions with roughly symmetric character on both the neutral and cation surfaces, as seen in Fig. 9 for an initial $\nu = 0$. This model assumes that the ionizing IR field has an equivalent effect on all orientations of the molecule. The extent to which this is a good model is a topic that is ripe for investigation by XTAS experiments.

In order to define the initial wavepackets for the neutral and the cation, we first solve for $\Psi_{\text{neutral}, \nu}$ on the UCCSD neutral electronic surface discussed in section 2.1 and shown in Figs. 2 and 9. The wavefunction is represented on the same grid as is used for computing the electronic potential-energy surface, having a spacing of 0.02 Å and a domain of 0.8–3.0 Å. The Fourier transform of the vibrational wavefunction on this grid (with periodic boundaries) is used to represent the momentum operator. With the vibrational Hamiltonian now completely defined, the ground-state ($\nu = 0$) wavefunction is converged upon by diagonalizing repeated Lanczos projections [55] of the Hamiltonian (with dimension of 50), until the norm of the change in the wavefunction between iterations is less than 1×10^{-3} , meaning the energy is converged to

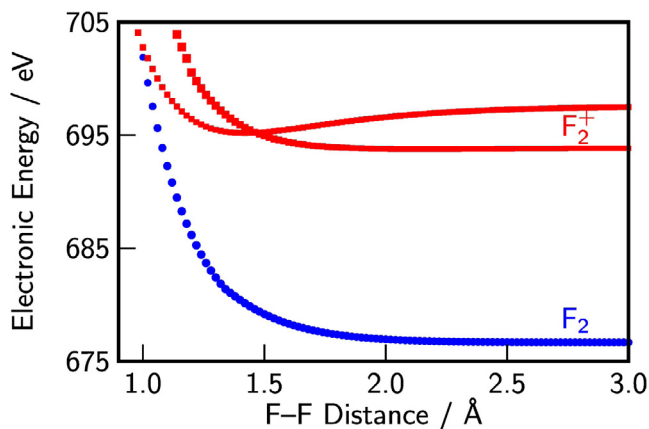


Fig. 8. The bright core-excited states for F_2 and F_2^+ that will be used in the XTAS simulation. The area of each point is proportional to its oscillator strength. There is one state for the neutral which is $1\sigma_g \rightarrow 2\sigma_u^{*(p)}$ in character at short separations, connecting to in-phase degenerate atomic excitations at long distances. The cation has two bright states, best described as $1\sigma_g \rightarrow 2\sigma_u^{*(p)}$ and $1\sigma_u \rightarrow 2\pi_g^*$, connecting to local excitations on the neutral and the cation, respectively.

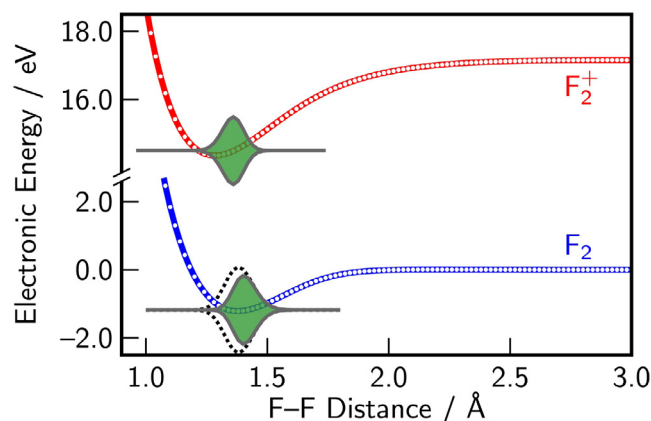


Fig. 9. The initial nuclear wavepackets on the neutral and cation electronic ground-state surfaces, assuming ionization from $\nu = 0$ of the neutral. The sum of their densities is equal to the density of the neutral $\nu = 0$ wavefunction. This has been divided into an “inner” portion on the cation surface and an “outer” portion on the neutral surface, in order to imitate the Lochfrass process. The $\nu = 0$ state is represented by the dotted envelope in the background. The vertical spread of each envelope is proportional to the wavefunction amplitude. Each state has uniform phase (same sign everywhere).

roughly six digits of precision. The same procedure is used to obtain the excited vibrational states ($\nu > 0$), with the previously obtained lower states projected out of each successive calculation. Since states only up to $\nu = 3$ will be needed for the temperatures in this work, this is sufficiently stable. Relative to the ground state, these lowest excitations are computed to be at (in cm^{-1}) 1079, 2129, and 3136 (the harmonic fundamental from Table 1 is 1070 cm^{-1}).

Once $\Psi_{\text{neutral}, \nu}$ has been obtained, eq. (1) defines the initial states at the time of coherent ionization. These have non-trivial time evolution on their respective UCCSD neutral and cation surfaces. The time evolution is handled via the split-operator algorithm [56]. A time-step of 5×10^{-3} fs was found to be sufficient to have no noticeable impact on the results.

4. Spectrum construction

An overview of the procedure by which the spectrum is constructed for a given delay between strong-field ionization and X-ray excitation is shown in Fig. 10, for a molecule that starts in vibrational state with $\nu = 0$ before ionization. To first order in the probe intensity, there are no interferences between the two possible excitations of the cation; similarly, excitation of the neutral and cation do not interfere. Therefore the absorption spectrum at any time is the incoherent sum of the contributions from these

three transitions. The background X-ray absorption from an un-ionized sample in the same vibrational state is also subtracted from this spectrum, in order to obtain a signal that is relatively independent of the total strong-field ionization rate (to the extent that this is decoupled from the shapes of the wavepackets produced), and reflective of the manner in which shot-to-shot noise is cancelled experimentally.

The XTAS from simulations beginning with different vibrational eigenstates of the neutral may be Boltzmann weighted and incoherently summed to construct finite-temperature results. Vibrational states with less than 1% population in the Boltzmann distribution are neglected, and the truncated distribution is normalized to unit probability.

We now turn to the manner in which an individual instantaneous XTAS contribution may be computed for a wavepacket on one of the two electronic ground-state surfaces (neutral or cation) being excited to (one of) the core-excited surface(s), for which we know both the energy and oscillator strength at every point. The definition of oscillator strength implicitly averages over rotational degrees of freedom for isotropically oriented molecules, which we have already presumed are all exhibiting the same vibrational dynamics. If we assume that nuclear motion is not important over the duration of the short X-ray pulse, then absorption $A(R)$ at each nuclear separation R contributes incoherently to the absorption at the respective photon energy $E(R)$ ($h\nu_{\text{X-ray}}$) for that separation. This contribution $A(R)$ should be proportional to the product of the instantaneous nuclear density and the oscillator strength as functions of R . The total (energy-integrated) absorption should therefore be obtainable by integrating $A(R)$ over the nuclear coordinate; however, the contributions from different nuclear separations will contribute to an energy-dispersed absorption signal $\tilde{A}(E)$ over an energy window of variable size, depending on the relative slopes of the ground and excited states, but this $\tilde{A}(E)$ should have the same integral as $A(R)$.

We now have one function $A(R)$ that describes absorption as a function of nuclear distance, and another $E(R)$ that describes excitation energy as a function of distance, and we must convert these to absorption as a function of energy $\tilde{A}(E)$. All of these are represented in Fig. 10 for a 100 fs delay after ionization, along with the corresponding nuclear wavepackets on each surface at that time, for a molecule that started with $\nu = 0$. Noting that $E(R)$ is a monotonic function over the region of interest, it can be inverted to define $R(E)$ such that the method of integration by substitution gives for the integrated absorption

$$\int_{R_1}^{R_2} dR A(R) = \int_{E(R_2)}^{E(R_1)} dE \tilde{A}(E) \quad (2)$$

where

$$\tilde{A}(E) = -A(R(E)) \left(\frac{dR}{dE} \right) \quad (3)$$

The negative sign here is a direct consequence of the negative slope of $E(R)$, which motivates us to switch the limits of integration on the right-hand side to run from low to high energy, thus also ensuring that we arrive at a positive distribution for $\tilde{A}(E)$. Eq. (3) is rather intuitive, as it indicates that a small slope of $E(R)$ will cause absorption over a large stretch domain to coalesce into a narrow energy window. This is illustrated in Fig. 10, where the component from the excitation with a shallower slope gives a sharper contribution.

A final technical detail concerns the summation of the component signals. Each of the $\tilde{A}(E)$ components at a given time was obtained by computing the necessary derivative in Eq. (3) at each R on the nuclear-coordinate grid (finite difference, using the two

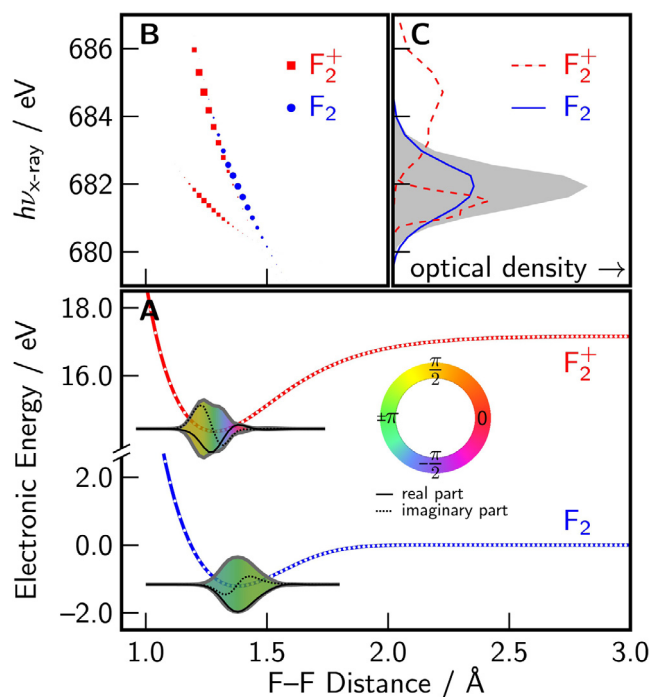


Fig. 10. Overview of the XTAS simulation procedure. (A) The initial wavepackets on the neutral and cation ground-state surfaces from Fig. 9 are shown now 100 fs after the model ionization event. The vertical spread of the envelope is proportional to wavefunction amplitude. Complex phase information is illustrated in two ways. First, the real and imaginary parts are plotted separately. Second, the phase angle (ϕ in the scalar $pe^{i\phi}$) is mapped onto the color wheel. (B) The excited state surfaces of Fig. 8 have their respective ground state curves subtracted from them, to obtain a vertical axis in photon energy $h\nu_{\text{X-ray}}$, retaining the horizontal axis in (A). The oscillator strengths at each point have been scaled by the relevant nuclear densities from panel (A), and this determines the areas of the points. (C) Instantaneous values of the components comprising the differential XTAS. Each of the three components from panel (B) is projected onto the vertical photon-energy axis using the procedure discussed in the text. The horizontal axis is used to represent the optical density as a function of energy. A shallower slope of the excited-state surface results in a sharper absorption feature. The shaded curve represents the background signal from a completely un-ionized sample.

adjacent points), taking the product of this with the oscillator strength and density at that R , and assigning the resulting \tilde{A} value to the appropriate energy. This results in points that are not evenly spaced in energy, with different energy points used for different spectrum contributions. For the purposes of adding several functions together, a uniform grid was introduced with a spacing of 0.01 eV, and the value of each spectrum component was computed for each energy grid point via simple linear interpolation from the existing data at the nearest non-uniform points on either side of that energy.

The three components of the signal and the subtractive background were computed at intervals of 0.2 fs. Fig. 11 shows the individual components for a 0 K simulation (starting with the molecule in $v = 0$ only). The components arising from the core excitations of F_2^+ are rendered as purely positive. The background has been subtracted only from the component arising from F_2 core excitations, leaving an almost purely negative feature, since this is thought of as leading to missing absorption in the overall background-subtracted signal. The combination of all three absorption components with the background subtraction is shown in Fig. 12 for both 0 K and for 950 K.

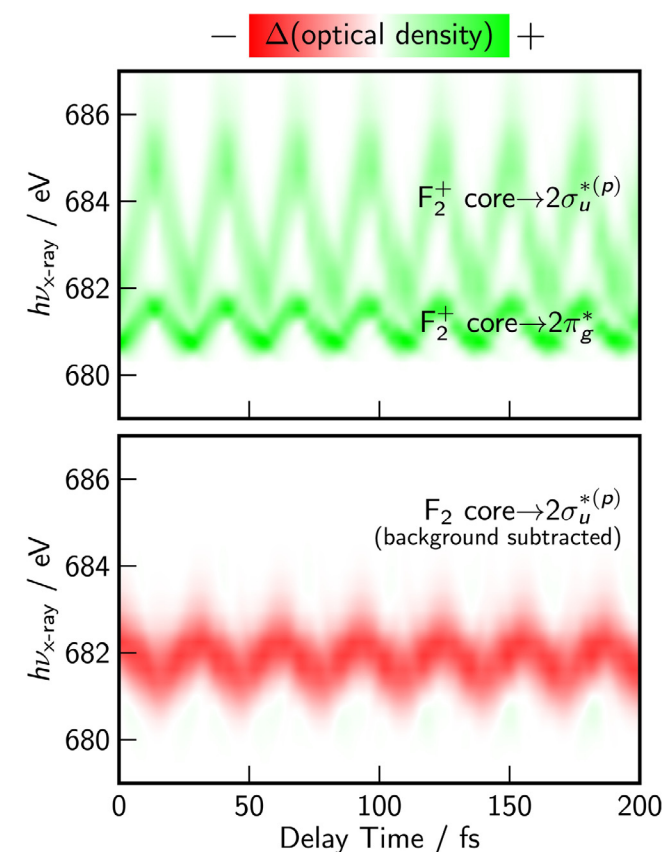


Fig. 11. The first 200 fs of the XTAS components due to F_2 and F_2^+ , obtained from the initial wavepackets shown in Fig. 9. These components will overlap in the complete signal. The background absorption from an un-ionized sample has been subtracted from the signal for the neutral, leaving an almost uniformly negative contribution. Oscillations of the two cation core-excitation signals share the vibrational frequency of that surface, which is slightly higher than that of the neutral surface. Both neutral and cation wavepackets begin with an inward trajectory. The neutral feature moves to lower energies in the first moments because the region of lower-energy excitation is increasingly depleted. The cation enhancement features move to higher energies because the inward trajectory is towards increasing core-valence excitation energy. The F_2^+ core $\rightarrow 2\pi_g^*$ feature is the sharpest because its excitation-energy slope is the most shallow. While the slopes for the two core $\rightarrow 2\sigma_u^{(p)*}$ excitations energies are similar, the motion on the cation surface is larger in amplitude, spreading the corresponding XTAS feature over a broader energy range. The change in optical density is in arbitrary units on a linear scale.

5. Discussion and conclusions

Figs. 11 and 12 show the behavior of the XTAS signal for F_2 as a function of delay since the coherent ionization event that initiates the dynamics on both surfaces. Fig. 11 shows the separate components from the cation and the neutral, which are summed in Fig. 12. As the wavepackets on each of the neutral and cation surfaces oscillate, the energies of the core-valence absorption maxima oscillate with them, since the core-excited states are all steeply repulsive near the neutral and cation equilibrium distances.

Fig. 11 shows the two most important distinctions between the neutral and cation dynamics; the cation oscillation has a larger amplitude and a slightly higher frequency. The sizes of the oscillations in the spectra are functions of both the amplitude of oscillations along the F–F coordinate and also the slopes of the core-valence energies with respect to this distance. This is well illustrated in Fig. 11, in light of what is seen in Fig. 10, by comparison of the two enhancement features from the cation, and by comparison of the $2\sigma_u^{(p)}$ excitations of the neutral and the cation. The two cation features share the same underlying vibrational dynamics, but have different slopes for the core-valence energies, resulting in different size oscillations in the energy of the absorption maxima. The two $2\sigma_u^{(p)}$ excitations have similar slope, but the vibrational oscillations themselves differ in amplitude.

In the overall signal in Fig. 12, the same features are present as in Fig. 11, but with some clear overlap of the signals from the cation and neutral molecules. Interestingly, this overlap is less after a delay of about 100 fs, as the oscillations of absorption maxima come temporarily into phase. A simulation was also done for molecules at 950 K, where the first vibrationally excited state of F_2 is roughly 20% populated, commensurate with the thermal excitation of Br_2 molecules at room temperature. The fact that the spectrum is not substantially different than at 0 K is easy to rationalize, since the dominant character is simply oscillation of the absorption maxima. For the stretches represented in the lowest few vibrational states, the electronic surfaces are still nearly harmonic, such that oscillation frequency is largely independent of amplitude. Ionization from a molecule in a higher vibrational wavefunction will result in wavepackets with slightly different centers (leading to some blurring and loss of signal coherence), but they will still oscillate with roughly the same frequencies and have the same phases (because they start on the same side of the potential minima).

Finer details of the dynamic, including the extent to which the surfaces are not harmonic may be drawn out by Fourier transformation. The spectral density (squared norm) of the Fourier transform is shown in Fig. 13 for the 0 K signal of Fig. 12; the result at 950 K is substantially similar. The artificial coloration of the all-positive data in Fig. 13 was actually produced by separately Fourier transforming the cation and neutral plots in Fig. 11 and overlaying the results, providing a plot that is imperceptibly different from the result of directly transforming the data in Fig. 12 (except for color). This works because the frequency content of the neutral and cation contributions are non-overlapping here. In order to obtain reasonably sharp peaks, a simulation of 2000 fs was transformed, with data for the transform being generated at intervals of 2 fs. This corresponds to roughly 70 oscillation periods for either the cation or neutral, over which it should be reasonable to neglect vibrational relaxation and dephasing in a dilute gas, especially since the systems have no dipoles. Since the discrete Fourier transform must assume boundary conditions that are not necessarily met by an arbitrary signal (in this case, periodicity of 2000 fs), the signal was first multiplied by the window function $[\sin(\pi t/(2000\text{fs}))]^2$, which damps the signal and its derivative to zero at the time-domain ends. Applied to a perfectly periodic signal, this window

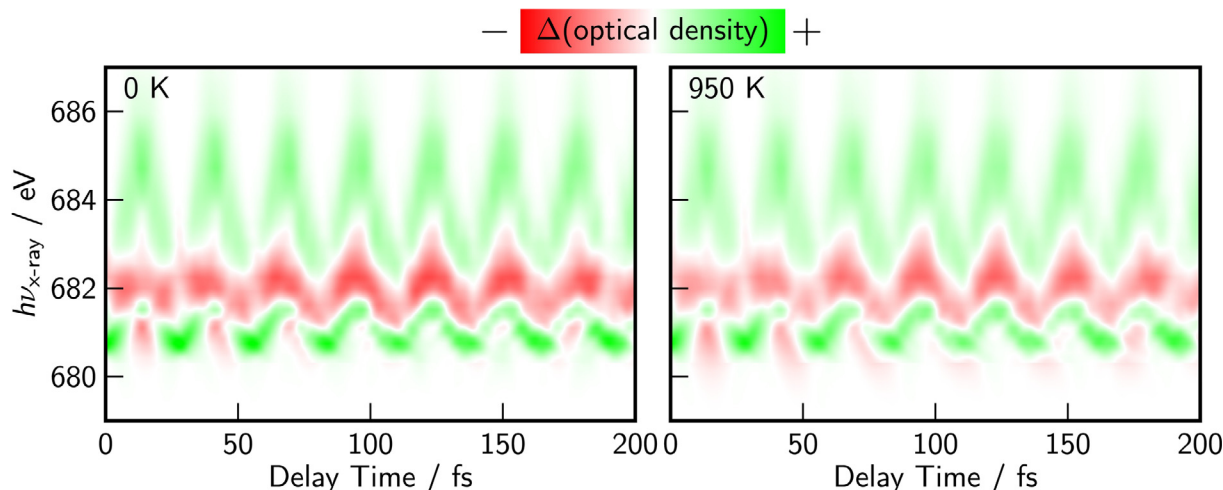


Fig. 12. The first 200 fs of the background-subtracted XTAS at 0 K and at 950 K (initially 20% vibrationally excited). In both spectra, the features from the individual F_2 and F_2^+ systems are most easily identified at times around 100–150 fs, when their oscillations come into phase. Due to the approximate amplitude independence of oscillation frequency on nearly harmonic potentials, the difference between the 0 K and finite-temperature cases is minimal.

would introduce line broadenings on the order of 500 GHz (20 cm^{-1}), but here it prevents the inherent broadenings from transforming finite data from being even larger and more jagged. Only positive frequencies are shown in Fig. 13, since the spectral density is a symmetric function.

Though it will be interesting to explore the phase information of the Fourier transform in future work, there is already much that can be seen from only the magnitude we have plotted. The structure of the Fourier transform is best understood by considering the time dependence of the nuclear density along the vibrational coordinate for either the neutral or cation. If the vibrational eigenfunctions on one of the surfaces belong to the set $\{\phi_\nu(R)\}$, each with energy E_ν , and the coherent wavepacket at $t=0$ is $\Psi(R, 0) = \sum_\nu c_\nu \phi_\nu(R)$, then the dynamic density $\rho(R, t)$ associated with the evolution of this wavepacket may be decomposed as

$$\rho(R, t) = \sum_{\nu_1, \nu_2} [c_{\nu_1}^* c_{\nu_2}] [\phi_{\nu_1}^*(R) \phi_{\nu_2}(R)] e^{i2\pi t(E_{\nu_1} - E_{\nu_2})/h} \quad (4)$$

Therefore, every significantly represented pair of vibrational eigenstates makes a contribution to the dynamic density in the form of the product (interference) distribution $\phi_{\nu_1}^*(R) \phi_{\nu_2}(R)$, which has nodes at any location where either of the individual wavefunctions have nodes. Due to the monotonic mapping of X-ray energy to bond stretch over the region of interest, as seen in Fig. 10, the spatial structure of the density is projected onto the X-ray energy axis in the spectrum. In a system with anharmonicities, each such interference contributes to a unique Fourier component of frequency $(E_{\nu_1} - E_{\nu_2})/h$, giving rise to the individual vertical lines in Fig. 13.

Drawing on the above, any difference of two vibrational energies will always be roughly an integer multiple of the $\nu = 0 \rightarrow 1$ fundamental for that surface, but slightly shifted. This explains why the vertical lines of Fig. 13 separate into clusters spaced approximately by the fundamental frequencies, with a larger spacing between the clusters of cation lines. For the overtones ($|\Delta\nu| > 1$), the anharmonicities are more evident as higher states are accessed, causing multiple lines to appear. The lines indeed carry the expected structure of the vibrational interferences. For example, there are three clearly visible $|\Delta\nu| = 3$ overtone lines for the neutral at around 3000 cm^{-1} . The $0 \rightarrow 3$ line at highest energy has $0 + 3$ zeros, consistent with the combined number of nodes for the $\nu = 0$ and $\nu = 3$ wavefunctions, whereas the $1 \rightarrow 4$ line immediately to its left has $1 + 4$ zeros, with two occurring very close together. Such structure is evident elsewhere too, but node counting is often obscured if the nodes are close together, which is very likely for even overtones ($|\Delta\nu| = 2, 4, \dots$). The $0 \rightarrow \nu$ lines (highest in energy for each overtone cluster) for $|\Delta\nu| = 1, 2, 3$, and 4 of the neutral all have the expected number of zeros clearly visible.

The implication of the foregoing for imaging vibrational wavepackets is immediate, if it would be possible to extract such a clean Fourier transform from an experiment. With the magnitude and phase of each of the density interference terms, along with quantitative theory for both the ground and core-excited electronic surfaces, complete wavepacket reconstruction should be possible, and this will be the subject of future work. This offers a route forward for using XTAS as a means to unravel the structure of the vibrational states produced by a strong-field ionization, thereby probing the strong-field ionization process itself. This article then makes clear the role that the modeling of ground-state and core-valence surfaces may play in using XTAS experiments to

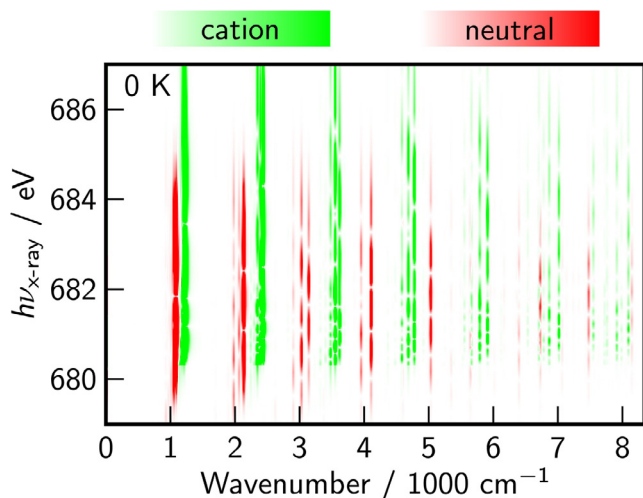


Fig. 13. Spectral density of the Fourier transform of the 0 K plot in Fig. 12 along the time axis (extended and windowed to remove artifacts – see text); the frequency axis is converted to wavenumber. Two families of overtone clusters (red and green) can be seen, having spacings of roughly the harmonic fundamentals of the neutral or cation (1070 cm^{-1} and 1260 cm^{-1} , respectively, in this work). The formation of clusters is the result of anharmonicities, giving different overtones with the same $|\Delta\nu|$ slightly different frequencies. Nodal structures of the vibrational wavefunctions are reflected in the structure of each line along the X-ray energy axis. The coloration is artificial and does not indicate sign.

investigate and even image nuclear dynamics. With regard to extension beyond coherently ionized dihalogens, one feature that is expected to be common is precisely the presence of a core–valence resonance with a steep dependence of the energy on the bond length, which allows for the important mapping of stretch coordinate to energy. This rests only on the assumption that, in the neighborhood of any bond, one of the low-lying resonances is to an antibonding state, making it potentially quite general.

The approximations made in this work for the electronic structure may have rendered it non-quantitative, but the conclusions drawn from the qualitative picture are on solid footing. The biggest unknown with regard to predictive simulations is how line-widths due to autoionization will impact the conclusions. Some of what has been learned about dealing with the substantial broken-bond character of F_2 will also translate directly to techniques for proceeding to Br_2 (though spin–orbit coupling will need to be modeled in order to compare to experiment, since the core electrons come from a near-degenerate d shell). The insight into how to deal with XTAS simulation when bonds are breaking should prove to be generally valuable as well. However, given the difficulty of constructing the spectrum here from symmetry-broken methods, more general investigations will likely require advances in using linear-response (and perhaps multi-reference) coupled-cluster theory to converge properties of high-lying states.

We now have a fully functioning framework by which diatomic nuclear dynamics and *ab initio* calculations may be translated into a time-domain XTAS signal, subject to refinement of the model for producing the initial wavepackets, which may also be extended to include orientational dependencies. This can now be used to provide insight into results from experiments focused on understanding the production of these wavepackets.

Acknowledgements

This manuscript is dedicated to Lorenz S. Cederbaum, on the occasion of his 70th birthday. A special acknowledgement goes to Michael Wormit, an academic grandchild of Cederbaum, for all of his help in recent years, both theoretical and practical; since his untimely departure, he has been missed as a friend as well as a colleague. The authors would also like to thank Andrew Attar, Evgeny Epifanovsky, Jan Wenzel, Andreas Dreuw, Anna Krylov, Joseph Subotnik, and Martin Head-Gordon for helpful conversations. This work was supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Visiting Faculty Program (VFP). A. D. D. gratefully acknowledges start-up funding from the University of the Pacific. S. R. L. gratefully acknowledges support by the Director, Office of Science, Office of Basic Energy Sciences, of the US Department of Energy under contract DE-AC02-05CH11231 and by the National Science Foundation under grant CHE-1361226.

Appendix A. Application of sum rules to state mixings

Let us define a “brightness operator” \hat{B} so that an expectation value taken with respect to it for state $|n\rangle$ is roughly proportional to its oscillator strength f_n against the ground state $|0\rangle$.

$$\begin{aligned} f_n &\propto \sum_{w \in \{x,y,z\}} |\langle n | \hat{\mu}_w | 0 \rangle|^2 \\ &\propto \langle n | \hat{B} | n \rangle \\ \hat{B} &= \sum_{w \in \{x,y,z\}} \hat{\mu}_w | 0 \rangle \langle 0 | \hat{\mu}_w \end{aligned} \quad (5)$$

where $\hat{\mu}_w$ is the component of the dipole operator in the w direction. In addition to some fundamental constants, a factor of the

transition energy in the definition of the oscillator strength has been suppressed. As long as we will only consider states whose energies differ only by a small amount (relative to the ground state), this too may be considered part of the proportionality constant. This is a somewhat special-purpose operator in that it references the ground state of whatever system is under discussion, but we may now use some well-paved operator logic to make the necessary point in two sentences. If, in some set of orthonormal states $\{|m\rangle\}$, there is only one bright state $|n\rangle$, then, for the oscillator strength of $|n\rangle$, we have

$$f_n = \langle n | \hat{B} | n \rangle = \sum_{|m\rangle} \langle m | \hat{B} | m \rangle = \text{Tr}_V[\hat{B}] \quad (6)$$

where Tr_V denotes the trace in the subspace $V = \text{span}\{|m\rangle\}$. Since a trace is invariant with respect to basis representation, oscillator strengths of several states which are mixtures of the basis states $\{|m\rangle\}$ may be added to obtain the oscillator strength of $|n\rangle$, so long as (1) the states summed fully span the bright state of interest $|n\rangle$, and (2) there are no other bright states orthogonal to $|n\rangle$ in the space V .

The application of this rule will be to situations in appendix B where oscillator strength transfers artifactually between two energy adiabats, whereby the bright portion of each curve would be adiabatically connected to each other in a theory with a less restricted state space, and where these two halves are themselves faithful representations of the state of interest away from this region of mixing. In such situations of artificial state mixing, the oscillator strength of the state that should connect the two halves (at the higher level of theory) may be approximately reconstructed in this region of mixing by simply summing the oscillator strengths of the two states involved at the lower level. The rule can also be applied when a single bright state is artifactually split, due to a symmetry breaking.

Appendix B. Symmetry breaking and state mixings

B.1. Notation and definitions for neutral F_2

In order to efficiently discuss the connection between X-ray excitations in the proper, symmetric CI picture of F_2 bond breaking and the results of symmetry-broken calculations, we first condense the notation. The $1\sigma_g$ and $1\sigma_u^*$ MOs will be labeled by their inversion symmetry alone, in lower case, as $|g\rangle$ and $|u\rangle$, respectively. The atom-localized core orbitals will be labeled as $|a\rangle$ and $|b\rangle$, and these may be constructed from the core MOs as

$$\begin{aligned} |a\rangle &= 2^{-1/2} [|g\rangle + |u\rangle] \\ |b\rangle &= 2^{-1/2} [|g\rangle - |u\rangle] \end{aligned} \quad (7)$$

Upper-case labels $|G\rangle$ and $|U\rangle$ will be used for the valence σ -symmetry MOs formed from the atomic $2p_z$ orbitals, $2\sigma_g^{(p)}$ and $2\sigma_u^{*(p)}$, respectively. For these, it will be necessary to define general polarized orbitals that continuously connect the bonding MO and the atomic $2p_z$ orbitals. Let the extent of polarization be characterized by the real, positive number $0 \leq d \leq 2^{-1/2}$, which is a coefficient in the expansion of these orbitals.

$$\begin{aligned} |A\rangle &= c|G\rangle + d|U\rangle \\ |B\rangle &= c|G\rangle - d|U\rangle \end{aligned} \quad (8)$$

where $c = \sqrt{1 - d^2}$. It is presumed that the real-valued MOs form an orthonormal set, and that their global phases are such that $|a\rangle$ and $|A\rangle$ localize to the same atom. Since we will be exciting core electrons into the empty part of the valence space, it is also necessary to define polarized versions of the antibonding MO as

orthogonal complements to the polarized bonding orbitals $|B\rangle$ and $|A\rangle$, respectively, as

$$\begin{aligned} |A^*\rangle &= c|U\rangle + d|G\rangle \\ |B^*\rangle &= c|U\rangle - d|G\rangle \end{aligned} \quad (9)$$

When $d = 0$, we have $|A\rangle = |B\rangle = |G\rangle$, and $|A^*\rangle = |B^*\rangle = |U\rangle$. If $d = 2^{-1/2}$, then, to within a phase, $|A\rangle$ is the same as $|A^*\rangle$, and $|B\rangle$ is the same as $|B^*\rangle$, these being the atomic $2p_z$ orbitals at infinite separation.

The textbook MO-theory electronic configuration that describes two bonded F atoms is now written as

$$\begin{aligned} |\text{MO-theory}\rangle &= |g\bar{g}u\bar{u}\dots G\bar{G}\rangle \\ &= |a\bar{a}b\bar{b}\dots G\bar{G}\rangle \end{aligned} \quad (10)$$

where the single bond is the result of double occupancy of the bonding orbital $|G\rangle$ and non-occupancy of the antibonding orbital $|U\rangle$. Collection of an ordered string of orbital labels into a single ket implies an antisymmetrized product, and a bar over an orbital label indicates that an electron in that spatial state is spin down, and it is spin-up otherwise. Orbitals that participate neither in the bond nor the core excitation are omitted from the discussion entirely (i.e., $2\sigma_g^{(s)}, 2\sigma_u^{(s)}, 2\pi_u^{(x)}, 2\pi_g^{*(x)}, 2\pi_u^{(y)}, 2\pi_g^{*(y)}$), referenced here only by the ellipsis, which will also henceforth be suppressed. Complete occupancy of both bonding and antibonding linear combinations of a pair of atomic orbitals in a diatomic is mathematically equivalent to placing a pair of electrons in each of the original atomic orbitals, and this is the origin of the second line of eq. (10), which is a reflection of the lack of a core-electron bond.

As the molecule dissociates, $|G\rangle$ and $|U\rangle$ become degenerate, and the powerful repulsion between electrons introduces a configuration interaction (CI). The bond-breaking singlet wavefunction is dominated by the following mixing

$$|\text{CI}(R)\rangle = \sqrt{1-t^2}|g\bar{g}u\bar{u}G\bar{G}\rangle - t|g\bar{g}u\bar{u}U\bar{U}\rangle \quad (11)$$

where the real, positive number $0 \leq t \leq 2^{-1/2}$ depends on the inter-nuclear distance R . Asymptotically, $t \rightarrow 2^{-1/2}$, and we obtain after some algebra

$$|\text{CI}(\infty)\rangle = 2^{-1/2} [|a\bar{a}b\bar{b}A\bar{B}\rangle - |a\bar{a}b\bar{b}A\bar{B}\rangle] \quad (12)$$

in which $|A\rangle$ and $|B\rangle$ are the completely polarized orbitals ($d = 2^{-1/2}$). Equal occupation of both the doubly bonded and doubly antibonded configurations results in no net bond, and the wave interferences localize the previously bonding electrons on two radical fragments that are spin coupled to form an overall singlet, leaving the spin on either site indefinite.

We have previously noted that the singlet F_2 molecule has substantial broken bond character even at equilibrium, and the configurational mixing just discussed is the proper way to describe this. However, the limited number of correlation amplitudes available in an affordable calculation has rendered a spin-polarized UHF wavefunction as our best choice for a single reference configuration. At all separations, there are two degenerate choices, which may be written as

$$\begin{aligned} |\text{UHF}\downarrow(R)\rangle &= |a\bar{a}b\bar{b}A\bar{B}\rangle \\ |\text{UHF}\uparrow(R)\rangle &= |a\bar{a}b\bar{b}A\bar{B}\rangle \end{aligned} \quad (13)$$

where the extent of polarization d of the orbitals $|A\rangle$ and $|B\rangle$ depends on the distance R . At small separation, $d = 0$, and both solutions are identical to the symmetric $|g\bar{g}u\bar{u}G\bar{G}\rangle$ configuration, but, beyond a critical distance only one of the two symmetry-broken configurations is selected (randomly, by numerical noise). As the orbitals polarize ($d \neq 0$), the doubly excited MO configuration $|g\bar{g}u\bar{u}U\bar{U}\rangle$ is mixed in implicitly, as well as the singly excited triplet, leading to

what is known as spin contamination. Asymptotically, $d \rightarrow 2^{-1/2}$, and the wavefunction converges to one of the two major configurations that comprises the proper CI wavefunction, approximated in eq. (12). In a minimal basis, it would have the same energy as the wavefunction in eq. (12), but it would be an equal mixture of singlet- and triplet-recoupled atoms, whereby the singlet and triplet are themselves degenerate as $R \rightarrow \infty$.

Let us now introduce a condensed notation for the core-excitation basis in either the symmetric or symmetry-broken picture. For example, let

$$\begin{aligned} |\bar{a} \rightarrow \bar{A}^*\rangle &= \hat{\psi}_A^\dagger \hat{\psi}_a |a\bar{a}b\bar{b}A\bar{B}\rangle = |a\bar{A}^*b\bar{b}A\bar{B}\rangle \\ |g \rightarrow U\rangle &= \hat{\psi}_U^\dagger \hat{\psi}_g |g\bar{g}u\bar{u}G\bar{G}\rangle = |U\bar{g}u\bar{u}G\bar{G}\rangle \end{aligned} \quad (14)$$

where the operators $\hat{\psi}$ and $\hat{\psi}^\dagger$ are the usual Fermionic field operators for electrons. The left-hand side labels the transition in a notation similar to that used in spectroscopy, and the right-hand side provides the formal definitions needed to execute the algebra. In the ‘‘spectroscopic shorthand,’’ it is permissible here to suppress writing the reference onto which the excitation is applied because, for example, if $|\bar{A}^*\rangle$ is the destination orbital, it means that the reference was $|\dots A\bar{B}\rangle$ because $|\bar{B}\rangle$ is the spin-down one-electron state that is orthogonal to the target $|\bar{A}^*\rangle$ (so they can both be occupied in the final configuration). The only configurations ever acted upon by an excitation operator here are the references $|\dots A\bar{B}\rangle, |\dots \bar{A}\bar{B}\rangle, |\dots G\bar{G}\rangle$, and $|\dots U\bar{U}\rangle$, with fully occupied cores.

Regardless of whether the symmetric MO basis or the basis of polarized orbitals is chosen, for this minimal model, we have a space of eight core-valence excitations that conserve M_S (two reference configurations in either basis, each with two valence vacancies, and two core electrons of appropriate spin to fill each vacancy). Let us first introduce spin-pure excitations in the MO basis. This provides four singlets, $|S_{g\rightarrow U}\rangle, |S_{u\rightarrow U}\rangle, |S_{g\rightarrow C}\rangle$, and $|S_{u\rightarrow C}\rangle$, and four triplets, $|T_{g\rightarrow U}\rangle, |T_{u\rightarrow U}\rangle, |T_{g\rightarrow C}\rangle$, and $|T_{u\rightarrow C}\rangle$, of the forms

$$\begin{aligned} |S_{g\rightarrow U}\rangle &= 2^{-1/2} [|g \rightarrow U\rangle + |\bar{g} \rightarrow \bar{U}\rangle] = 2^{-1/2} [|U\bar{g}u\bar{u}G\bar{G}\rangle - |\bar{U}\bar{g}u\bar{u}G\bar{G}\rangle] \\ |T_{g\rightarrow U}\rangle &= 2^{-1/2} [|g \rightarrow U\rangle - |\bar{g} \rightarrow \bar{U}\rangle] = 2^{-1/2} [|U\bar{g}u\bar{u}G\bar{G}\rangle + |\bar{U}\bar{g}u\bar{u}G\bar{G}\rangle] \end{aligned} \quad (15)$$

This basis provides a bedrock in terms of spin-symmetry and oscillator strength, into which we may resolve the eight excitations from the symmetry-broken UHF references. The excitations from the configuration $|\text{UHF}\downarrow\rangle$ are

$$\begin{aligned} |\bar{a} \rightarrow \bar{A}^*\rangle &= \frac{1}{2} \{c [|S_{g\rightarrow U}\rangle - |T_{g\rightarrow U}\rangle + |S_{u\rightarrow U}\rangle - |T_{u\rightarrow U}\rangle] - d [|S_{g\rightarrow C}\rangle - |T_{g\rightarrow C}\rangle + |S_{u\rightarrow C}\rangle - |T_{u\rightarrow C}\rangle]\} \\ |b \rightarrow B^*\rangle &= \frac{1}{2} \{c [|S_{g\rightarrow U}\rangle + |T_{g\rightarrow U}\rangle - |S_{u\rightarrow U}\rangle - |T_{u\rightarrow U}\rangle] + d [|S_{g\rightarrow C}\rangle + |T_{g\rightarrow C}\rangle - |S_{u\rightarrow C}\rangle - |T_{u\rightarrow C}\rangle]\} \\ |a \rightarrow B^*\rangle &= \frac{1}{2} \{c [|S_{g\rightarrow U}\rangle + |T_{g\rightarrow U}\rangle + |S_{u\rightarrow U}\rangle + |T_{u\rightarrow U}\rangle] + d [|S_{g\rightarrow C}\rangle + |T_{g\rightarrow C}\rangle + |S_{u\rightarrow C}\rangle + |T_{u\rightarrow C}\rangle]\} \\ |\bar{b} \rightarrow \bar{A}^*\rangle &= \frac{1}{2} \{c [|S_{g\rightarrow U}\rangle - |T_{g\rightarrow U}\rangle - |S_{u\rightarrow U}\rangle + |T_{u\rightarrow U}\rangle] - d [|S_{g\rightarrow C}\rangle - |T_{g\rightarrow C}\rangle - |S_{u\rightarrow C}\rangle + |T_{u\rightarrow C}\rangle]\} \end{aligned} \quad (16)$$

and the excitations from the configuration $|\text{UHF}\uparrow\rangle$ are

$$\begin{aligned} |a \rightarrow A^*\rangle &= -\frac{1}{2} \{c [|S_{g\rightarrow U}\rangle + |T_{g\rightarrow U}\rangle + |S_{u\rightarrow U}\rangle + |T_{u\rightarrow U}\rangle] - d [|S_{g\rightarrow C}\rangle + |T_{g\rightarrow C}\rangle + |S_{u\rightarrow C}\rangle + |T_{u\rightarrow C}\rangle]\} \\ |\bar{b} \rightarrow \bar{B}^*\rangle &= -\frac{1}{2} \{c [|S_{g\rightarrow U}\rangle - |T_{g\rightarrow U}\rangle - |S_{u\rightarrow U}\rangle - |T_{u\rightarrow U}\rangle] + d [|S_{g\rightarrow C}\rangle - |T_{g\rightarrow C}\rangle - |S_{u\rightarrow C}\rangle - |T_{u\rightarrow C}\rangle]\} \\ |\bar{a} \rightarrow \bar{B}^*\rangle &= -\frac{1}{2} \{c [|S_{g\rightarrow U}\rangle - |T_{g\rightarrow U}\rangle + |S_{u\rightarrow U}\rangle - |T_{u\rightarrow U}\rangle] + d [|S_{g\rightarrow C}\rangle - |T_{g\rightarrow C}\rangle + |S_{u\rightarrow C}\rangle - |T_{u\rightarrow C}\rangle]\} \\ |b \rightarrow A^*\rangle &= -\frac{1}{2} \{c [|S_{g\rightarrow U}\rangle + |T_{g\rightarrow U}\rangle - |S_{u\rightarrow U}\rangle - |T_{u\rightarrow U}\rangle] - d [|S_{g\rightarrow C}\rangle + |T_{g\rightarrow C}\rangle - |S_{u\rightarrow C}\rangle - |T_{u\rightarrow C}\rangle]\} \end{aligned} \quad (17)$$

In order to relate the excitations in the symmetry-broken picture to those in the properly symmetric picture, we immediately note that we may eliminate the triplet components by defining the following superpositions

$$\begin{aligned} |S_{a\rightarrow A^*}\rangle &= 2^{-1/2} [|\bar{a} \rightarrow \bar{A}^*\rangle - |a \rightarrow A^*\rangle] \\ |S_{b\rightarrow B^*}\rangle &= 2^{-1/2} [|b \rightarrow B^*\rangle - |\bar{b} \rightarrow \bar{B}^*\rangle] \\ |S_{\bar{a}\rightarrow B^*}\rangle &= 2^{-1/2} [|a \rightarrow B^*\rangle - |\bar{a} \rightarrow \bar{B}^*\rangle] \\ |S_{\bar{b}\rightarrow A^*}\rangle &= 2^{-1/2} [|\bar{b} \rightarrow \bar{A}^*\rangle - |b \rightarrow A^*\rangle] \end{aligned} \quad (18)$$

These states mix excitations from each of the two symmetry-broken references, and the form of these is recognized as being that of spin-angular-momentum-conserving local excitation operators (in-phase promotion of spins) acting on the subtractive superposition of polarized references that asymptotes to that given in eq. (12). These states represent pure singlets that are either atom-local or CT in character asymptotically, and they provide a good basis for thinking about the excitation process at large interatomic separation.

At large separation, the two local excitations $|S_{a \rightarrow A^*}\rangle$ and $|S_{b \rightarrow B^*}\rangle$ and the two CT excitations $|S_{a \rightarrow B^*}\rangle$ and $|S_{b \rightarrow A^*}\rangle$ each form degenerate pairs. Taking only the bright, in-phase linear combination of these (which can be determined by detailed consideration of orbital phases) provides superpositions of the only two MO-basis excitations that are bright on account of spin and space symmetry, $|S_{g \rightarrow U}\rangle$ and $|S_{u \rightarrow G}\rangle$.

$$\begin{aligned} |\text{local}\rangle &= 2^{-1/2} [|S_{a \rightarrow A^*}\rangle + |S_{b \rightarrow B^*}\rangle] = c|S_{g \rightarrow U}\rangle - d|S_{u \rightarrow G}\rangle \\ |\text{CT}\rangle &= 2^{-1/2} [|S_{a \rightarrow B^*}\rangle + |S_{b \rightarrow A^*}\rangle] = c|S_{g \rightarrow U}\rangle + d|S_{u \rightarrow G}\rangle \end{aligned} \quad (19)$$

These two states are not orthogonal on account of the overlap of $|A^*\rangle$ with $|B^*\rangle$, and they both contain a larger component of $|S_{g \rightarrow U}\rangle$ than $|S_{u \rightarrow G}\rangle$, except at infinite separation where orbital polarization is complete.

B.2. Discussion of neutral F_2

The foregoing gives us now crucial insight into the relationship of local and CT excitations at large separation to the MO-picture excitations at short distance. At short distance, the $|\text{local}\rangle$ and the $|\text{CT}\rangle$ state mix as the energy difference between them lessens, and the intuitive bright $|S_{g \rightarrow U}\rangle$ excitation in the MO picture may be recovered as

$$|S_{g \rightarrow U}\rangle \propto |\text{local}\rangle + |\text{CT}\rangle \quad (20)$$

The fact that $|\text{local}\rangle$ and $|\text{CT}\rangle$ become highly overlapping as the bond distance shortens simply means that the orthogonal complement

$$|S_{u \rightarrow G}\rangle \propto |\text{local}\rangle - |\text{CT}\rangle \quad (21)$$

is not well represented as a superposition of single-electron excitations on top of the ground-state. Indeed, per the definition of $|S_{u \rightarrow G}\rangle$ as a single excitation relative to the $|\dots UU\rangle$ configuration, this is best thought of as a multiply excited state, relative to the ground state. Although it is a valid member of the Hilbert space, it is spectroscopically dim (formally allowed but effectively dark). At short distance, the core-excited eigenstates are best described as $|S_{g \rightarrow U}\rangle$ as a lower bright state and $|S_{u \rightarrow G}\rangle$ as an upper dim state. As the bond distance increases, $|S_{u \rightarrow G}\rangle$ is less dim because the bond breaking CI opens vacancy in the $|G\rangle$ orbital in the ground state. Similarly $|S_{g \rightarrow U}\rangle$ decreases in brightness as $|U\rangle$ increases occupancy in the ground state. Simultaneously, these two excitations mix, concentrating their combined spectroscopic strength in the lower $|\text{local}\rangle$ state, with the transition dipoles interfering in the orthogonal complement $|\text{CT}\rangle$, such that it is dim, as we intuitively expect from a CT state.

These are then the best qualitative descriptions of the characters of the two most well-represented LR-UCCSD curves seen in Fig. 3. The lower asymptotic state $|\text{local}\rangle$ adiabatically connects to the intuitive MO-picture state $|S_{g \rightarrow U}\rangle$ at short distance, and this is the only bright state. This state traces a simple curve that is flat at the asymptote and rises steeply at short distances, and it has the lowest energy at all distances. The upper asymptotic $|\text{CT}\rangle$ state adiabatically connects to the multiply excited $|S_{u \rightarrow G}\rangle$ at short distance, and this state is both dim and higher in energy at all distances. Both of these states can be represented in an LR-UCCSD calculation

because even the multiply excited state at short distance needs only double excitations above the reference. Both of these states will also mix with states of other character as they rise in energy outside the window of interest near 1.4 Å, and this explains why the CT state appears to have a flat asymptote (because the remainder of the Coulombic part has been raised off of the top of the plot by such a mixing).

The forgoing has considered the situation in a minimal basis but with no restrictions as to what excitations might be present, commenting only at the end that the excitations that arose naturally in the discussion are available in LR-UCCSD. In a UADC calculation, however, the state $|S_{u \rightarrow G}\rangle$ is not present at short distances, due to its doubly excited character (relative to the reference). The triplet state $|T_{u \rightarrow U}\rangle$ is present though, and, due to broken spin symmetries, it and other triplets mix with the singlet states. Repeating the above algebra with the states available from one of the symmetry broken references, say $|\text{UHF}|\uparrow\rangle$, we obtain

$$\begin{aligned} |\text{local}|\uparrow\rangle &= 2^{-1/2} [\bar{a} \rightarrow \bar{A}^* + |b \rightarrow B^*] \\ &= 2^{-1/2} \{ [c|S_{g \rightarrow U}\rangle + d|T_{g \rightarrow G}\rangle] - [c|T_{u \rightarrow U}\rangle + d|S_{u \rightarrow G}\rangle] \} \\ |\text{CT}|\uparrow\rangle &= 2^{-1/2} [a \rightarrow B^* + |\bar{b} \rightarrow \bar{A}^*] \\ &= 2^{-1/2} \{ [c|S_{g \rightarrow U}\rangle + d|T_{g \rightarrow G}\rangle] + [c|T_{u \rightarrow U}\rangle + d|S_{u \rightarrow G}\rangle] \} \end{aligned} \quad (22)$$

As in the fully symmetric picture, at short bond length, the local and CT states mix. Accounting also for the fact that the orbitals are not polarized at short distances ($d = 0, c = 1$), we have

$$\begin{aligned} |S_{g \rightarrow U}\rangle &\propto |\text{local}|\uparrow\rangle + |\text{CT}|\uparrow\rangle \\ |T_{u \rightarrow U}\rangle &\propto |\text{local}|\uparrow\rangle - |\text{CT}|\uparrow\rangle \end{aligned} \quad (23)$$

Since the orbitals $|g\rangle$ and $|u\rangle$ are degenerate at this geometry, only the exchange contribution differentiates the energies of these two states, and, in this case, the lower state is the triplet; this is dark with respect to excitation from the ground state (which tends toward a spatially symmetric singlet at short bond distance).

As the atoms separate, however, the singlet $|S_{g \rightarrow U}\rangle$ and triplet $|T_{u \rightarrow U}\rangle$ excitations must mix, in order to isolate the $|\text{local}|\uparrow\rangle$ and $|\text{CT}|\uparrow\rangle$ states as they become very different in energy. Simultaneously components of the singlet $|S_{u \rightarrow G}\rangle$ and the triplet $|T_{g \rightarrow G}\rangle$ also begin to mix in as a result of orbital polarization. Meanwhile, the triplet components are no longer dark because the ground-state wavefunction is also spin contaminated. The transition dipoles of all these components nevertheless interfere constructively in the $|\text{local}|\uparrow\rangle$ state, which is bright, leaving the asymptotically higher $|\text{CT}|\uparrow\rangle$ state dim. This means the bright and the dim state must cross energetically, due to the spin contamination. This then explains the appearance of the curves in Fig. 4. Before the point of symmetry breaking, the two curves are best thought of as $|S_{g \rightarrow U}\rangle$ and $|T_{u \rightarrow U}\rangle$, which are both descending steeply in energy, with the dark triplet being lower. When the orbitals begin to polarize, singlets and triplets mix, the local excitations separate from the CT states, and the oscillator strength transfers from the upper $|S_{g \rightarrow U}\rangle$ state to the lower $|\text{local}|\uparrow\rangle$ state over an interval of roughly 0.5 Å.

One important thing that the above discussion shows us is that the lower two states in the UADC calculations are well represented as a superposition of the same two many-body basis states at all distances, these being $|\text{local}|\uparrow\rangle$ and $|\text{CT}|\uparrow\rangle$, which possess all of the oscillator strength inherent to the low-lying single-electron transitions. At each bond-distance limit, there is only one bright state. At short bond distances, the upper state is purely $|S_{g \rightarrow U}\rangle$ which has the same oscillator strength as the lowest-energy singlet in the proper, symmetric picture. At long distance, the lower state $|\text{local}|\uparrow\rangle$ has the same oscillator strength as $|\text{local}\rangle$ in the symmetric picture.

Therefore, by the argument forwarded in appendix A, the oscillator strength of the state of maximum brightness in this space that connects these two end points is given at all points by the sum of the oscillator strengths of these two states at each point. It further stands to reason that this approximates the oscillator strength of the lowest state in the fully symmetric picture, this being the only bright one.

B.3. Informal extrapolation to F_2^+ cation case

The situation with the cation excitations is in some ways much more complicated, but many features are wholly analogous.

In the completely symmetric picture (such that dissociation is to a superposition of $F \cdots F^+$ and $F^+ \cdots F$), the process can largely be thought of as containing the same σ -bond breaking as for neutral F_2 , but where a π -symmetry hole of mostly definite spin is entangled with the spin states of the two σ -symmetry electrons that are unpairing, in order to maintain that the cation (on whichever side it is located) is a triplet and that the neutral is a doublet, these being recoupled to an overall doublet of definite M_S . In fact, it is more complicated than that, since proper spin recoupling demands that neither the spin projection of the π hole at short distance nor of the atomic fragments asymptotically can be definite. This means also that the projection of the wavefunction in the σ bond-breaking space will also contain triplet components.

In spite of the formal increase in complexity, most features of interest may be qualitatively discussed in analogy to the neutral without needing to introduce the heavy notation necessary to discuss them formally. Asymptotically, excitation from the core to the σ valence space must contain four primitive excitations, two of which are local to the neutral or the cationic atoms, and two of which are CT-like in character. The local excitations are, as before, the asymptotically bright excitations, but these are no longer degenerate; the lower of the two is the excitation on the neutral. At short bond distance the only bright excitation to the σ valence space is from the $1\sigma_g$ orbital to the $2\sigma_u^{*(p)}$ orbital, which is also the lowest energy excitation. Therefore, the local excitation on the neutral atom (on whichever side it is located) will connect adiabatically to the $1\sigma_g \rightarrow 2\sigma_u^{*(p)}$ excitation as the lowest state at all distances, and this state will be bright. The higher state that asymptotes to the bright local excitation on the cation must therefore adiabatically connect to a state that is dim at short distance.

In light of the prior discussion of the UADC excitations of the neutral, the UADC curves for the excitations to the σ space of the cation in Fig. 6 have a relatively transparent structure. At short distance the lowest (only) two curves for excitations to the σ space are a bright doublet (the higher state), generated by the in-phase excitation of the two spins out of the $1\sigma_g$ orbital and a dark state generated by the out-of-phase spin excitations. The out-of-phase transition (lower state) is predominantly a quartet, but there is some contamination from a dark doublet with a similar exchange energy. At the point when the symmetry breaks ($\sim 1.2 \text{ \AA}$), the valence orbitals polarize only slowly, but the degenerate core orbitals localize immediately, due to the charge asymmetry. At this point excitations from each of the localized cores rapidly differentiate in energy, causing both the upper and lower state to bifurcate, even as they are only slightly spin contaminated (locality of excitation and spin purity are not strictly related, as demonstrated above explicitly for the neutral). As the distance increases, the valence orbitals continue to polarize and eventually localize letting each of the four excitations be well described as either local or CT states. Similar to what happened with neutral F_2 , this orbital polarization, and separation of local and CT character, must be accompanied by mixing of doublets and quartets, leaving all states heavily spin contaminated at the asymptote. In this mixing, it is the lower (after

bifurcation) doublet that mixes with the lower quartet as these both describe an excitation from the core that has localized to the atom that will eventually be neutral; the upper of these two states is asymptotically a dark CT state, but it is energetically flat because the result of the excitation is still a cation and a neutral; the lower of these two states is a bright local excitation on the neutral. When the upper doublet mixes with the upper quartet, the result is a bright local excitation on the atomic cation, which has the second highest asymptote overall (only the aforementioned local excitation on the neutral atom is lower), and a dark CT state which transfers an electron from the cation to the neutral, thus rising quickly in energy.

As with the neutral, the lowest-energy M_S -conserving core excitation (a doublet here) should be the sole bright state at short distance, but the asymptotically lowest-energy state that it should connect to is now one of two bright states. These two asymptotic bright local excitations are well represented in the UADC calculation, but, at short separation, they coalesce and erroneously connect to the dark predominantly-quartet state. The oscillator strength for the doublet state that it should connect to is, after the symmetry breaking point, split over two upper states (distinguished by the location of the core hole), which themselves erroneously connect to the two CT states. This is in full analogy to what we have seen with the neutral, except that there are two CT states, differentiated by the cation–neutral asymmetry. Therefore, it is reasonable to do as we did with the neutral and add the oscillator strengths from the lowest state that starts as a dark quartet and the bright upper two states that lose intensity as they transition to being CT states, such that this smoothly describes the intensity of the correct state that progresses from being the single bright MO-picture excitation to being the lower of the two bright local states asymptotically. The oscillator strength for the higher local excitation (on the cation) separates out of this sum naturally, as this state that is not included picks up its own intensity asymptotically.

The situation with the core $\rightarrow 2\pi_g^*$ transitions (arbitrary x, y orientation) is easier to understand, and these transitions are completely separate by symmetry from excitations to the valence σ space. In the properly symmetric picture, the only bright primitive excitations are $1\sigma_u^* \rightarrow 2\pi_g^*$ and $1\sigma_g \rightarrow 2\pi_u$. The $1\sigma_g \rightarrow 2\pi_u$ excitation is only relevant at longer distances when configurational mixings localize the π hole to whichever atom is the cation in a given half of the overall superposition. At short distances, only the $1\sigma_u^* \rightarrow 2\pi_g^*$ is available. At the point at which the two core states suddenly localize, the π -symmetry valence hole is still a good approximation of a delocalized MO, as suggested by Fig. 7 (at 1.4 \AA). It is then clear that the $1\sigma_u^* \rightarrow 2\pi_g^*$ excitation can be approximately constructed as a superposition of the transition from each local core orbital to this delocalized valence orbital, and so, by the arguments in appendix A, the oscillator strength of the $1\sigma_u^* \rightarrow 2\pi_g^*$ transition is simply the sum of the oscillator strengths of these two UADC states. In the properly symmetric picture, this $1\sigma_u^* \rightarrow 2\pi_g^*$ transition adiabatically connects to an in-phase superposition of local core $\rightarrow 2p$ excitations on the cation, and the lower relevant UADC asymptote is indeed a faithful representation of this transition (only where the side on which the cation is located on is definite), so that adding to its oscillator strength the zero intensity from the other curve continues to provide the correct oscillator strength. The higher state represents the dark, neutral–cation CT state asymptotically.

It is worth noting for completeness of discussion the double degeneracy of the core $\rightarrow 2p$ excitations on cation asymptotically (at both the UADC and LR-UCCSD levels), where one of the excitations is to the σ -symmetry space and one is to the π -symmetry space.

References

- [1] M. Lewenstein, P. Balcou, M.Y. Ivanov, A. L'Huillier, P.B. Corkum, *Phys. Rev. A* 49 (1994) 2117.
- [2] R. Kienberger, E. Goulielmakis, M. Uiberacker, A. Baltuska, V. Yakovlev, F. Bammer, A. Scrinzi, T. Westerwalbesloh, U. Kleineberg, U. Heinzmann, M. Drescher, F. Krausz, *Nature* 427 (2004) 817.
- [3] T. Popmintchev, M.-C. Chen, P. Arpin, M.M. Murnane, H.C. Kapteyn, *Nat. Photonics* 4 (2010) 822.
- [4] Z.-H. Loh, S.R. Leone, *J. Chem. Phys.* 128 (2008) 204302.
- [5] E. Goulielmakis, Z.-H. Loh, A. Wirth, R. Santra, N. Rohringer, V.S. Yakovlev, S. Zherebtsov, T. Pfeifer, A.M. Azzeer, M.F. Kling, S.R. Leone, F. Krausz, *Nature* 466 (2010) 739.
- [6] E.R. Hosler, S.R. Leone, *Phys. Rev. A* 88 (2013) 023420.
- [7] J.H. Posthumus, J. Plumridge, P.F. Taday, J.H. Sanderson, A.J. Langley, K. Codling, W.A. Bryan, *J. Phys. B: At. Mol. Opt. Phys.* 32 (1999) L93.
- [8] E. Goll, G. Wunner, A. Saenz, *Phys. Rev. Lett.* 97 (2006) 103003.
- [9] T. Ergler, B. Feuerstein, A. Rudenko, K. Zrost, C.D. Schröter, R. Moshammer, J. Ullrich, *Phys. Rev. Lett.* 97 (2006) 103004.
- [10] L. Fang, G.N. Gibson, *Phys. Rev. A* 75 (2007) 063410.
- [11] L. Fang, G.N. Gibson, *Phys. Rev. Lett.* 100 (2008) 103003.
- [12] L. Fang, G.N. Gibson, *Phys. Rev. A* 78 (2008) 051402(R).
- [13] M.S. Gordon, D.G. Truhlar, *Theor. Chim. Acta* 71 (1987) 1.
- [14] J.D. Watts, R.J. Bartlett, *J. Chem. Phys.* 95 (1991) 6652.
- [15] J. Gräfenstein, D. Cremer, *Phys. Chem. Chem. Phys.* 2 (2000) 2091.
- [16] Y. Shao, M. Head-Gordon, A.I. Krylov, *J. Chem. Phys.* 118 (2003) 4807.
- [17] Y. Shao, Z. Gan, E. Epifanovsky, A.T.B. Gilbert, M. Wormit, J. Kussmann, A.W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P.R. Horn, L.D. Jacobson, I. Kaliman, R.Z. Khaliullin, T. Kus, A. Landau, J. Liu, E.I. Proynov, Y.M. Rhee, R.M. Richard, M.A. Rohrdanz, R.P. Steele, E.J. Sundstrom, H.L. Woodcock III, P.M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G.J.O. Beran, Y.A. Bernard, E. Berquist, K. Brandhorst, K.B. Bravaya, S.T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S.H. Chien, K.D. Closser, D.L. Crittenden, M. Diedenhofen, R.A. DiStasio Jr., H. Do, A.D. Dutoi, R.G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M.W.D. Hanson-Heine, P.H.P. Harbach, A. W. Hauser, E.G. Hohenstein, Z.C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyayev, J. Kim, J. Kim, R.A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K.U. Lao, A. Laurent, K.V. Lawler, S.V. Levchenko, C.Y. Lin, F. Liu, E. Livshits, R.C. Lochan, A. Luenser, P. Manohar, S.F. Manzer, S.-P. Mao, N. Mardirossian, A.V. Marenich, S.A. Maurer, N.J. Mayhall, E. Neuscamman, C.M. Oana, R. Olivares-Amaya, D.P. O'Neill, J.A. Parkhill, T.M. Perrine, R. Peverati, A. Prociuk, D.R. Rehn, E. Rosta, N.J. Russ, S.M. Sharada, S. Sharma, D.W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A.J.W. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M.A. Watson, J. Wenzel, A. White, C.F. Williams, J. Yang, S. Yeganeh, S.R. Yost, Z.-Q. You, I.Y. Zhang, X. Zhang, Y. Zhao, B.R. Brooks, G.K.L. Chan, D.M. Chipman, C.J. Cramer, W.A. Goddard III, M.S. Gordon, W.J. Hehre, A. Klamt, H.F. Schaefer III, M.W. Schmidt, C.D. Sherrill, D.G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A.T. Bell, N.A. Besley, J.-D. Chai, A. Dreuw, B.D. Dunietz, T.R. Furlani, S.R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D.S. Lambrecht, W. Liang, C. Ochsenfeld, V.A. Rassolov, L.V. Slipchenko, J.E. Subotnik, T. Van Voorhis, J.M. Herbert, A.I. Krylov, P.M.W. Gill, M. Head-Gordon, *Mol. Phys.* 113 (2015) 184.
- [18] O. Živný, J. Czernek, *Chem. Phys. Lett.* 308 (1999) 165.
- [19] R.D. Cohen, C.D. Sherrill, *J. Chem. Phys.* 114 (2001) 8257.
- [20] G.J.O. Beran, S.R. Gwaltney, M. Head-Gordon, *Phys. Chem. Chem. Phys.* 5 (2003) 2488.
- [21] P.R. Tentscher, J.S. Arey, *J. Chem. Theory Comput.* 8 (2012) 2165.
- [22] A.W. Potts, W.C. Price, *Trans. Faraday Soc.* 67 (1971) 1242.
- [23] R.P. Tuckett, A.R. Dale, D.M. Jaffey, P.S. Jarrett, T. Kelly, *Mol. Phys.* 49 (1983) 475.
- [24] A.J. Cormack, A.J. Yench, R.J. Donovan, K.P. Lawley, A. Hopkirk, G.C. King, *Chem. Phys.* 213 (1996) 439.
- [25] J. Yang, Y. Hao, J. Li, C. Zhou, Y. Mo, *J. Chem. Phys.* 122 (2005) 134308.
- [26] K.P. Huber, G. Herzberg, *Molecular Spectra and Molecular Structure: Constants of Diatomic Molecules*, Van Nostrand-Reinhold, New York, 1979.
- [27] A. Kramida, Y. Ralchenko, J. Reader, and NIST ASD Team, *NIST Atomic Spectra Database (ver. 5.3)*, [Online]. Available: <http://physics.nist.gov/asd> [2015, November 30]. National Institute of Standards and Technology, Gaithersburg, MD, 2015.
- [28] N. Moiseyev, *Non-Hermitian Quantum Mechanics*, Cambridge University Press, New York, 2011.
- [29] J. Schirmer, *Phys. Rev. A* 26 (1982) 2395.
- [30] A.B. Trofimov, J. Schirmer, *J. Phys. B* 28 (1995) 2299.
- [31] A. Barth, J. Schirmer, *J. Phys. B: At. Mol. Opt. Phys.* 18 (1985) 867.
- [32] E.M.-L. Ohrendorf, L.S. Cederbaum, F. Tarantelli, *Phys. Rev. A* 44 (1991) 205.
- [33] J. Schirmer, A.B. Trofimov, K.J. Randall, J. Feldhaus, A.M. Bradshaw, Y. Ma, C.T. Chen, F. Sette, *Phys. Rev. A* 47 (1993) 1136.
- [34] G. Handke, F. Tarantelli, A. Tarantelli, L.S. Cederbaum, *J. Electron. Spectrosc. Relat. Phenom.* 75 (1995) 109.
- [35] H.D. Schulte, L.S. Cederbaum, *J. Chem. Phys.* 103 (1995) 698.
- [36] H. Köppel, F.X. Gadea, G. Klatt, J. Schirmer, L.S. Cederbaum, *J. Chem. Phys.* 106 (1997) 4415.
- [37] L.S. Cederbaum, J. Zobeley, F. Tarantelli, *Phys. Rev. Lett.* 79 (1997) 4778.
- [38] M.S. Deleuze, L.S. Cederbaum, *Int. J. Quantum Chem.* 63 (1997) 465.
- [39] J. Zobeley, L.S. Cederbaum, F. Tarantelli, *J. Chem. Phys.* 108 (1998) 9737.
- [40] A.B. Trofimov, T.E. Moskovskaya, E.V. Gromov, H. Köppel, J. Schirmer, *Phys. Rev. A* 64 (2001) 022504.
- [41] N. Moiseyev, R. Santra, J. Zobeley, L.S. Cederbaum, *J. Chem. Phys.* 114 (2001) 7351.
- [42] H. Hennig, J. Breidbach, L.S. Cederbaum, *J. Phys. Chem. A* 109 (2005) 409.
- [43] A.I. Kuleff, L.S. Cederbaum, *Chem. Phys.* 338 (2007) 320.
- [44] A.I. Kuleff, L.S. Cederbaum, *Phys. Rev. Lett.* 98 (2007) 083201.
- [45] I. Bâldea, B. Schimmelpfennig, M. Plaschke, J. Rothe, J. Schirmer, A.B. Trofimov, T. Fanghanel, *J. Electron. Spectrosc. Relat. Phenom.* 154 (2007) 109.
- [46] R. Santra, N.V. Kryzhevoi, L.S. Cederbaum, *Phys. Rev. Lett.* 103 (2009) 013002.
- [47] V. Feyrer, O. Plekan, R. Richter, M. Coreno, M. de Simone, K.C. Prince, A.B. Trofimov, I.L. Zaytseva, J. Schirmer, *J. Phys. Chem.* 114 (2010) 10270.
- [48] A.D. Dutoi, L.S. Cederbaum, *Phys. Rev. A* 90 (2014) 023414.
- [49] J. Wenzel, M. Wormit, A. Dreuw, *J. Comput. Chem.* 35 (2014) 1900.
- [50] J. Wenzel, M. Wormit, A. Dreuw, *J. Chem. Theory Comput.* 10 (2014) 4583.
- [51] K. Gokhberg, P. Kolorenč, A.I. Kuleff, L.S. Cederbaum, *Nature* 505 (2014) 661.
- [52] E. Fasshauer, P. Kolorenč, M. Pernpointner, *J. Chem. Phys.* 142 (2015) 144106.
- [53] J. Wenzel, A. Dreuw, *J. Chem. Theory Comput.* 12 (2016) 1314.
- [54] V. Stumpf, K. Gokhberg, L.S. Cederbaum, *Nat. Chem.* 8 (2016) 237.
- [55] C.C. Paige, *IMA, J. Appl. Math.* 10 (1972) 373.
- [56] C. Leforestier, R.H. Bisseling, C. Cerjan, M.D. Feit, R. Friesner, A. Guldberg, A. Hammerich, G. Jolicard, W. Karrlein, H.-D. Meyer, N. Lipkin, O. Roncero, R. Kosloff, *J. Comput. Phys.* 94 (1991) 59.