Statistical analysis of correlations in the x-ray induced Coulomb explosion of iodopyridine

Benoît Richard[1,2,3], Julia M. Schäfer[1,4], Zoltan Jurek[1,2], Robin Santra[1,2,3,4], Ludger Inhester[1,2]

[1] Center for Free-Electron Laser Science CFEL, Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany

[2]The Hamburg Centre for Ultrafast Imaging, Luruper Chaussee 149, 22761 Hamburg, Germany

[3]Department of Physics, Universität Hamburg, Notkestr. 9-11, 22607 Hamburg, Germany

[4]Department of Chemistry, Universität Hamburg, Martin-Luther-King-Platz 6, 20146 Hamburg, Germany

E-mail: benoit.richard@desy.de,ludger.inhester@desy.de

Abstract. Coulomb explosion imaging is a promising experimental tool to study individual molecules. In this work simulation data for the x-ray induced Coulomb explosion of 2-iodopyridine is analyzed and the involved fragmentation dynamics are described. It is found that particular final ion momenta show correlations that reflect a collision of two atoms during the explosion. Variations of the proton kinetic energies can be associated with variations in the charge build-up time. Covariances of the forces during the explosion can be utilized to simplify the description of the dynamics in reduced dimensionality using only four collective coordinates.

Submitted to: *J. Phys. B: At. Mol. Opt. Phys.*

1. Introduction

A powerful tool for gaining knowledge on the structure and dynamics of molecules is the Coulomb explosion imaging (CEI) technique [1]. The basic principle behind CEI is that a molecule is very rapidly charged up such that the strong Coulomb repulsion among the atomic ions overcomes any chemical binding energy. If charge buildup is realized faster than the timescales of structural rearrangement and the charge is high enough, the molecule dissociates rapidly into atomic ions mainly driven by their Coulomb interaction. In such a scenario, the final ion momenta contain detailed information on the undistorted initial geometry of the molecule before the initiation of the explosion.

Originally, the rapid buildup of charge in the molecule has been achieved by sending molecular ion beams through a thin metal foil [2]. In the collision of the molecular ions with the foil, electrons are stripped away on a sub-femtosecond timescale, resulting in a quasi-instantaneous explosion. In order to conduct CEI also with neutral molecules, techniques have been developed to ionize molecules through collision with highly charged ions [3, 4] and with intense femtosecond optical laser pulses [5–37]. These approaches also allow embedding CEI in a pump-probe setup to also image photo-induced structural changes [7].

A key limitation of CEI induced by an optical laser is that it requires enormous laser intensity to achieve a sufficiently high degree of ionization [20, 37, 38]. Ionization by strong electric field leads to polarization dependencies for the emission direction of ions [8–10]. Furthermore, ionization rates may sensitively depend on the molecular geometry [11]. These effects complicate an interpretation of ion momenta in terms of initial geometries considerably.

The advent of x-ray free-electron lasers (XFEL) [39–44] offers new opportunities for CEI. Today, x-ray pulses at XFELs can be created with an intensity leading to the absorption of multiple x-ray photons [45] and a pulse duration of few femtoseconds or even below [46]. Since x-ray light predominantly ionizes inner-shell electrons, each absorbed photon implies a cascade of Auger decay relaxation processes that further ionize the molecule, resulting in a very efficient and rapid creation of charge [45, 47, 48]. In contrast to ionization via intense optical light, ionization via x-ray light can be considered almost insensitive to molecular geometry or orientation with respect to polarization. So far, x-ray induced Coulomb explosion has been employed in a number of experiments involving mostly small molecules [49]. Specifically, time-resolved studies investigated dissociation [50], isomerization dynamics [51, 52], and charge transfer [53] dynamics.

If molecules with multiple atoms are considered, the resulting interaction among the ions during the Coulomb explosion can lead to non-trivial collision effects [16] and rotation dynamics that may lead to misinterpretation of the final momentum data [54]. For the correct interpretation of ion momenta, a crucial ingredient is the timescale for the charge buildup in the molecule that is not only governed by the x-ray pulse duration, but also involves Auger decay cascades building up charge in the molecule

on an intrinsic timescale of a few femtoseconds. Thus, in addition to the temporal fluctuations due to the finite pulse duration, the spatio-temporal fluctuations associated with relaxation cascades also lead to variations in the ionization dynamics that are mapped into the distribution of the final ion momenta. Here, simulations of the x-ray induced fragmentation of molecules play a crucial role in understanding how the charge in the molecule builds up and the individual ions emerge. In particular, they give insight into the correlation of ion momenta and uncover how these correlations relate to dynamics during the explosion process, thereby providing important insights for accurate interpretations of coincidently detected ion momenta. Considering the possibly large number of emitted ions from a single molecule and given the variety of different ion charge states, it is often a non-trivial task to read the essential information that can be gained from such a multi-dimensional ion simulation data set.

In order to facilitate the interpretation of Coulomb explosion data, we conduct a statistical analysis of simulation data for the x-ray induced Coulomb explosion of the 2-iodopyridine molecule. This simulation data set has been created in the context of a recent experiment [55] at the European X-ray Free-Electron Laser facility. We show how relevant dynamical patterns occurring during the explosion can be uncovered through principal component analysis (PCA) of the final ion momentum data and the forces. We demonstrate that individual ion pairs exhibit particularly strong correlated momenta due to their collision during the molecular explosion. We further show how the explosion dynamics can be decomposed into few collective coordinates that we obtain from PCA components of the forces. Furthermore, we show that particular information on the temporal evolution of the charging up of individual parts of the molecule is encoded in the ion momentum distribution.

The paper is structured as follows: In section 2 we briefly introduce our methods for performing simulations of x-ray induced Coulomb explosion and summarize the employed analysis techniques. In section 3 we present our results and in section 4 we draw final conclusions.

2. Methods

2.1. Simulation of XFEL induced Coulomb Explosion

We perform the simulation using the XMDYN toolkit [56], which is designed to model matter irradiated by high-intensity x-rays, such as XFEL pulses. The approach consists of a quantum-classical hybrid. On one hand, atoms and ionized (free) electrons move as classical particles in the Coulomb field generated by atomic ions and free electrons. For this molecular dynamics (MD) simulation, particles are propagated using the velocity Verlet algorithm with 1-attosecond time steps for a total time of up to 1 ps. All trajectories are initialized with the same neutral ground-state geometry of the 2-iodopyridine molecule at zero initial velocities. The equilibrium molecular geometry is computed using GAMESS [57] at the restricted Hartree-Fock level of theory with a 3-

Parameter	Value
Photon energy	2 keV
Pulse shape	Gaussian
Full-width at half-maximum	10 fs
Fluence	7.5×10^{10} photons/μm^2
Polarization	Linear, perpendicular to the molecular plane

Table 1: Pulse parameters.

21G* basis set. On the other hand, bound electrons and their interaction with light are treated quantum mechanically. Electronic configurations are considered individually for each atom, via the occupation numbers of the atomic orbitals, which also means that no molecular bonds or molecular orbitals are taken into account. However, atoms and atomic ions can exchange electrons based on a classical over-the-barrier model that leads to the change of occupation numbers involved in the process and to the change of atomic charges. An electron is transferred between two atoms if it has an orbital energy higher than the potential barrier separating the two atoms. The height of this barrier is determined by the maximum of the classical electrostatic potential an electron sees along the direct line connecting the two atoms [53, 58]. Ab-initio calculations using the XATOM toolkit [56] are employed to obtain orbital binding energies, photoionization cross sections, and rates for Auger and fluorescence decays, taking into account all atomic orbitals which are occupied in the neutral ground state. These quantities are then used by XMDYN in a Monte-Carlo algorithm to incorporate the corresponding events into the MD simulations according to their computed probability and to track the occupation of the atomic orbitals. Emitted electrons are turned into classical free electrons and are placed around the emitting atom according to their angular emission probability. In each electron emission process, the electron-emitting ion is given a corresponding recoil momentum to satisfy momentum conservation.

The pulse parameter used in the simulation are listed in table 1.

2.2. Principal component analysis for non-equilibrium process

Principal component analysis (PCA) is a standard analysis technique for high dimensional data [59]. It decomposes its input space into multiple uncorrelated components by diagonalizing the covariance matrix. Given the input data represented as vector $\boldsymbol{x}$ with N components, the covariance matrix elements $\boldsymbol{\Sigma}_{ij}$ are given by the pairwise covariances, that is $\boldsymbol{\Sigma}_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle$ where $\langle . \rangle$ reprsents the ensemble average. PCA decomposes the covariance matrix $\boldsymbol{\Sigma}$ as

$$\boldsymbol{\Sigma} = \mathbf{U}\,\boldsymbol{\Lambda}\,\mathbf{U}^T \tag{1}$$

where the columns of matrix $\mathbf{U}$ are the principal component vectors $\boldsymbol{u}_i$ and the diagonal matrix $\boldsymbol{\Lambda}$ contains the eigenvalues λ_i that are the corresponding explained variances [60].

The principal components $\boldsymbol{u}_i$ (i.e., the eigenvectors of $\boldsymbol{\Sigma}$) are ordered according to λ_i in decreasing order. Projecting variations of a data set into its first principal components thus provides a description of the data with maximal variance for a given reduced dimensionality. The study of the first few components is therefore often sufficient to capture most of the variations. Moreover, thanks to $\boldsymbol{\Lambda}$ being diagonal, projecting the original data on the PCA components results in uncorrelated quantities. The algorithm can thus be viewed as linearly decorrelating the input data. PCA is a useful tool to analyze MD simulations, for both steady state [59, 61, 62] and non-equilibrium [63–65] dynamics. In our analysis, we do not perform any rescaling of the input data (i.e., giving all input components the same variance) and conduct the PCA analysis on the plain covariance matrix.

In this work, we apply PCA to two different inputs: the final momenta and the dynamical forces. Given a set of N atoms, we first perform PCA on the $3N$-dimensional set of final momenta of all atoms. This data can in principle be measured in coincident CEI experiments [55]. Therefore, final momentum PCA may realistically be performed on experimental data in the near future, by leveraging PCA methods working with missing data [66] to compensate for the finite detection efficiency. Understanding how it can be interpreted and used to elucidate the dynamics of the Coulomb explosion is thus of prime interest.

Moreover, we apply PCA to the dynamical data, taking into account the full time-dependent data from the explosion simulation. This way, we aim to describe not only the variation of the ensemble in the final state of the explosion (i.e., the final momenta), but also cover temporal variations during the explosion dynamics. A common choice for analysis of MD trajectories is to use atom positions as input. However, in the current context, the topology of the positional data is problematic due to the far-from-equilibrium nature of Coulomb explosions. The atoms are evolving towards constant velocities, and thus, depending on the length of the simulation, the resulting components may be dominated by the asymptotic behavior of the system. This suppresses information about the dynamics during the ionization and explosion and defeats the purpose of using dynamical data. As a more natural alternative, we employ forces as input. Forces are bounded, and they are zero at the start and end of the simulation. They are also directly proportional to the charge of each atom, allowing to emphasize the effect of the charge distribution on the explosion. From the simulation data, we compute the forces on the individual atoms by second order finite differences from the position data employing a femtosecond time step.

3. Results

3.1. Description of the simulation data

The simulation data consists of 15,000 trajectories describing the time-dependent position and charge of all atoms. At the final time-frame, the trajectories show a

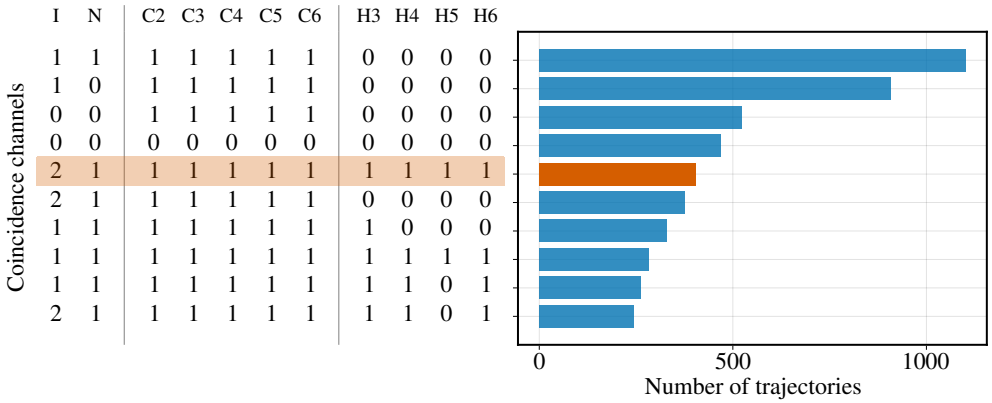

Figure 1: Number of occurrence of the most abundant final charge distributions obtained in the simulation. The numbers are the final charges of each atoms. The coincidence condition used for the analysis is highlighted in orange.

variety of different charge states for a given atom. The simulations result in 1739 distinct combinations of final charges. The trajectory counts for the most abundant ones are depicted in figure 1. If not stated otherwise, we restrict our analysis to a single coincidence condition, i.e., we select trajectories where the atoms have the final charge state highlighted in figure 1. As can be seen, this set of coincident ion charges is the most abundant in which all atoms have a charge of at least $+1$, which is a prerequisite for their experimental detection. This coincidence channel contains 404 trajectories. For simplicity, our analysis focuses on this subset of trajectories, because the data for other final charge coincidence conditions exhibit *quantitative* differences that undermine a joint analysis. We note, however, that most of the other coincident conditions are associated with *qualitatively* similar dynamics and analysis of the other sets of coincident ion charges yields analogous results.

We start by describing general properties of the data. Iodopyridine is a planar molecule and in the following the molecular plane is defined by the xy plane. The equilibrium geometry of the neutral molecule is represented in figure 2(b) together with the atom labels. The ionization cross section for ionization on iodine is significantly larger compared to ionization on the other atoms such that the charge is primarily created on the iodine through photoionization and subsequent Auger decays. The created charge is then distributed among the atoms by charge rearrangement. After the molecule has been exposed to the x-ray pulse it rapidly dissociates into individual atomic ions. Snapshots of the position distribution at selected times after the pulse peak are shown in figure 2(a). After $\sim 100\,\mathrm{fs}$ the momenta of all atomic ions converge to their final asymptotic values. The distribution of these final momenta is shown in figure 2(c). As can be seen, the final ion momentum distribution of each atom is relatively sharp and most of the momenta can be directly linked to the geometrical position of the atom in the initial geometry. Compared to the other heavier atoms, the C2 momentum stands out by pointing to the left, whereas in the initial geometry the C2 atom is next to the

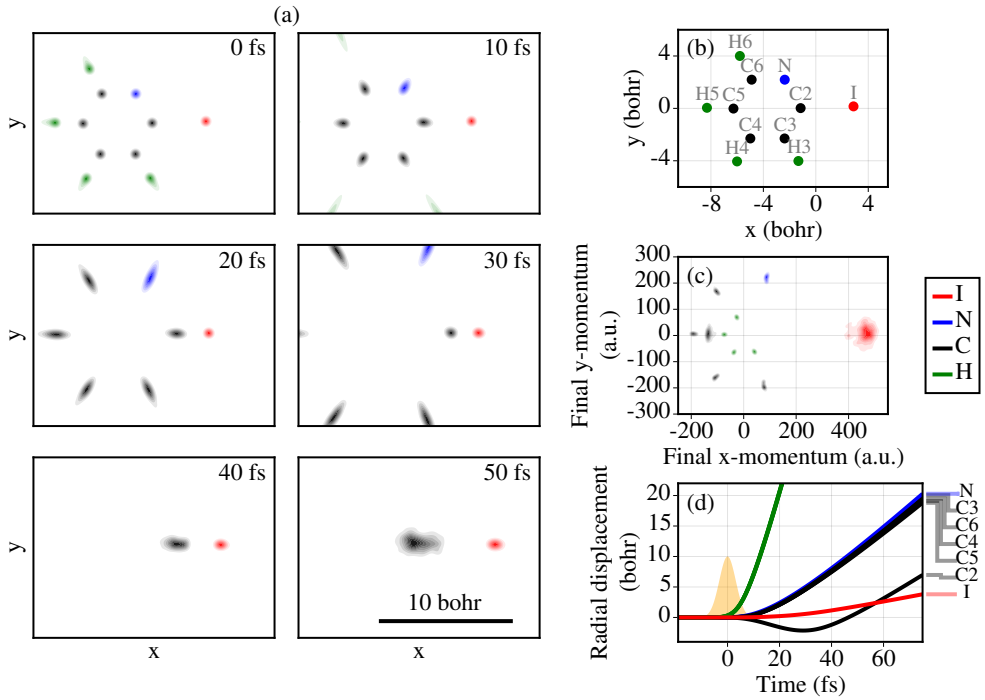

Figure 2: (a) Distribution of the ion positions at selected times. (b) 2-Iodopyridine equilibrium geometry used as initial condition for the simulations. (c) Final momentum distribution. (d) Radial displacements. The yellow area indicates the temporal shape of the x-ray pulse.

iodine atom on the right-hand side of the molecule. The C2 momentum also displays somewhat higher variance compared to other carbon atoms.

As indicated by the snapshots in figure 2(a), the explosion is mainly radial. In figure 2(d) we show the radial displacement of all atoms as a function of time, with the radial direction defined by the final average momentum of each atom. As can be seen, already during the pulse exposure, the hydrogens are expelled. This is followed by the radial dynamics of the atoms N, C3, C4, C5, C6, and, eventually, the I atom. An exception to this purely radial explosion is the C2 atom that undergoes a reversion of its motion during the explosion. As depicted in figures 2(a) and (d) this reversion occurs around 30 fs.

3.2. Out-of-plane dynamics

We note that while the molecule is initially planar, a substantial momentum spread appears along the out-of-plane direction for the I and C2 atoms. Because in the simulation the initial conditions do not contain any out-of-plane distortion, this motion must be attributed to the recoil momentum caused by emission of electrons from the iodine atom. Because ionization occurs predominantly on the iodine atom, the momentum perpendicular to the molecular plane is originally only on the I atom. It

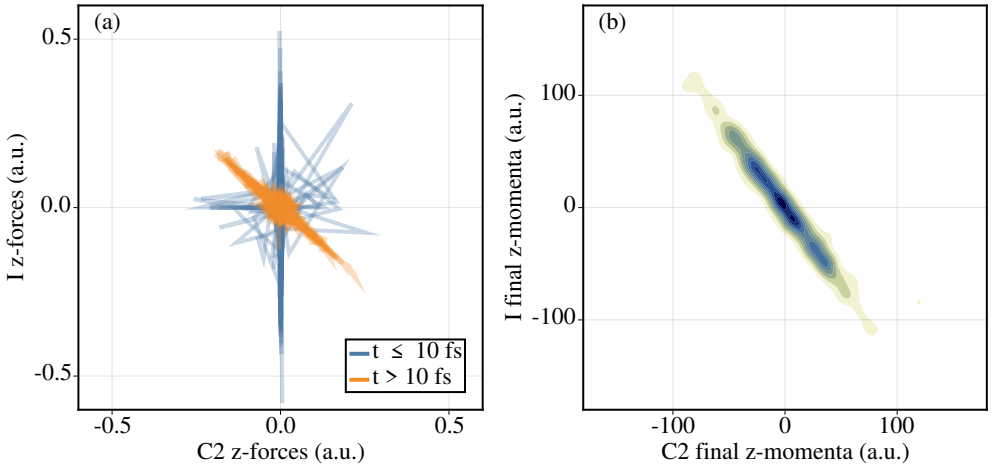

Figure 3: (a) Out-of-plane forces for all trajectories and all times for the C2 and I atoms. Each point on a trajectory is colored according to time. (b) Distribution of the final momentum of the I atom along the z axis against the final momentum of the C2 atom along the z axis.

is then transferred to other atoms and further enhanced by the repulsion between the atomic ions, especially for I and C2 that come close to each other. These two processes can be associated with different time periods during the simulation. This is illustrated in figure 3(a) which shows the out-of-plane forces for I and C2. Two distinct patterns can be recognized in the figure by coloring the forces according to time. For early time $< 10\,\mathrm{fs}$ (blue lines), i.e., during or shortly after the x-ray pulse, the out-of-plane forces dominantly appear for I and are considerably lower for the C2 atom. For later times (orange lines), i.e., after the x-ray pulse, the out-of-plane forces for the two atoms show a strong anti-correlation as a consequence of their mutual repulsion. This out-of-plane momentum is particularly strong for I and C2 and the involved collision dynamics eventually result in a strong anti-correlation in the final momenta that is shown in figure 3(b). This strong anti-correlation in the final momenta can thus be seen as a strong fingerprint of the collision between the two atoms, in which the two atoms are initially pushed against each other by the Coulomb-forces due to the surrounding ions and later repel each other. The observation of out-of-plane momenta that are amplified due to the collision between the two colliding atoms represents a remarkable manifestation of the photoelectron-recoil effect [67], whose effects have been studied before via photoelectron and Auger spectra [68–80] .

3.3. PCA of final momenta

To further elucidate how the explosion is mapped into correlations in the final momentum data, we now inspect the principal components of the final momenta. The first four components are illustrated in panels (a)-(d) of figure 4 showing projections

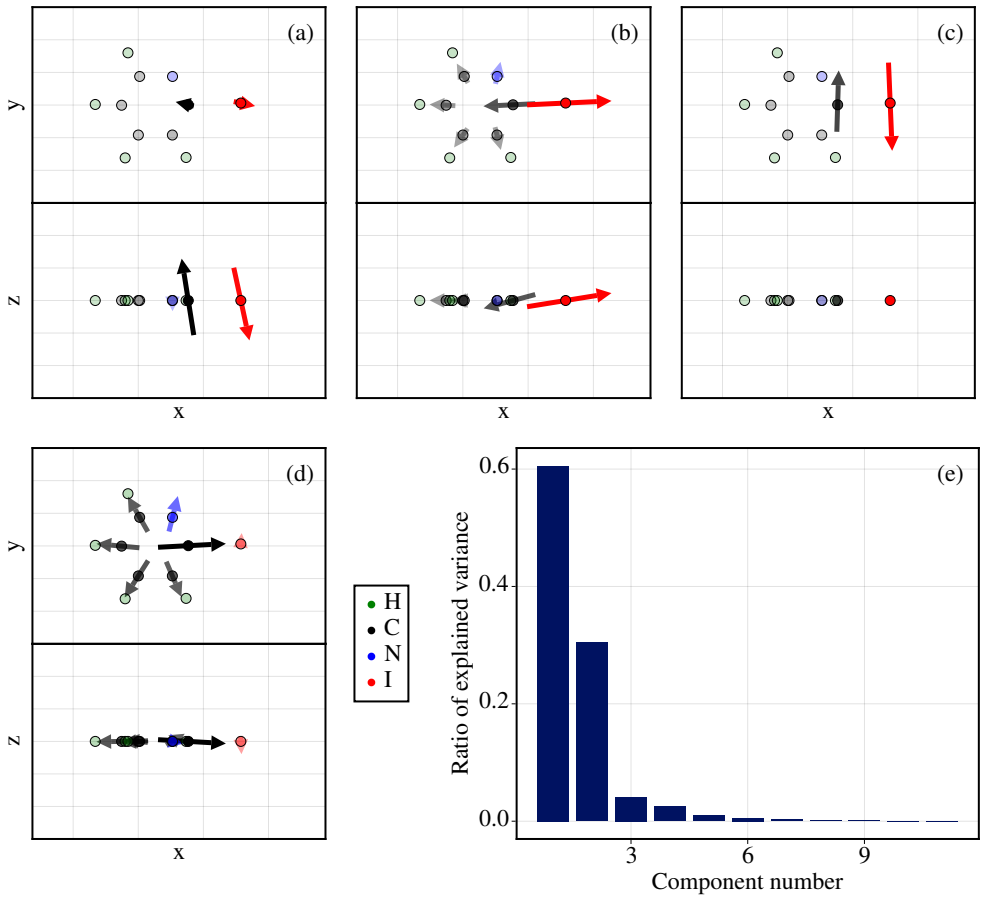

Figure 4: Results of final momentum PCA. (a)-(d): First four PCA components, superimposed on the equilibrium atom positions and ordered by component number. Transparency of the color is set relative to the magnitude of the component at the given atom. Arrows smaller than the atom marker are omitted. (e) Ratio of explained variance of the PCA components.

in the xy and xz planes. The components are depicted by arrows on each atom of the iodopyridine molecule at its ground-state equilibrium geometry. The relative length of the arrows and their direction indicates the respective amplitude of each atom for the component. We choose to orient the components such that the projection of the final velocities on the component is on average positive.

In figure 4(e) we report the ratio of explained momentum variance for the first PCA components. In the following, we discuss the components individually.

The first component describes an anti-correlated out-of-plane movement of the C2 and I atom that has already been discussed in figure 3. As can be seen, the variance explained by this component is quite significant compared to the total variance in the final momentum data. The predominant occurrence of this component points to the strong interactions between the two atoms arising from their collision described above. The strong interaction between the C2 and I atoms is further emphasized by the two next

components. Component 2 reflects anti-correlated variance of the C2 and I atom along the x-axis, i.e., along the direction of the I-C2 collision. Component 3 corresponds to variations of the I and C2 momenta in the xy plane transverse to the collision direction. Component 4 corresponds to the expulsion of the carbon atoms. A notable feature absent from the first four components is the momenta of the hydrogen atoms.

The predominant occurrence of correlated I-C2 momenta in the dominant components indicates a strong impact of the collision dynamics of the two atoms on the final momentum variations. This result suggests that studying covariances of final ion momenta (e.g., via PCA decomposition) gives a clear indication on collisions during the explosion.

3.4. PCA of forces

We now turn to the study of the PCA used on forces. We applied the same procedure as for the final momentum PCA, but now also take into account the dynamical information at each femtosecond time step. Specifically, we consider forces for the time frames between 20 fs before and 150 fs after the pulse center and for all trajectories fulfilling the coincidence condition highlighted in figure 1. Each snapshot is considered equally when constructing the covariance matrix. Thus the variance captures not only the ensemble variations but also the changes in the forces in time during each trajectory. The results are reported in figure 5 in a similar fashion as in figure 4. As can be seen, the first component corresponds to the radial expulsion of the iodine, nitrogen, and all the carbon atoms, but C2. The next component is very similar to final momentum component 2, as it mostly represents the C2-I collision. Component 3 describes the ejection of the hydrogen atoms, coupled with a slight push of the carbon atoms toward the center. This anti-correlated force between the hydrogens and the carbons can be associated with the recoil momentum that the hydrogens imprint on the carbon atoms when they are radially expelled out of the molecule. Accordingly, the corresponding inward contribution for nitrogen atom that has no hydrogen attached to it is absent. Finally, component 4 depicts the out-of-plane dynamics for C2 and I that have been discussed already in figure 3 and are rooted in the electron emission recoil and the C2-I collision.

The first four components obtained by the force PCA exhibit a quite complete picture of the dynamics during the Coulomb explosion. They separate the observed dynamics into remarkably insightful contributions.

3.5. Decomposition of the dynamics

To underline the advantage of the principal components obtained by force PCA in decomposing the explosion dynamics, we demonstrate to what degree the overall dynamics can be reconstructed using only the four components presented in figure 5. In figure 6 we describe the overall dynamics employing projections of time-dependent

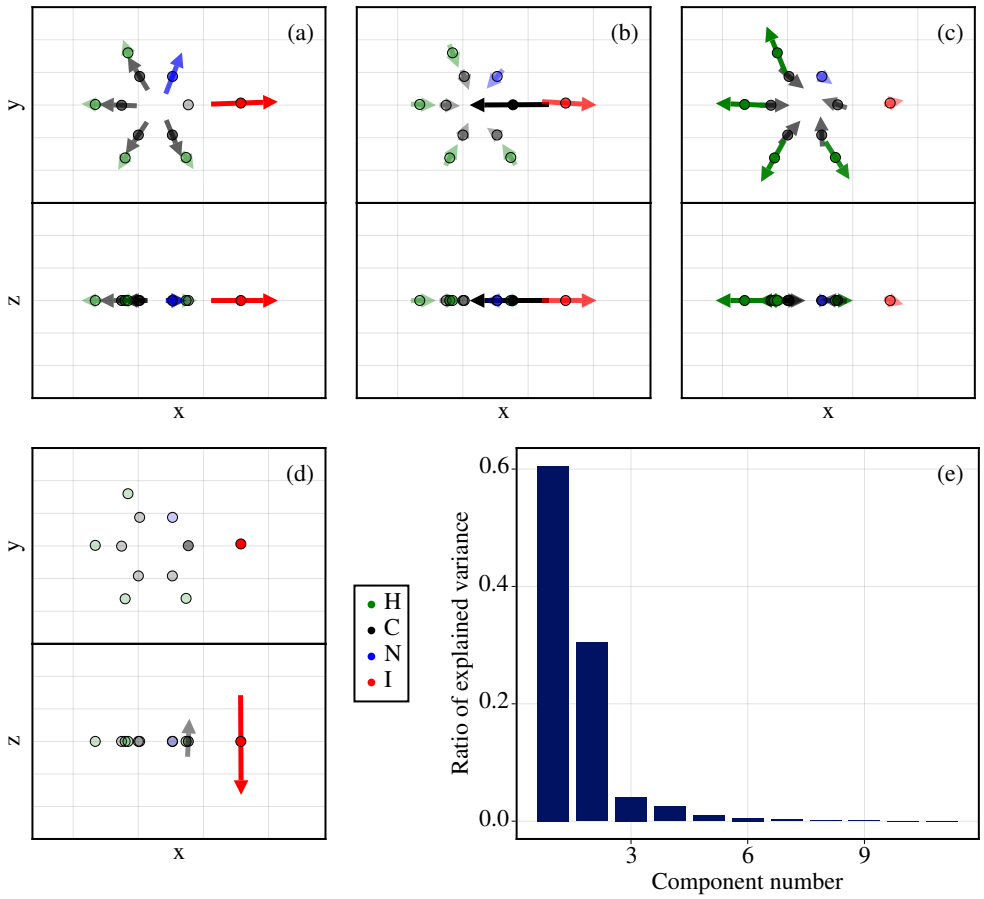

Figure 5: Results of force PCA. (a)-(d) First four PCA components, superimposed on the equilibrium geometry and ordered by component number. Transparency of the color is set relative to the size of the component at the given atom. Arrows smaller than the atom marker are omitted. (e) Ratio of explained variance of the PCA components.

quantities onto the first four components. The left panel shows average projected displacement, velocities, and forces.

Figure 6(a) shows the displacements projected onto the components. We see that displacement along the first 3 components increases steadily. Displacement along component 3 increases more rapidly compared to component 1 and 2. Component 4 only shows a marginal increase. The corresponding velocities are shown in figure 6. As can be seen, they increase up to a certain threshold. For component 2 we see an intermediate maximum followed by a slight decrease until 30 fs. The velocities along component 4 remain very low. The forces projected along the components are shown in figure 6(c). Whereas the force projected along component 1 rapidly increases at the onset of the explosion and decreases again afterwards, component 2 and 3 change their sign during the explosion.

In accordance to what we have discussed in figure 2, the decomposition of the dynamics shown in figures 6(a), 6(b), and 6(c) shows that the Coulomb explosion of the

molecule can be mostly described by the concerted acceleration of the atoms into the direction of component 1 (i.e., the radial explosion) having the largest absolute force at all times. On top of that, the average dynamics are modulated by components 2 and 3. For component 2, the change in sign of force reflects the reversion of movement of C2. The modulation by component 3 adds the temporal sequence in which hydrogens are emitted first, followed later by the heavier carbon-nitrogen ring and finally by the C2-I pair. The additional component 4 describes the out-of-plane dynamics of the I and C2 atoms. This component should be zero by symmetry; the relatively small number (below 0.03 a u) must be attributed to finite sampling.

The right panel in figure 6 shows the variances in displacement [figure 6(d)], velocity [figure 6(e)], and force [figure 6(f)] at each femtosecond time step. The fraction of the variance that can be explained with the first four components obtained by force PCA is indicated in blue, whereas the variance that is not covered by the first four components is indicated in yellow. For the position [figure 6(d)], we can see that due to the explosion the variances steadily increase. At the same time, the variance in velocity [figure 6(e)] peaks shortly after the pulse and then decreases to a constant plateau value. The variance in force [figure 6(f)] shows two peaks, one at ~ 0 fs, where the molecule starts to build up charge, and the second is at ~ 30 fs, where the collision between C2 and I occurs.

As indicated by the color code in the right panel of figure 6, a large contribution of variance in the trajectories occurs within the first four components obtained by force PCA and the variance outside the space spanned by these components (the *unexplained* variance indicated in yellow) is relatively small. This result demonstrates that the first four components obtained by force PCA cover most of the dynamics during the explosion. We are thus able to describe the relevant mechanisms during the explosion with a reduction in dimensionality from 33 to four.

3.6. Relation of ion kinetic energy to charge buildup time

Finally, we look at the relation between the charging up of the molecule and the final momenta. As a measure of the time an atom needs to charge up, we use the time difference between the first and last charge transfer events from or to a given atom as modeled in our simulation via over-the-barrier charge rearrangement processes. We show in figure 7 the relation between the charge buildup time of atom C5 and the kinetic energy of the connected hydrogen atom H5. To ensure a strong effect, we look at higher charges and in contrast to the data inspected before, the figure contains data from all trajectories filtered for a final charge of $+2$ on C5 and $+1$ on H5 (irrespective of the charge states of all other atoms). As can be seen, the hydrogen momentum shows a strong anti-correlation to the charge buildup time of the associated heavier atom. For larger charge buildup time of C5, we can see that the H5 momentum is considerably smaller, whereas with fast charging-up of C5 the H5 momentum is larger. Given the fact that the charge is predominantly created on the iodine atom on the opposite side of

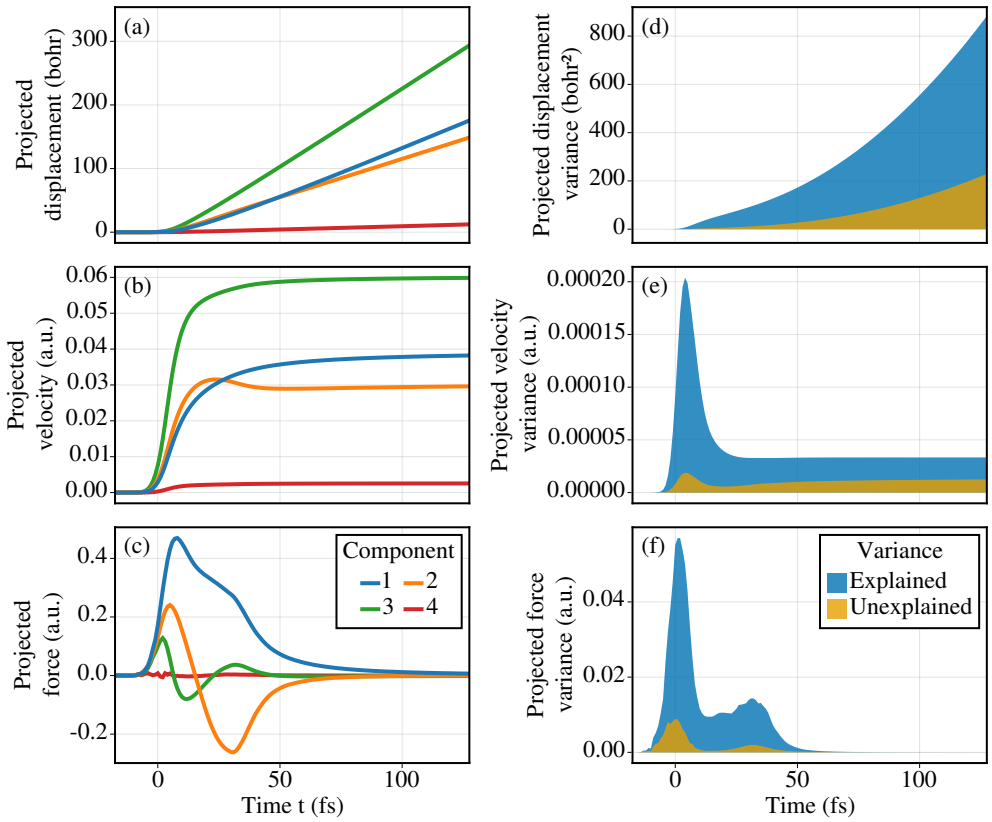

Figure 6: Left panel: Average of the dynamics projected on the first 4 force PCA components in time. (a) Projected displacements. (b) Projected velocities. (c) Projected forces. Right panel: Variance explained by the projections on the first 4 force PCA components in time. (d) Variance of projected displacements. (e) Variance of projected velocities. (f) Variance of projected forces.

the molecule, the relation of the kinetic energy of H5 and charge timing of the carbon atom C5 suggests an opportunity to probe the timescale of charge-buildup on specific atoms and charge redistribution in the molecule.

4. Conclusions

We have simulated the XFEL-induced Coulomb explosion of the 2-iodopyridine molecule and performed an analysis of the resulting simulation data. The x-ray induced Coulomb explosion of the molecule involves an interplay between the expulsion of hydrogens, the expulsion of heavier atoms, and the collision dynamics of the C2 atom with the iodine atom. Among other effects, the collision amplifies out-of-plane motion, producing strongly correlated out-of-plane momenta for the I and C2 ions.

PCA applied to final momenta, which in principle could be available from experimental data, results in components that contain fingerprints of the atomic collision between the I and C2. This result suggests that collisions between two atoms may be

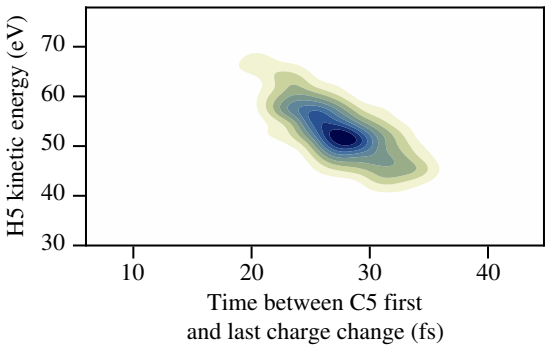

Figure 7: Kinetic energy release of the H5 atom against the time difference between the first and last charge transfer events to the C5 atom for final charge +2 on C5 and +1 on H5.

identified via correlation in their final momenta.

Studying the data employing force PCA reveals components that facilitate the description of the dynamics during the explosion. They allow to interpret the overall dynamics using a limited number of components. Employing only 4 components, we are able to follow the main average dynamics and describe the dominant part of the variances during the Coulomb explosion. Finally, we have demonstrated the strong relation between charge timing of carbon atoms and the kinetic energy release of the associated hydrogen atom.

Our analysis demonstrates that representing data from Coulomb explosion simulations employing components obtained from force PCA is a useful tool allowing to highlight the most relevant dynamical motifs and achieve a decomposition of the dynamics in reduced dimensionality. The results discussed here for iodopyridine may also be applicable to other molecules.

Simulations of x-ray induced Coulomb explosion and their appropriate analysis are essential tools for interpretation of experimental data. Prospectively, the techniques applied here are useful for further investigation towards reconstructing initial molecular geometries from final momentum data,which is one the crucial challenges for Coulomb explosion imaging.

Acknowledgments

This work was supported by the Cluster of Excellence 'Advanced Imaging of Matter' of the Deutsche Forschungsgemeinschaft (DFG) - EXC 2056 - project ID 390715994. L.I. and R.S. acknowledge support from the Chemical Sciences, Geosciences, and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy, grant No. DE-SC0019451.

Bibliography

[1] Yatsuhashi T and Nakashima N 2018 *J. Photochem. Photobiol. C* **34** 52–84

[2] Vager Z, Naaman R and Kanter E P 1989 *Science* **244** 426–431

[3] Sanderson J H, Nishide T, Shiromaru H, Achiba Y and Kobayashi N 1999 *Phys. Rev. A* **59** 4817–4820

[4] Kitamura T, Nishide T, Shiromaru H, Achiba Y and Kobayashi N 2001 *J. Chem. Phys.* **115** 5–6

[5] Cornaggia C, Schmidt M and Normand D 1995 *Phys. Rev. A* **51** 1431–1437

[6] Stapelfeldt H, Constant E and Corkum P B 1995 *Phys. Rev. Lett.* **74** 3780–3783

[7] Stapelfeldt H, Constant E, Sakai H and Corkum P B 1998 *Phys. Rev. A* **58** 426–433

[8] Kou J, Zhakhovskii V, Sakabe S, Nishihara K, Shimizu S, Kawato S, Hashida M, Shimizu K, Bulanov S, Izawa Y, Kato Y and Nakashima N 2000 *J. Chem. Phys.* **112** 5012–5020

[9] Shimizu S, Zhakhovskii V, Sato F, Okihara S, Sakabe S, Nishihara K, Izawa Y, Yatsuhashi T and Nakashima N 2002 *J. Chem. Phys.* **117** 3180–3189

[10] Shimizu S, Zhakhovskii V, Murakami M, Tanaka M, Yatsuhashi T, Okihara S, Nishihara K, Sakabe S, Izawa Y and Nakashima N 2005 *Chem. Phys. Lett.* **404** 379–383

[11] Légaré F, Lee K F, Litvinyuk I V, Dooley P W, Wesolowski S S, Bunker P R, Dombi P, Krausz F, Bandrauk A D, Villeneuve D M and Corkum P B 2005 *Phys. Rev. A* **71** 013415

[12] Ergler T, Rudenko A, Feuerstein B, Zrost K, Schröter C D, Moshammer R and Ullrich J 2005 *Phys. Rev. Lett.* **95** 093001

[13] Rudenko A, Ergler T, Feuerstein B, Zrost K, Schröter C D, Moshammer R and Ullrich J 2006 *Chem. Phys.* **329** 193–202

[14] Gagnon J, Lee K F, Rayner D M, Corkum P B and Bhardwaj V R 2008 *J. Phys. B: At. Mol. Opt. Phys.* **41** 215104

[15] Bocharova I, Karimi R, Penka E F, Brichta J P, Lassonde P, Fu X, Kieffer J C, Bandrauk A D, Litvinyuk I, Sanderson J and Légaré F 2011 *Phys. Rev. Lett.* **107** 063201

[16] Yatsuhashi T, Mitsubayashi N, Itsukashi M, Kozaki M, Okada K and Nakashima N 2011 *ChemPhysChem* **12** 122–126

[17] Matsuda A, Fushitani M, Takahashi E J and Hishikawa A 2011 *PCCP* **13** 8697–8704

[18] Hansen J L, Nielsen J H, Madsen C B, Lindhardt A T, Johansson M P, Skrydstrup T, Madsen L B and Stapelfeldt H 2012 *J. Chem. Phys.* **136** 204310

[19] Corrales M E, Gitzinger G, González-Vázquez J, Loriot V, de Nalda R and Bañares L 2012 *J. Phys. Chem. A* **116** 2669–2677

[20] Pitzer M, Kunitski M, Johnson A S, Jahnke T, Sann H, Sturm F, Schmidt L P H, Schmidt-Böcking H, Dörner R, Stohner J, Kiedrowski J, Reggelin M, Marquardt S, Schießer A, Berger R and Schöffler M S 2013 *Science* **341** 1096–1100

[21] Wales B, Bisson É, Karimi R, Beaulieu S, Ramadhan A, Giguère M, Long Z, Liu W K, Kieffer J C, Légaré F and Sanderson J 2014 *J. Electron Spectrosc. Relat. Phenom.* **195** 332–336

[22] Matsuda A, Takahashi E J and Hishikawa A 2014 *J. Electron Spectrosc. Relat. Phenom.* **195** 327–331

[23] Ibrahim H, Wales B, Beaulieu S, Schmidt B E, Thiré N, Fowe E P, Bisson É, Hebeisen C T, Wanie V, Giguére M, Kieffer J C, Spanner M, Bandrauk A D, Sanderson J, Schuurman M S and Légaré F 2014 *Nat. Commun.* **5** 4422

[24] Kunitski M, Zeller S, Voigtsberger J, Kalinin A, Schmidt L P H, Schöffler M, Czasch A, Schöllkopf W, Grisenti R E, Jahnke T, Blume D and Dörner R 2015 *Science* **348** 551–555

[25] Slater C S, Blake S, Brouard M, Lauer A, Vallance C, Bohun C S, Christensen L, Nielsen J H, Johansson M P and Stapelfeldt H 2015 *Phys. Rev. A* **91** 053424

[26] Pickering J D, Amini K, Brouard M, Burt M, Bush I J, Christensen L, Lauer A, Nielsen J H, Slater C S and Stapelfeldt H 2016 *J. Chem. Phys.* **144** 161105

[27] Burt M, Boll R, Lee J W L, Amini K, Köckert H, Vallance C, Gentleman A S, Mackenzie S R, Bari S, Bomme C, Düsterer S, Erk B, Manschwetus B, Müller E, Rompotis D, Savelyev E,

Schirmel N, Techert S, Treusch R, Küpper J, Trippel S, Wiese J, Stapelfeldt H, de Miranda B C, Guillemin R, Ismail I, Journel L, Marchenko T, Palaudoux J, Penent F, Piancastelli M N, Simon M, Travnikova O, Brausse F, Goldsztejn G, Rouzée A, Géléoc M, Geneaux R, Ruchon T, Underwood J, Holland D M P, Mereshchenko A S, Olshin P K, Johnsson P, Maclot S, Lahl J, Rudenko A, Ziaee F, Brouard M and Rolles D 2017 *Phys. Rev. A* **96** 043415

[28] Amini K, Boll R, Lauer A, Burt M, Lee J W L, Christensen L, Brauβe F, Mullins T, Savelyev E, Ablikim U, Berrah N, Bomme C, Düsterer S, Erk B, Höppner H, Johnsson P, Kierspel T, Krecinic F, Küpper J, Müller M, Müller E, Redlin H, Rouzée A, Schirmel N, Thøgersen J, Techert S, Toleikis S, Treusch R, Trippel S, Ulmer A, Wiese J, Vallance C, Rudenko A, Stapelfeldt H, Brouard M and Rolles D 2017 *J. Chem. Phys.* **147** 013933

[29] Burt M, Amini K, Lee J W L, Christiansen L, Johansen R R, Kobayashi Y, Pickering J D, Vallance C, Brouard M and Stapelfeldt H 2018 *J. Chem. Phys.* **148** 091102

[30] Amini K, Savelyev E, Braußе F, Berrah N, Bomme C, Brouard M, Burt M, Christensen L, Düsterer S, Erk B, Höppner H, Kierspel T, Krecinic F, Lauer A, Lee J W L, Müller M, Müller E, Mullins T, Redlin H, Schirmel N, Thøgersen J, Techert S, Toleikis S, Treusch R, Trippel S, Ulmer A, Vallance C, Wiese J, Johnsson P, Küpper J, Rudenko A, Rouzée A, Stapelfeldt H, Rolles D and Boll R 2018 *Struct. Dyn.* **5** 014301

[31] Pickering J D, Shepperson B, Christiansen L and Stapelfeldt H 2018 *J. Chem. Phys.* **149** 154306

[32] Allum F, Burt M, Amini K, Boll R, Köckert H, Olshin P K, Bari S, Bomme C, Braußе F, Cunha de Miranda B, Düsterer S, Erk B, Géléoc M, Geneaux R, Gentleman A S, Goldsztejn G, Guillemin R, Holland D M P, Ismail I, Johnsson P, Journel L, Küpper J, Lahl J, Lee J W L, Maclot S, Mackenzie S R, Manschwetus B, Mereshchenko A S, Mason R, Palaudoux J, Piancastelli M N, Penent F, Rompotis D, Rouzée A, Ruchon T, Rudenko A, Savelyev E, Simon M, Schirmel N, Stapelfeldt H, Techert S, Travnikova O, Trippel S, Underwood J G, Vallance C, Wiese J, Ziaee F, Brouard M, Marchenko T and Rolles D 2018 *J. Chem. Phys.* **149** 204313

[33] Corrales M E, González-Vázquez J, de Nalda R and Bañares L 2019 *J. Phys. Chem. Lett.* **10** 138–143

[34] Wu H, Xue Y, Wen J, Wang H, Bai L, He W, Sun R and Zheng W 2019 *RSC Adv.* **9** 31853–31859

[35] Lee J W L, Köckert H, Heathcote D, Popat D, Chapman R T, Karras G, Majchrzak P, Springate E and Vallance C 2020 *Commun. Chem.* **3** 1–6

[36] Endo T, Neville S P, Wanie V, Beaulieu S, Qu C, Deschamps J, Lassonde P, Schmidt B E, Fujise H, Fushitani M, Hishikawa A, Houston P L, Bowman J M, Schuurman M S, Légaré F and Ibrahim H 2020 *Science* **370** 1072–1077

[37] Schouder C A, Chatterley A S, Madsen L B, Jensen F and Stapelfeldt H 2020 *Phys. Rev. A* **102** 063125

[38] Zhou W, Ge L, Cooper G A, Crane S W, Evans M H, Ashfold M N R and Vallance C 2020 *J. Chem. Phys.* **153** 184201

[39] Emma P, Akre R, Arthur J, Bionta R, Bostedt C, Bozek J, Brachmann A, Bucksbaum P, Coffee R, Decker F J, Ding Y, Dowell D, Edstrom S, Fisher A, Frisch J, Gilevich S, Hastings J, Hays G, Hering P, Huang Z, Iverson R, Loos H, Messerschmidt M, Miahnahri A, Moeller S, Nuhn H D, Pile G, Ratner D, Rzepiela J, Schultz D, Smith T, Stefan P, Tompkins H, Turner J, Welch J, White W, Wu J, Yocky G and Galayda J 2010 *Nat. Photonics* **4** 641–647

[40] Ishikawa T, Aoyagi H, Asaka T, Asano Y, Azumi N, Bizen T, Ego H, Fukami K, Fukui T, Furukawa Y, Goto S, Hanaki H, Hara T, Hasegawa T, Hatsui T, Higashiya A, Hirono T, Hosoda N, Ishii M, Inagaki T, Inubushi Y, Itoga T, Joti Y, Kago M, Kameshima T, Kimura H, Kirihara Y, Kiyomichi A, Kobayashi T, Kondo C, Kudo T, Maesaka H, Maréchal X M, Masuda T, Matsubara S, Matsumoto T, Matsushita T, Matsui S, Nagasono M, Nariyama N, Ohashi H, Ohata T, Ohshima T, Ono S, Otake Y, Saji C, Sakurai T, Sato T, Sawada K, Seike T, Shirasawa K, Sugimoto T, Suzuki S, Takahashi S, Takebe H, Takeshita K, Tamasaku K, Tanaka H, Tanaka R, Tanaka T, Togashi T, Togawa K, Tokuhisa A, Tomizawa H, Tono K, Wu S, Yabashi M, Yamaga M, Yamashita A, Yanagida K, Zhang C, Shintake T, Kitamura H and Kumagai N 2012 *Nat.*

Photonics **6** 540–544

[41] Allaria E, Castronovo D, Cinquegrana P, Craievich P, Dal Forno M, Danailov M B, D'Auria G, Demidovich A, De Ninno G, Di Mitri S, Diviacco B, Fawley W M, Ferianis M, Ferrari E, Froehlich L, Gaio G, Gauthier D, Giannessi L, Ivanov R, Mahieu B, Mahne N, Nikolov I, Parmigiani F, Penco G, Raimondi L, Scafuri C, Serpico C, Sigalotti P, Spampinati S, Spezzani C, Svandrlik M, Svetina C, Trovo M, Veronese M, Zangrando D and Zangrando M 2013 *Nat. Photonics* **7** 913–918

[42] Kang H S, Min C K, Heo H, Kim C, Yang H, Kim G, Nam I, Baek S Y, Choi H J, Mun G, Park B R, Suh Y J, Shin D C, Hu J, Hong J, Jung S, Kim S H, Kim K, Na D, Park S S, Park Y J, Han J H, Jung Y G, Jeong S H, Lee H G, Lee S, Lee S, Lee W W, Oh B, Suh H S, Parc Y W, Park S J, Kim M H, Jung N S, Kim Y C, Lee M S, Lee B H, Sung C W, Mok I S, Yang J M, Lee C S, Shin H, Kim J H, Kim Y, Lee J H, Park S Y, Kim J, Park J, Eom I, Rah S, Kim S, Nam K H, Park J, Park J, Kim S, Kwon S, Park S H, Kim K S, Hyun H, Kim S N, Kim S, Hwang S m, Kim M J, Lim C y, Yu C J, Kim B S, Kang T H, Kim K W, Kim S H, Lee H S, Lee H S, Park K H, Koo T Y, Kim D E and Ko I S 2017 *Nat. Photonics* **11** 708–713

[43] Decking W, Abeghyan S, Abramian P, Abramsky A, Aguirre A, Albrecht C, Alou P, Altarelli M, Altmann P, Amyan K, Anashin V, Apostolov E, Appel K, Auguste D, Ayvazyan V, Baark S, Babies F, Baboi N, Bak P, Balandin V, Baldinger R, Baranasic B, Barbanotti S, Belikov O, Belokurov V, Belova L, Belyakov V, Berry S, Bertucci M, Beutner B, Block A, Blöcher M, Böckmann T, Bohm C, Böhnert M, Bondar V, Bondarchuk E, Bonezzi M, Borowiec P, Bösch C, Bösenberg U, Bosotti A, Böspflug R, Bousonville M, Boyd E, Bozhko Y, Brand A, Branlard J, Briechle S, Brinker F, Brinker S, Brinkmann R, Brockhauser S, Brovko O, Brück H, Brüdgam A, Butkowski L, Büttner T, Calero J, Castro-Carballo E, Cattalanotto G, Charrier J, Chen J, Cherepenko A, Cheskidov V, Chiodini M, Chong A, Choroba S, Chorowski M, Churanov D, Cichalewski W, Clausen M, Clement W, Cloué C, Cobos J A, Coppola N, Cunis S, Czuba K, Czwalinna M, D'Almagne B, Dammann J, Danared H, de Zubiaurre Wagner A, Delfs A, Delfs T, Dietrich F, Dietrich T, Dohlus M, Dommach M, Donat A, Dong X, Doynikov N, Dressel M, Duda M, Duda P, Eckoldt H, Ehsan W, Eidam J, Eints F, Engling C, Englisch U, Ermakov A, Escherich K, Eschke J, Saldin E, Faesing M, Fallou A, Felber M, Fenner M, Fernandes B, Fernández J M, Feuker S, Filippakopoulos K, Floettmann K, Fogel V, Fontaine M, Francés A, Martin I F, Freund W, Freyermuth T, Friedland M, Fröhlich L, Fusetti M, Fydrych J, Gallas A, García O, Garcia-Tabares L, Geloni G, Gerasimova N, Gerth C, Geßler P, Gharibyan V, Gloor M, Głowinkowski J, Goessel A, Gołębiewski Z, Golubeva N, Grabowski W, Graeff W, Grebentsov A, Grecki M, Grevsmuehl T, Gross M, Grosse-Wortmann U, Grünert J, Grunewald S, Grzegory P, Feng G, Guler H, Gusev G, Gutierrez J L, Hagge L, Hamberg M, Hanneken R, Harms E, Hartl I, Hauberg A, Hauf S, Hauschildt J, Hauser J, Havlicek J, Hedqvist A, Heidbrook N, Hellberg F, Henning D, Hensler O, Hermann T, Hidvégi A, Hierholzer M, Hintz H, Hoffmann F, Hoffmann M, Hoffmann M, Holler Y, Hüning M, Ignatenko A, Ilchen M, Iluk A, Iversen J, Iversen J, Izquierdo M, Jachmann L, Jardon N, Jastrow U, Jensch K, Jensen J, Jeżabek M, Jidda M, Jin H, Johansson N, Jonas R, Kaabi W, Kaefer D, Kammering R, Kapitza H, Karabekyan S, Karstensen S, Kasprzak K, Katalev V, Keese D, Keil B, Kholopov M, Killenberger M, Kitaev B, Klimchenko Y, Klos R, Knebel L, Koch A, Koepke M, Köhler S, Köhler W, Kohlstrunk N, Konopkova Z, Konstantinov A, Kook W, Koprek W, Körfer M, Korth O, Kosarev A, Kosiński K, Kostin D, Kot Y, Kotarba A, Kozak T, Kozak V, Kramert R, Krasilnikov M, Krasnov A, Krause B, Kravchuk L, Krebs O, Kretschmer R, Kreutzkamp J, Kröplin O, Krzysik K, Kube G, Kuehn H, Kujala N, Kulikov V, Kuzminych V, La Civita D, Lacroix M, Lamb T, Lancetov A, Larsson M, Le Pinvidic D, Lederer S, Lensch T, Lenz D, Leuschner A, Levenhagen F, Li Y, Liebing J, Lilje L, Limberg T, Lipka D, List B, Liu J, Liu S, Lorbeer B, Lorkiewicz J, Lu H H, Ludwig F, Machau K, Maciocha W, Madec C, Magueur C, Maiano C, Maksimova I, Malcher K, Maltezopoulos T, Mamoshkina E, Manschwetus B, Marcellini F, Marinkovic G, Martinez T, Martirosyan H, Maschmann W, Maslov M, Matheisen A, Mavric U, Meißner J, Meissner K,

Messerschmidt M, Meyners N, Michalski G, Michelato P, Mildner N, Moe M, Moglia F, Mohr C, Mohr S, Möller W, Mommerz M, Monaco L, Montiel C, Moretti M, Morozov I, Morozov P, Mross D, Mueller J, Müller C, Müller J, Müller K, Munilla J, Münnich A, Muratov V, Napoly O, Näser B, Nefedov N, Neumann R, Neumann R, Ngada N, Noelle D, Obier F, Okunev I, Oliver J A, Omet M, Oppelt A, Ottmar A, Oublaid M, Pagani C, Paparella R, Paramonov V, Peitzmann C, Penning J, Perus A, Peters F, Petersen B, Petrov A, Petrov I, Pfeiffer S, Pflüger J, Philipp S, Pienaud Y, Pierini P, Pivovarov S, Planas M, Pławski E, Pohl M, Polinski J, Popov V, Prat S, Prenting J, Priebe G, Pryschelski H, Przygoda K, Pyata E, Racky B, Rathjen A, Ratuschni W, Regnaud-Campderros S, Rehlich K, Reschke D, Robson C, Roever J, Roggli M, Rothenburg J, Rusiński E, Rybaniec R, Sahling H, Salmani M, Samoylova L, Sanzone D, Saretzki F, Sawlanski O, Schaffran J, Schlarb H, Schlösser M, Schlott V, Schmidt C, Schmidt-Foehre F, Schmitz M, Schmökel M, Schnautz T, Schneidmiller E, Scholz M, Schöneburg B, Schultze J, Schulz C, Schwarz A, Sekutowicz J, Sellmann D, Semenov E, Serkez S, Sertore D, Shehzad N, Shemarykin P, Shi L, Sienkiewicz M, Sikora D, Sikorski M, Silenzi A, Simon C, Singer W, Singer X, Sinn H, Sinram K, Skvorodnev N, Smirnow P, Sommer T, Sorokin A, Stadler M, Steckel M, Steffen B, Steinhau-Kühl N, Stephan F, Stodulski M, Stolper M, Sulimov A, Susen R, Świerblewski J, Sydlo C, Syresin E, Sytchev V, Szuba J, Tesch N, Thie J, Thiebault A, Tiedtke K, Tischhauser D, Tolkiehn J, Tomin S, Tonisch F, Toral F, Torbin I, Trapp A, Treyer D, Trowitzsch G, Trublet T, Tschentscher T, Ullrich F, Vannoni M, Varela P, Varghese G, Vashchenko G, Vasic M, Vazquez-Velez C, Verguet A, Vilcins-Czvitkovits S, Villanueva R, Visentin B, Viti M, Vogel E, Volobuev E, Wagner R, Walker N, Wamsat T, Weddig H, Weichert G, Weise H, Wenndorf R, Werner M, Wichmann R, Wiebers C, Wiencek M, Wilksen T, Will I, Winkelmann L, Winkowski M, Wittenburg K, Witzig A, Wlk P, Wohlenberg T, Wojciechowski M, Wolff-Fabris F, Wrochna G, Wrona K, Yakopov M, Yang B, Yang F, Yurkov M, Zagorodnov I, Zalden P, Zavadtsev A, Zavadtsev D, Zhirnov A, Zhukov A, Ziemann V, Zolotov A, Zolotukhina N, Zummack F and Zybin D 2020 *Nat. Photonics* **14** 391–397

[44] Prat E, Abela R, Aiba M, Alarcon A, Alex J, Arbelo Y, Arrell C, Arsov V, Bacellar C, Beard C, Beaud P, Bettoni S, Biffiger R, Bopp M, Braun H H, Calvi M, Cassar A, Celcer T, Chergui M, Chevtsov P, Cirelli C, Citterio A, Craievich P, Divall M C, Dax A, Dehler M, Deng Y, Dietrich A, Dijkstal P, Dinapoli R, Dordevic S, Ebner S, Engeler D, Erny C, Esposito V, Ferrari E, Flechsig U, Follath R, Frei F, Ganter R, Garvey T, Geng Z, Gobbo A, Gough C, Hauff A, Hauri C P, Hiller N, Hunziker S, Huppert M, Ingold G, Ischebeck R, Janousch M, Johnson P J M, Johnson S L, Juranić P, Jurcevic M, Kaiser M, Kalt R, Keil B, Kiselev D, Kittel C, Knopp G, Koprek W, Laznovsky M, Lemke H T, Sancho D L, Löhl F, Malyzhenkov A, Mancini G F, Mankowsky R, Marcellini F, Marinkovic G, Martiel I, Märki F, Milne C J, Mozzanica A, Nass K, Orlandi G L, Loch C O, Paraliev M, Patterson B, Patthey L, Pedrini B, Pedrozzi M, Pradervand C, Radi P, Raguin J Y, Redford S, Rehanek J, Reiche S, Rivkin L, Romann A, Sala L, Sander M, Schietinger T, Schilcher T, Schlott V, Schmidt T, Seidel M, Stadler M, Stingelin L, Svetina C, Treyer D M, Trisorio A, Vicario C, Voulot D, Wrulich A, Zerdane S and Zimoch E 2020 *Nat. Photonics* **14** 748–754

[45] Young L, Kanter E P, Kraessig B, Li Y, March A M, Pratt S T, Santra R, Southworth S H, Rohringer N, DiMauro L F, Doumy G, Roedig C A, Berrah N, Fang L, Hoener M, Bucksbaum P H, Cryan J P, Ghimire S, Glownia J M, Reis D A, Bozek J D, Bostedt C and Messerschmidt M 2010 *Nature* **466** 56–61

[46] Huang S, Ding Y, Feng Y, Hemsing E, Huang Z, Krzywinski J, Lutman A A, Marinelli A, Maxwell T J and Zhu D 2017 *Phys. Rev. Lett.* **119** 154801

[47] Rudek B, Son S K, Foucar L, Epp S W, Erk B, Hartmann R, Adolph M, Andritschke R, Aquila A, Berrah N, Bostedt C, Bozek J, Coppola N, Filsinger F, Gorke H, Gorkhover T, Graafsma H, Gumprecht L, Hartmann A, Hauser G, Herrmann S, Hirsemann H, Holl P, Hömke A, Journel L, Kaiser C, Kimmel N, Krasniqi F, Kühnel K U, Matysek M, Messerschmidt M, Miesner D, Möller T, Moshammer R, Nagaya K, Nilsson B, Potdevin G, Pietschner D, Reich C, Rupp D,

Schaller G, Schlichting I, Schmidt C, Schopper F, Schorb S, Schröter C D, Schulz J, Simon M, Soltau H, Strüder L, Ueda K, Weidenspointner G, Santra R, Ullrich J, Rudenko A and Rolles D 2012 *Nat. Phot.* **6** 858–865

[48] Rudenko A, Inhester L, Hanasaki K, Li X, Robatjazi S J, Erk B, Boll R, Toyota K, Hao Y, Vendrell O, Bomme C, Savelyev E, Rudek B, Foucar L, Southworth S H, Lehmann C S, Kraessig B, Marchenko T, Simon M, Ueda K, Ferguson K R, Bucher M, Gorkhover T, Carron S, Alonso-Mori R, Koglin J E, Correa J, Williams G J, Boutet S, Young L, Bostedt C, Son S K, Santra R and Rolles D 2017 *Nature* **546** 129–132

[49] Kukk E, Motomura K, Fukuzawa H, Nagaya K and Ueda K 2017 *Appl. Sci.* **7** 531

[50] Schnorr K, Senftleben A, Kurka M, Rudenko A, Schmid G, Pfeifer T, Meyer K, Kübel M, Kling M, Jiang Y, Treusch R, Düsterer S, Siemer B, Wöstmann M, Zacharias H, Mitzner R, Zouros T, Ullrich J, Schröter C and Moshammer R 2014 *Phys. Rev. Lett.* **113** 073001

[51] Jiang Y H, Rudenko A, Herrwerth O, Foucar L, Kurka M, Kühnel K U, Lezius M, Kling M F, van Tilborg J, Belkacem A, Ueda K, Düsterer S, Treusch R, Schröter C D, Moshammer R and Ullrich J 2010 *Phys. Rev. Lett.* **105** 263002

[52] Liekhus-Schmaltz C E, Tenney I, Osipov T, Sanchez-Gonzalez A, Berrah N, Boll R, Bomme C, Bostedt C, Bozek J D, Carron S, Coffee R, Devin J, Erk B, Ferguson K R, Field R W, Foucar L, Frasinski L J, Glownia J M, Gühr M, Kamalov A, Krzywinski J, Li H, Marangos J P, Martinez T J, McFarland B K, Miyabe S, Murphy B, Natan A, Rolles D, Rudenko A, Siano M, Simpson E R, Spector L, Swiggers M, Walke D, Wang S, Weber T, Bucksbaum P H and Petrovic V S 2015 *Nat. Commun.* **6** 8199

[53] Erk B, Boll R, Trippel S, Anielski D, Foucar L, Rudek B, Epp S W, Coffee R, Carron S, Schorb S, Ferguson K R, Swiggers M, Bozek J D, Simon M, Marchenko T, Küpper J, Schlichting I, Ullrich J, Bostedt C, Rolles D and Rudenko A 2014 *Science* **345** 288–291

[54] Li Z, Inhester L, Liekhus-Schmaltz C, Curchod B F E, Snyder J W, Medvedev N, Cryan J, Osipov T, Pabst S, Vendrell O, Bucksbaum P and Martinez T J 2017 *Nat. Commun.* **8** 453

[55] Boll R, Schäfer J M, Richard B, Fehre K, Kastirke G, Jurek Z, Schoffler M S, Abdullah M M, Anders N, Baumann T M, Eckart S, Erk B, Fanis A D, Dorner R, Grundmann S, Grychtol P, Hartung A, Hofmann M, Ilchen M, Inhester L, Janke C, Jin R, Kircher M, Kubicek K, Kunitski M, Li X, Mazza T, Meister S, Melzer N, Montano J, Music V, Nalin G, Ovcharenko Y, Passow C, Pier A, Rennhack N, Rist J, Rivas D E, Rolles D, Schlichting I, Schmidt L P H, Schmidt P, Siebert J, Strenger N, Trabert D, Trinter F, Vela-Perez I, Wagner R, Walter P, Weller M, Ziolkowski P, Son S K, Rudenko A, Meyer M, Santra R and Jahnke T 25

[56] Jurek Z, Son S K, Ziaja B and Santra R 2016 *J. Appl. Crystallogr.* **49** 1048–1056

[57] Schmidt M W, Baldridge K K, Boatz J A, Elbert S T, Gordon M S, Jensen J H, Koseki S, Matsunaga N, Nguyen K A, Su S, Windus T L, Dupuis M and Montgomery J A 1993 *J. Comput. Chem.* **14** 1347–1363

[58] Ryufuku H, Sasaki K and Watanabe T 1980 *Phys. Rev. A* **21** 745–750

[59] Stein S A M, Loccisano A E, Firestine S M and Evanseck J D 2006 Chapter 13 Principal Components Analysis: A Review of its Application on Molecular Dynamics Data *ARCC* vol 2 ed Spellmeyer D C (Elsevier) pp 233–261

[60] 2002 Mathematical and Statistical Properties of Sample Principal Components *Principal Component Analysis* Springer Series in Statistics ed Jolliffe I T (New York, NY: Springer) pp 29–61 ISBN 978-0-387-22440-4

[61] Amadei A, Linssen A B M and Berendsen H J C 1993 *Proteins Struct. Funct. Bioinforma.* **17** 412–425

[62] Hub J S and de Groot B L 2009 *PLOS Comput. Biol.* **5** e1000480

[63] Li X, Xie Y, Hu D and Lan Z 2017 *J. Chem. Theory Comput.* **13** 4611–4623

[64] Li X, Hu D, Xie Y and Lan Z 2018 *J. Chem. Phys.* **149** 244104

[65] Post M, Wolf S and Stock G 2019 *J. Chem. Phys.* **150** 204110

[66] Taylor M H, Losch M, Wenzel M and Schröter J 2013 *J. Clim.* **26** 9194–9205

[67] Domcke W and Cederbaum L S 1978 *J. Electron Spectrosc. Relat. Phenom.* **13** 161–173
[68] Kukk E, Ueda K, Hergenhahn U, Liu X J, Prümper G, Yoshida H, Tamenori Y, Makochekanwa C, Tanaka T, Kitajima M and Tanaka H 2005 *Phys. Rev. Lett.* **95** 133001
[69] Thomas T D, Kukk E, Sankari R, Fukuzawa H, Prümper G, Ueda K, Püttner R, Harries J, Tamenori Y, Tanaka T, Hoshino M and Tanaka H 2008 *J. Chem. Phys.* **128** 144311
[70] Kreidi K, Demekhin P V, Jahnke T, Weber T, Havermeier T, Liu X J, Morisita Y, Schössler S, Schmidt L P H, Schöffler M, Odenweller M, Neumann N, Foucar L, Titze J, Ulrich B, Sturm F, Stuck C, Wallauer R, Voss S, Lauter I, Kim H K, Rudloff M, Fukuzawa H, Prümper G, Saito N, Ueda K, Czasch A, Jagutzki O, Schmidt-Böcking H, Scheit S, Cederbaum L S and Dörner R 2009 *Phys. Rev. Lett.* **103** 033001
[71] Thomas T D, Kukk E, Fukuzawa H, Ueda K, Püttner R, Tamenori Y, Asahina T, Kuze N, Kato H, Hoshino M, Tanaka H, Meyer M, Plenge J, Wirsing A, Serdaroglu E, Flesch R, Rühl E, Gavrilyuk S, Gel'mukhanov F, Lindblad A and Sæthre L J 2009 *Phys. Rev. A* **79** 022506
[72] Thomas T D, Kukk E, Ouchi T, Yamada A, Fukuzawa H, Ueda K, Püttner R, Higuchi I, Tamenori Y, Asahina T, Kuze N, Kato H, Hoshino M, Tanaka H, Lindblad A and Sæthre L J 2010 *J. Chem. Phys.* **133** 174312
[73] Kukk E, Ayuso D, Thomas T D, Decleva P, Patanen M, Argenti L, Plésiat E, Palacios A, Kooser K, Travnikova O, Mondal S, Kimura M, Sakai K, Miron C, Martín F and Ueda K 2013 *Phys. Rev. A* **88** 033412
[74] Simon M, Püttner R, Marchenko T, Guillemin R, Kushawaha R K, Journel L, Goldsztejn G, Piancastelli M N, Ablett J M, Rueff J P and Céolin D 2014 *Nat. Commun.* **5** 4069
[75] Kukk E, Thomas T D, Ueda K, Céolin D, Granroth S, Kooser K, Travnikova O, Iablonsky D, Decleva P, Ayuso D, Püttner R, Levola H, Goldsztejn G, Marchenko T, Piancastelli M N and Simon M 2017 *Phys. Rev. A* **95** 042509
[76] Kukk E, Thomas T D, Céolin D, Granroth S, Travnikova O, Berholts M, Marchenko T, Guillemin R, Journel L, Ismail I, Püttner R, Piancastelli M N, Ueda K and Simon M 2018 *Phys. Rev. Lett.* **121** 073002
[77] Céolin D, Liu J C, da Cruz V V, Ågren H, Journel L, Guillemin R, Marchenko T, Kushawaha R K, Piancastelli M N, Püttner R, Simon M and Gel'mukhanov F 2019 *PNAS* **116** 4877–4882
[78] Kircher M, Rist J, Trinter F, Grundmann S, Waitz M, Melzer N, Vela-Pérez I, Mletzko T, Pier A, Strenger N, Siebert J, Janssen R, Schmidt L P H, Artemyev A N, Schöffler M S, Jahnke T, Dörner R and Demekhin P V 2019 *Phys. Rev. Lett.* **123** 243201
[79] Grundmann S, Kircher M, Vela-Perez I, Nalin G, Trabert D, Anders N, Melzer N, Rist J, Pier A, Strenger N, Siebert J, Demekhin P V, Schmidt L P H, Trinter F, Schöffler M S, Jahnke T and Dörner R 2020 *Phys. Rev. Lett.* **124** 233201
[80] Thomas T D 2020 *Phys. Rev. A* **101** 043415