Imaging collective quantum fluctuations of the structure of a complex molecule

Benoît Richard[1,2,3], Rebecca Boll[4#], Sourav Banerjee[1], Julia M. Schäfer[1,5,‡], Zoltan Jurek[1], Gregor Kastirke[6,7], Kilian Fehre[6], Markus S. Schöffler[6], Nils Anders[6], Thomas M. Baumann[4], Sebastian Eckart[6], Benjamin Erk[8], Alberto De Fanis[4], Reinhard Dörner[6,7], Sven Grundmann[6], Patrik Grychtol[4], Max Hofmann[6], Markus Ilchen[1,2,4], Max Kircher[6], Katharina Kubicek[2,3,4], Maksim Kunitski[6], Xiang Li[10,14], Tommaso Mazza[4], Severin Meister[11], Niklas Melzer[6], Jacobo Montano[4], Valerija Music[4,9], Yevheniy Ovcharenko[4], Christopher Passow[8], Andreas Pier[6], Nils Rennhack[4], Jonas Rist[6], Daniel E. Rivas[4], Daniel Rolles[10], Ilme Schlichting[12], Lothar Ph. H. Schmidt[6], Philipp Schmidt[4,9], Daniel Trabert[6], Florian Trinter[6,13], Rene Wagner[4], Peter Walter[14], Pawel Ziolkowski[4], Artem Rudenko[10], Michael Meyer[3,4], Robin Santra[1,2,3,5], Ludger Inhester[1,3,$], and Till Jahnke[4,11†]

[1] Center for Free-Electron Laser Science CFEL, Deutsches Elektronen-Synchrotron DESY, 22603 Hamburg, Germany

[2] Department of Physics, University of Hamburg, 22603 Hamburg, Germany

[3] The Hamburg Centre for Ultrafast Imaging, 22761 Hamburg, Germany

[4] European XFEL, 22869 Schenefeld, Germany

[5] Department of Chemistry, University of Hamburg, 20146 Hamburg, Germany

[6] Institut für Kernphysik, Goethe-Universität Frankfurt, 60438 Frankfurt am Main, Germany

[7] Helmholtz Forschungsakademie Hessen für FAIR (HFHF) GSI Helmholtzzentrum für Schwerionenforschung Campus Fr

[8] Deutsches Elektronen-Synchrotron DESY, 22603 Hamburg, Germany

[9] Institut für Physik und CINSaT, Universität Kassel, 34132 Kassel, Germany

[10] J. R. Macdonald Laboratory, Department of Physics, Kansas State University, Manhattan, KS 66506, USA

[11] Max-Planck-Institut für Kernphysik, 69117 Heidelberg, Germany

[12] Max-Planck-Institut für medizinische Forschung, 69120 Heidelberg, Germany

[13] Molecular Physics, Fritz-Haber-Institut der Max-Planck-Gesellschaft, 14195 Berlin, Germany

[14] SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

‡ Present Affiliation: Heliatek GmbH, 01139 Dresden, Germany [#] Email: rebecca.boll@xfel.eu

[$] Email: ludger.inhester@cfel.de [†] Email: till.jahnke@xfel.

One sentence summary: Coulomb explosion imaging enhanced by a new data-science approach captures high-dim. molecular ground-state fluctuations.

Due to the Heisenberg uncertainty principle, the structure of a molecule fluctuates about its mean geometry, even in the ground state. Observing this fundamental quantum effect experimentally, particularly revealing the collective nature of the structural quantum fluctuations, remains an unmet challenge for complex molecules. In this work, we achieved this for an 11-atom molecule by inducing its Coulomb explosion with an x-ray free-electron laser. We show that the structural fluctuations manifest themselves in correlated variations of ion momenta obtained via coincident detection of the atomic fragments from individual molecules. Our analysis scheme allows extracting these variations, despite our measurements covering only a fraction of the full 33-dimensional momentum space, thereby establishing a general approach for extracting information on high-dimensional structural dynamics using Coulomb explosion.

Near absolute zero temperature, matter exhibits structural fluctuations that are fully absent in a classical description and that can significantly affect its properties. These quantum fluctuations,

termed *ground-state structural fluctuations* (GSFs) (or *zero-point fluctuations*), arise from the vibrational wave function of the system and play a key role in many phenomena: at high pressure, causing various types of quantum phase transitions (*1, 2*), for example, to a symmetric phase of ice (*3*), or to an antiferroelectric phase of anhydrous sodium hydroxide (*4*); inhibiting or promoting glass transitions (*5*); imposing structural changes that lead to superconductivity at high temperatures (*6*); suppressing the formation of charge-density-wave phases in layered transition-metal dichalcogenides (*7*); strengthening or weakening hydrogen bonds (*8*); or influencing cross sections of chemical reactions (*9*).

GSFs are complex, as the atoms of bound systems fluctuate in a concerted manner following specific collective patterns. This intricate behavior emerges already in individual molecules: Their structural properties are determined by their respective ground-state potential energy surfaces, which are many-body potentials that are challenging to visualize and characterize. Although the equilibrium structure of a molecule corresponds to the global minimum on its potential energy surface, its structure is in reality fuzzy, as the aforementioned quantum fluctuations explore the potential energy landscape. Matching the high dimensionality of the potential energy surface, these fluctuations involve many different degrees of freedom, which results in their strongly collective nature. In this work, we demonstrate that this collective nature of structural quantum fluctuations in complex molecules can be directly observed.

Indirect evidence for this quantum phenomenon was found already one hundred years ago, when Robert S. Mulliken compared energy levels of different isotopologues (*10*). By now, the measurement of discrete molecular energy levels and the identification of zero-point energies are routine (*11*). However, typically one only measures the consequences of the ground-state quantum fluctuations on molecular energy levels, but cannot observe the underlying structural manifestation of these fluctuations, and the relevance of GSFs is usually inferred from theoretical modeling (*3–7*). To experimentally image the GSFs themselves and obtain critical insight into collective quantum phenomena such as the ones mentioned above, it is necessary to target individual molecules in order to access their entire high-dimensional phase space and thus the structural dependencies among the atoms constituting the molecules. This is out of reach for traditional imaging techniques such as crystallography, because those sample ensembles of molecules and thus inherently only provide information about the average structure.

In contrast, Coulomb explosion imaging (CEI) is a method that is distinctly suited to performing the required single-molecule measurements. In CEI, a molecule fragments violently due to the Coulomb repulsion following from a rapid charge-up. In the ideal case, the molecule is fully fragmented into atomic ions on a short timescale, and the information on its initial geometry is retrieved by analyzing the momenta of the fragment ions. The charge required to trigger an effective fragmentation was first generated by passing molecular ions through a thin foil (stripping off electrons) (*12*), and later by using femtosecond laser pulses (*13–20*), synchrotron light (*21–24*), or intense (x-ray) free-electron laser pulses (*25–30*). Previous experiments were able to highlight structural fluctuations of small, symmetric molecules using CEI, including the correlations between atoms necessary to probe the GSFs. For example, CO_2 molecules, whose nominal ground-state geometry is linear, were found to deviate from linearity most of the time (*22*), and deuterated methane (CD_4) – highly symmetric in its equilibrium state – is in fact asymmetric and even chiral due to its structural fluctuations (*31*). These previous works relied on obtaining the full n-body distribution of ions by the measurement of all (or all but one) ions in coincidence. Such full measurement is impossible for complex molecules due to the finite detection efficiency of the ion detectors, limiting previous results to small molecules. Here, we overcome this barrier, opening the possibility to apply the method to a much larger class of complex and extended molecules. In this work, we show that, by using intense, femtosecond x-ray pulses and an elaborate analysis of only a subset of the fragment ions detected in coincidence, the full dimensionality of the structural fluctuations and of their intramolecular correlations can be probed, which goes far beyond the visualization of independent fragments in the molecular frame presented in our previous work (*30*). In particular, it allowed us to identify fingerprints of specific ground-state normal modes (NMs) in our data through correlated fluctuations involving more than five atoms. This result constitutes an important advance towards measuring the collective fluctuations of atoms in molecules.

Results and Discussion

Coulomb explosion of individual 2-iodopyridine molecules (C_5H_4NI, see equilibrium geometry in Fig. 1) was induced by intense, femtosecond soft x-ray pulses from the European X-ray Free-Electron Laser (EuXFEL). A COLTRIMS reaction microscope (*32, 33*) available at the Small

Quantum Systems (SQS) instrument was used to image the three-dimensional momenta of the emitted fragment ions in coincidence. More details on the experiment are given in the supplementary materials (SM) section S1 and in our earlier paper (*30*). The x-ray fluence at SQS is sufficiently high to trigger a full Coulomb explosion of the molecule, dissociating it into atomic fragments, and the repetition rate is high enough to detect many atomic fragments in coincidence. This setup provides a wealth of in-depth information, as will become evident in the following. Due to the finite ion detection efficiency in the experiment, the recorded multidimensional data is typically incomplete, which poses a fundamental analysis challenge (*34, 35*). We have overcome this problem for an N-atom molecule (here $N = 11$) by reconstructing the full N-fold coincidence information from sets of coincident measurements of n ions ($n < N$), with a sophisticated data analysis technique fitting a multivariate distribution to the measured data.

Figure 1 shows the experimental results of Coulomb explosion imaging of the iodopyridine molecule. Here and in the following, we focus on a specific coincidence channel where an I^{4+} ion, an N^{2+} ion, and at least one C^{2+} ion and one H^{+} ion were measured in coincidence. Panels A to D depict the carbon and hydrogen ions' three-dimensional momentum distributions projected onto two different planes in the molecular recoil frame. The emission direction of the iodine ion defines the z axis and spans the xz plane together with the emission direction of the nitrogen ion. All ion momenta have been normalized such that, for each coincident measurement, the magnitude of the iodine momentum equals one. The projections on the xz plane (Figs. 1A and 1C) clearly resemble the molecular structure, as reported earlier (*30*). The projections on the yz plane (Figs. 1B and 1D) reveal that, remarkably, many ions were emitted with substantial out-of-plane momentum, despite the planar equilibrium geometry of the molecule. To further study this feature, we switched to spherical coordinates, sketched in Fig. 1E, resulting in the angular emission distribution of the carbon ions (Fig. 1F). The four pronounced peaks correspond to the C3, C4, C5, and C6 carbon ions (see Fig. S6 in the supplemental materials for a sketch of the molecule with full labeling). The C2 ion has comparably low momentum (see panel A) and lies close to the origin of the spherical coordinate system; thus it appears as a broad feature in Fig. 1F overlapping partially with the peaks for the C4 and C5 carbon ions. This angular representation of the coincidence data allows visualizing the dependence between the carbon and hydrogen out-of-plane momenta: The panels in Fig. 1G depict carbon-ion angular emission distributions corresponding to specific emission

directions of the H3 hydrogen ion (marked with boxes in Fig. 1D). If the H3 ion was emitted within the molecular plane (middle box), the carbon ions were on average also emitted in plane. If the H3 ion was either emitted upwards (upper box) or downwards (lower box) with respect to the molecular plane, the carbon angular distribution became tilted, with C3 and C4 being emitted in the same direction as H3, and C6 in the opposite direction. C5 remained centered in plane. The full dependence between the H3 emission direction and the out-of-plane emission angle of C3 to C6 is summarized in Fig. 1H where we depict the mean emission angle Θ of these carbon ions in dependence of the out-of-plane momentum of the H3 proton. Similar correlations also exist for other hydrogen atoms, (SM section S2, where we extracted the full correlated behavior for all hydrogen and carbon ion pairs).

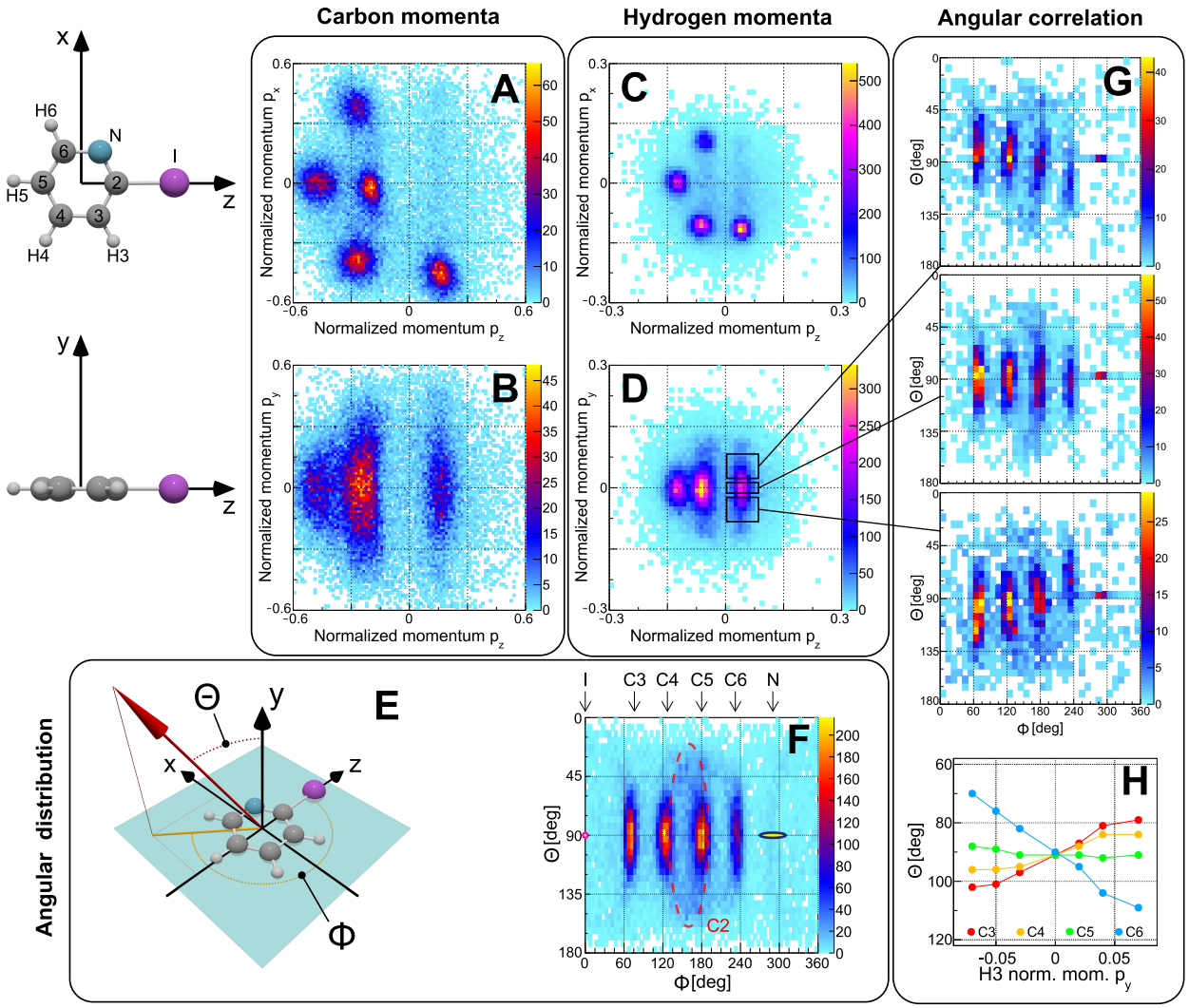

Figure 1: **Correlated out-of-plane fluctuations probed by Coulomb explosion.** The sketches on the left depict the ground-state equilibrium geometry of iodopyridine and the recoil frame, for reference. A. to D. Newton diagrams of ion momenta, depicting the C^{2+} ions (A,B) and protons (C,D) in the molecular frame. All momenta are normalized relative to the momentum of the I^{4+} ion. E. Definition of the spherical coordinates used. F. Angular emission distributions of the carbon ions, in the coordinates defined in panel E. G. Angular emission distributions of the carbon ions when restricted to different emission directions of the H3 proton. H. Angular emission dependence between H3 and C3 to C6. The error bars are smaller than the symbols. The plotted data contain four-fold ion coincidence events including I^{4+}, N^{2+}, C^{2+}, and H^{+}.

Origin of out-of-plane fluctuations

The experimental results presented in Fig. 1 show the rich information content available in such multidimensional data sets. In the following, we will demonstrate that the observed momentum fluctuations directly reflect quantum fluctuations present in the molecule in its vibrational ground state. Because the equilibrium geometry of iodopyridine is planar, the contributions of the quantum fluctuations can be expected to be particularly clear in the out-of-plane momenta, and we found that, indeed, initial out-of-plane GSFs are reflected in the final momenta after Coulomb explosion. In-plane momenta, in contrast, are more strongly impacted by additional variations in the mutual ion interaction during the explosion dynamics caused by variations in the timing of individual ionization events.

In order to elucidate how GSFs relate to the observed fluctuations in the measured ion momenta, we performed a simulation of the x-ray-induced ionization and the subsequent Coulomb explosion of the molecule. To this end, we utilized the XMDYN package (*36*) in combination with the reactive force field method (*37, 38*) implemented in PuReMD (*39*). We used an ensemble of initial atomic positions and velocities that samples the corresponding vibrational ground-state Wigner distribution (see SM section S3 for further details on the simulation). In Fig. 2, we examine the out-of-plane behavior as predicted by our simulation and compare the simulation results to the momenta observed in the experiment, as well as to the magnitude of the quantum fluctuations. Figures 2A, 2B, 2C, and 2D show the out-of-plane Newton diagrams for protons and carbon ions, respectively. The simulations took into consideration GSFs. The predicted momentum distributions after the Coulomb explosion are in good agreement with the measured data. The impact of the GSFs becomes evident by comparison with Figs. 2E and 2F, which depict simulation results where GSFs have been neglected (see supplementary material section S3.4 for a quantitative comparison). In this case, out-of-plane momenta can only be caused by the recoil from the electrons emitted during the charge-up, which has a comparatively minor effect. The GSFs themselves are depicted in Figs. 2G-J showing the out-of-plane momentum and position distribution in the initial, neutral equilibrium state. As can be seen, the initial-state momentum fluctuations (Figs. 2G and 2H) are much smaller than the variations in the final momenta after Coulomb explosion (Figs. 2A and 2B). The position fluctuations are depicted in Fig. 2I (hydrogen atoms) and Fig. 2J (carbon atoms).

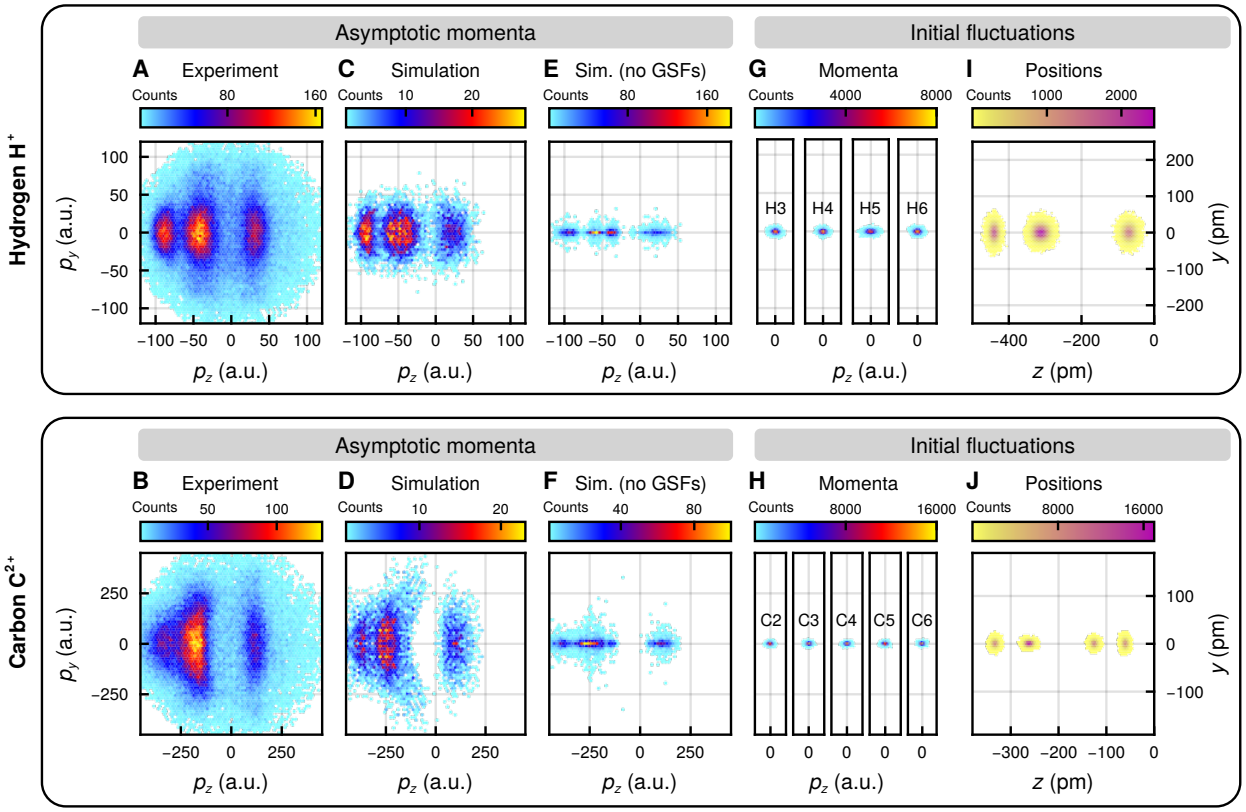

Figure 2: **Out-of-plane momenta of the hydrogen and carbon ions, from experiment and simulation.** A, B. Experiment. C, D. Simulation with initial conditions from the ground-state Wigner distribution. E, F. Simulation with the initial geometry restricted to the equilibrium geometry. G–J. Initial-state momentum and position fluctuations as sampled from the Wigner distribution. The $(0, 0)$ position is the center of mass of the molecule. ('a.u.' stands for 'atomic units', the asymptotic momenta belong to the I^{4+}, N^{2+}, C^{2+}, H^{+} coincidence channel.)

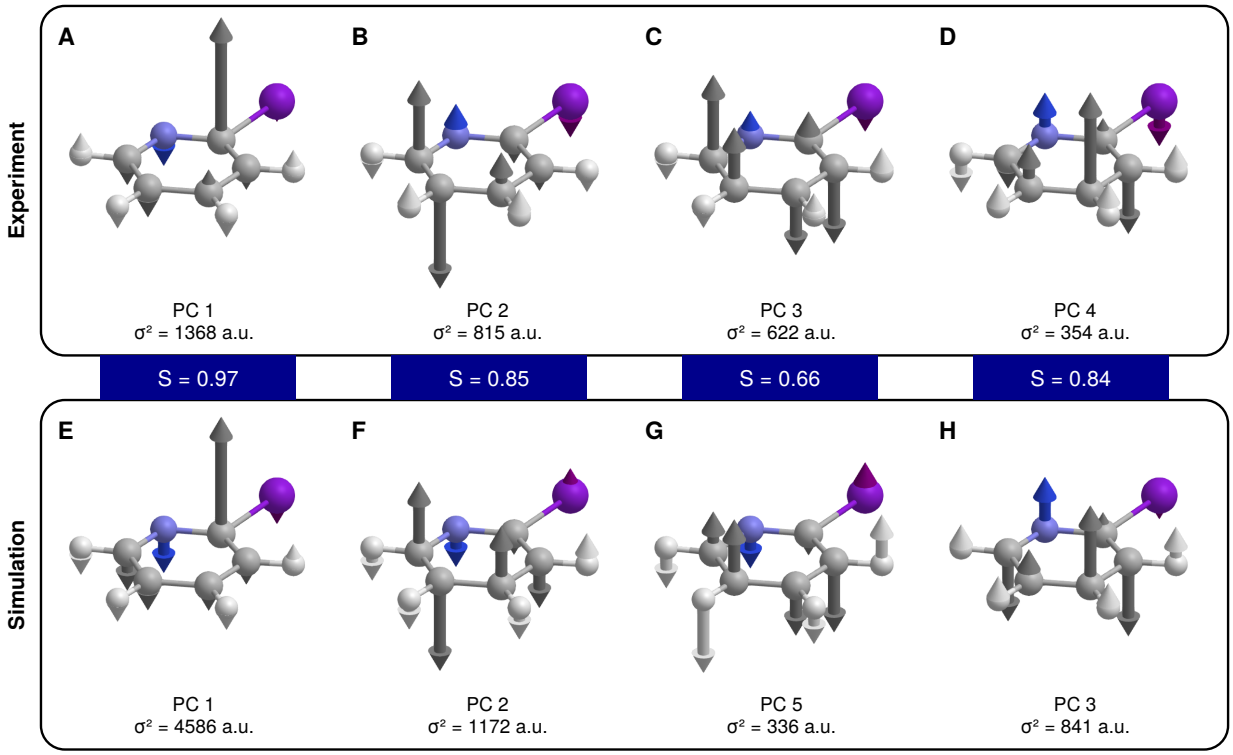

Figure 3: **Dominant out-of-plane principal components of the distribution of asymptotic momenta.** A-D. Reconstructed from the experimental data using the fitting algorithm described in the text. E-H. Reconstructed from the XMDYN simulation data. σ^2 is the variance associated with each PC (explained variance). For a better comparison, we changed the order of the simulation PCs in panels E-H. The blue boxes show the mutual overlap S between the two linked PCs, computed as their scalar product (maximum 1 for identical PCs).

Momentum-space reconstruction and correlation analysis

To access the correlations among the momenta in the final state, we have to determine the full 33-dimensional momentum distribution of all eleven ionic fragments of the Coulomb-exploded molecule. In order to do so, we need to overcome three challenges. (*i*) Although the mass-to-charge ratio of each detected ion is measured in the experiment, its initial position within the molecule is unknown (for those cases where multiple atoms of the same element are present, carbon and hydrogen atoms for iodopyridine). (*ii*) The orientation of each molecule in the laboratory frame is random and unknown. Therefore, the molecular frame has to be determined based on the recorded ion momenta. In Fig. 1, for example, we used the recoil frame defined by the iodine and nitrogen

ions (see SM section S5). (*iii*) Only a subset of all ions created by the x-ray pulse is detected. In the present experiment, relevant statistics could be achieved for up to 7 ions in coincidence (SM section S1.7), due to the typical detection efficiency of approximately 60% per ion. We overcame these problems and recovered the momentum distribution in its full dimensionality, by fitting a 33-dimensional Gaussian distribution to the measured data that involved 4 or more ions detected in coincidence. To make the fit independent of the initial (laboratory-frame) orientation of the molecule, we used scalar products between the measured momenta as internal coordinates. In this way, we avoided any bias from choosing specific ion momenta to define the molecular frame. The Gaussian distribution was then fitted by matching the statistical moments of the distribution of the scalar products, a procedure that is not biased by undetected ions and was performed in a manner that does not rely on the (unknown) identity of each ion. More details on the fitting procedure can be found in the supplementary materials (section S7).

We performed dimensional reduction of the resulting momentum distribution by principal component analysis (PCA) (*40, 41*), based on the fitted Gaussian model (see SM section S8.3). Heavy atoms stay longer in the repulsive Coulomb potential and acquire, in general, both a larger momentum and a larger variance in their momentum. Therefore, to compare light and heavy atoms on a more equal footing, we performed the fit on the momenta weighted by the inverse square root of the ion mass, as commonly done in normal-mode analysis. The resulting principal components (PCs) of the weighted momenta are collective, 33-dimensional vectors that describe the dominant variations in the final ion momenta.

The four dominant out-of-plane PCs extracted from the experimental data are shown in Figs. 3A-D together with their explained variance σ^2. The corresponding PCs obtained from the simulation, using the same procedure, are depicted in Figs. 3E-H. The order of the simulation components has been altered in order to facilitate the comparison.

The similarity of the PCs may be further characterized by their mutual overlaps (defined as the absolute value of their scalar product)given inbetween the respective pairs of PCs. Apart from the somewhat different values in the explained variances, the PCs extracted from the simulation data show strong similarity (overlap between 0.66 and 0.97) with the respective experimental PCs, indicating that the collective momentum fluctuations found in the experiment are in agreement with those obtained from the simulation.

Fingerprints of collective ground-state fluctuations

The good agreement with the simulation results allowed us to directly relate the observed correlated behavior of the ions in the momentum-space data (Fig. 3) to the collective nature of quantum fluctuations in the molecular ground state. To that end, we inspected the GSF along individual normal-mode coordinates before the molecule interacts with the x-ray pulse and determined the contribution of the GSF along each NM to the variance in the final momenta along a given PC (SM section S9).

For the three dominant out-of-plane PCs displayed in Figs. 4A, 4E, and 4I, the mode contributions are shown in Figs. 4B, 4F, and 4J. The analysis for the other dominant modes is given in the supplementary materials (section S9). The Figures 4C, 4D, 4G, 4H, 4K, 4L, show the two NMs with the highest contribution to the respective PC. Each of the three PCs can be associated with few NMs. PC 1 is dominated by the variance of the C2 ion, resulting from a near collision between this carbon ion and the iodine atom, as described in our previous works (*30, 40*). Accordingly, PC 1 is promoted by NMs where the C2 atom strongly shifts out-of-plane, such as the NMs at 16.5 THz (Fig. 4C) and 27.1 THz (Fig. 4D). In the other cases, the connection between the NMs and the PCs is directly visually evident. For example, the displacement of atoms along the NMs at 26.1 THz (Fig. 4G) and 36.3 THz (Fig. 4H) corresponds to a similar pattern of correlated variations in the atomic CEI momenta along PC 2 (Fig. 4E). In the explosion, some of the collective out-of-plane patterns inherent in the GSFs were amplified, and, thus, transferred into similar momentum patterns. This visual similarity is, however, lower when looking at the hydrogen atoms, as the contribution from the NMs are opposite and can cancel out. Similar observations can be made for PC 3 (Fig. 4I): there is a low-frequency mode (14.4 THz) and a high-frequency mode (34.9 THz), both having a matching carbon pattern with PC 3, but different contributions for the hydrogen atoms.

Figure 4 establishes that the correlated fluctuations occurring in the measured momenta can indeed be traced back to the effect of collective quantum fluctuations along individual NMs. The measured ion momenta thus provide access to the collective structural fluctuations present in the ground state of the molecule. Further analysis provided in the supplementary materials (section S11) also demonstrates that the patterns highlighted in Fig. 1 are associated primarily with the GSFs along the NMs with frequency 16.5 THz and 36.3 THz, respectively. Furthermore, our simulation results

indicate that such a clear mapping between initial-state quantum fluctuations and final-state CEI momentum distributions is only possible for the out-of plane momenta, when several dominating modes are considered. Variations of in-plane momenta, in contrast, can only partially be associated with GSFs, illustrating that they are much more influenced by the additional contributions imprinted on the explosion dynamics due to variations in the ionization dynamics and the strong nonlinear interaction among the ions during the charge-up and the explosion. These effects could in the future be partially overcome by reducing the fluctuations in the ionization dynamics through shorter x-ray pulses and by targeting even higher ion charge states. Along the same line it should be noted that a clear correspondence of the experimental data with the simulated explosion dynamics could only be reached when utilizing coincidence events with sufficiently high charge. Coincidence data sets consisting of singly charged ions are considerably impacted by the influence of chemical bonding during the explosion, such that discrepancies to the Coulomb explosion simulation appear (SM section S10). Only the high x-ray fluence available at the SQS instrument of the EuXFEL produced such a high charge in the molecule that a direct connection between the CEI data and GSFs could be demonstrated in this experiment.

Conclusion

We measured fingerprints of the ground-state fluctuations in a complex molecule, originating from the Heisenberg uncertainty principle. The fingerprints emerge as collective features in the momenta of the ions generated in an x-ray-free-electron-laser-induced Coulomb explosion. This achievement was made possible via an advanced analysis technique that turns Coulomb explosion imaging into a powerful tool for measuring highly collective structural fluctuations that are inaccessible to any ensemble-averaging imaging method. Our analysis demonstrated that the variances present in incomplete coincidence sets are sufficient to construct a Gaussian model for the full-dimensional data, thereby also capturing collective features. Thanks to refined simulations of the Coulomb explosion, the measured momentum variances could be linked to ground-state fluctuations along individual normal modes exhibiting correlations among multiple atoms. Our work thereby shows that it is possible to probe high-dimensional potential energy surfaces by observing the spontaneous structural fluctuations that molecules undergo due to quantum effects. Thanks to its ability to address

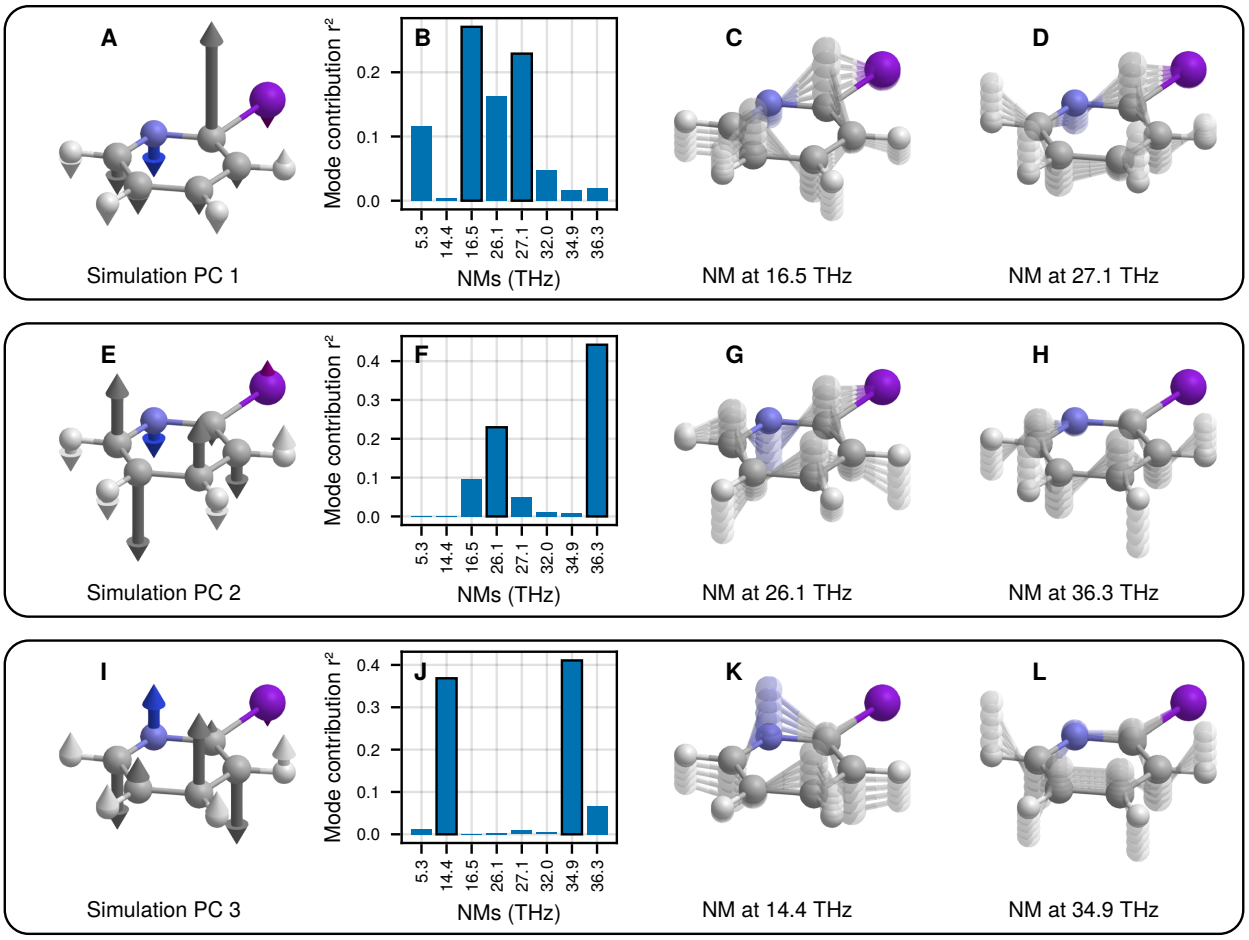

Figure 4: **Mapping of the principal components to the NMs.** A, E, I. Out-of-plane-momentum principal components computed from the simulation. B, F, J. Mode contributions r_{ij}^2 to the respective momentum principal components from each NM. The two leading contributions are highlighted and the corresponding NMs are shown in the next panels. C-D, G-H, K-L. NMs with the dominant contribution. In order to improve the visibility of the collective motion of hydrogen and heavier atoms, the NMs are shown via position variations that are weighted by the inverse square root of the atom mass. For the absolute scale of the GSFs see Table S1 and Fig. S6 in the SM.

highly collective effects and delocalized systems, our work also opens up new avenues to explore chemical dynamics, where the nuclear wave packet tends to spread and split over different reaction pathways. The visualization of the collective motion of the atoms rather than only the centroid of their distribution, as is the case for many other experimental methods, offers incomparably more detailed insight into the time evolution of molecular systems. In the future, one might be even able to simultaneously track electronic degrees of freedom by measuring, in addition, emitted electrons in coincidence with the fragment ions. Recent advances in the generation of more intense, few-femtosecond to sub-femtosecond x-ray pulses (*42–44*) further facilitate rapid and complete charge-up and fragmentation of larger systems, and thus put the long-awaited goal to image evolving wave packets in complex polyatomic systems within reach of the method presented in this work.

References and Notes

1. M. Vojta, Quantum phase transitions. *Rep. Prog. Phys.* **66** (12), 2069 (2003), doi:10.1088/0034-4885/66/12/R01, https://dx.doi.org/10.1088/0034-4885/66/12/R01.

2. S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press), 2 ed. (2011).

3. M. Benoit, D. Marx, M. Parrinello, Tunnelling and zero-point motion in high-pressure ice. *Nat* **392** (6673), 258–261 (1998), doi:10.1038/32609.

4. S. Schaack, *et al.*, When Quantum Fluctuations Meet Structural Instabilities: The Isotope- and Pressure-Induced Phase Transition in the Quantum Paraelectric NaOH. *Phys. Rev. Lett.* **131** (12), 126101 (2023), doi:10.1103/PhysRevLett.131.126101.

5. T. E. Markland, *et al.*, Quantum fluctuations can promote or inhibit glass formation. *Nat. Phys.* **7** (2), 134–137 (2011), doi:10.1038/nphys1865.

6. I. Errea, *et al.*, Quantum crystal structure in the 250-kelvin superconducting lanthanum hydride. *Nature* **578**, 66 – 69 (2019).

7. Y. Zheng, *et al.*, Nuclear Quantum Effects on the Charge-Density Wave Transition in NbX₂ (X = S, Se). *Nano Lett.* **22** (5), 1858–1865 (2022), doi:10.1021/acs.nanolett.1c04015.

8. X.-Z. Li, B. Walker, A. Michaelides, Quantum Nature of the Hydrogen Bond. *Proc. Natl. Acad. Sci. U.S.A.* **108** (16), 6369–6373 (2011), doi:10.1073/pnas.1016653108.

9. B. Csorba, P. Szabó, S. Góger, G. Lendvay, The Role of Zero-Point Vibration and Reactant Attraction in Exothermic Bimolecular Reactions with Submerged Potential Barriers: Theoretical Studies of the R + HBr $\rightarrow$ RH + Br (R = CH_3, HO) Systems. *J. Phys. Chem. A* **125** (38), 8386–8396 (2021), doi:10.1021/acs.jpca.1c05839, https://doi.org/10.1021/acs.jpca.1c05839.

10. R. S. Mulliken, The Band Spectrum of Boron Monoxide. *Nature* **114** (2862), 349–350 (1924), doi:10.1038/114349a0.

11. D. A. McQuarrie, J. D. Simon, *Physical Chemistry: A Molecular Approach* (University Science Books, Sausalito, Calif.) (1997).

12. Z. Vager, R. Naaman, E. P. Kanter, Coulomb Explosion Imaging of Small Molecules. *Science* **244** (4903), 426–431 (1989), doi:10.1126/science.244.4903.426, http://www.sciencemag.org/cgi/doi/10.1126/science.244.4903.426.

13. H. Stapelfeldt, E. Constant, H. Sakai, P. B. Corkum, Time-Resolved Coulomb Explosion Imaging: A Method to Measure Structure and Dynamics of Molecular Nuclear Wave Packets. *Phys. Rev. A* **58** (1), 426 (1998), doi:10.1103/PhysRevA.58.426.

14. F. Légaré, *et al.*, Laser Coulomb-explosion imaging of small molecules. *Phys. Rev. A* **71** (1), 013415 (2005), doi:10.1103/PhysRevA.71.013415, http://link.aps.org/doi/10.1103/PhysRevA.71.013415.

15. A. Hishikawa, A. Matsuda, M. Fushitani, E. J. Takahashi, Visualizing Recurrently Migrating Hydrogen in Acetylene Dication by Intense Ultrashort Laser Pulses. *Phys. Rev. Lett.* **99** (25), 258302 (2007), doi:10.1103/PhysRevLett.99.258302, http://link.aps.org/doi/10.1103/PhysRevLett.99.258302.

16. C. B. Madsen, *et al.*, A combined experimental and theoretical study on realizing and using laser controlled torsion of molecules. *J. Chem. Phys.* **130** (23), 234310 (2009), doi:10.1063/1.3149789, https://pubs.aip.org/aip/jcp/article/130/23/234310/924793/A-combined-experimental-and-theoretical-study-on.

17. T. Endo, *et al.*, Capturing roaming molecular fragments in real time. *Science* **370** (6520), 1072–1077 (2020), doi:10.1126/science.abc2960, https://www.sciencemag.org/lookup/doi/10.1126/science.abc2960.

18. S. Bhattacharyya, *et al.*, Strong-Field-Induced Coulomb Explosion Imaging of Tribromomethane. *J. Phys. Chem. Lett.* **13**, 5845–5853 (2022), doi:10.1021/acs.jpclett.2c01007.

19. A. J. Howard, *et al.*, Filming Enhanced Ionization in an Ultrafast Triatomic Slingshot. *Commun. Chem.* **6** (1), 81 (2023), doi:10.1038/s42004-023-00882-w.

20. H. V. S. Lam, *et al.*, Differentiating Three-Dimensional Molecular Structures Using Laser-Induced Coulomb Explosion Imaging. *Phys. Rev. Lett.* **132**, 123201 (2024), doi:10.1103/PhysRevLett.132.123201, https://link.aps.org/doi/10.1103/PhysRevLett.132.123201.

21. Y. Muramatsu, *et al.*, Direct Probe of the Bent and Linear Geometries of the Core-Excited Renner-Teller Pair States by Means of the Triple-Ion-Coincidence Momentum Imaging Technique. *Phys. Rev. Lett.* **88**, 133002 (2002).

22. N. Saito, *et al.*, Deformation, nuclear motion and fragmentation of core-excited CO_2 probed by multiple-ion coincidence momentum imaging. *J. Electron Spectrosc. Relat. Phenom.* **141** (2), 183–193 (2004), doi:https://doi.org/10.1016/j.elspec.2004.06.007, https://www.sciencedirect.com/science/article/pii/S0368204804003366.

23. U. Ablikim, *et al.*, Identification of absolute geometries of cis and trans molecular isomers by Coulomb Explosion Imaging. *Sci. Rep.* **6**, 38202 (2016), doi:10.1038/srep38202, http://www.nature.com/articles/srep38202.

24. M. Pitzer, *et al.*, Absolute Configuration from Different Multifragmentation Pathways in Light-Induced Coulomb Explosion Imaging. *ChemPhysChem* **17** (16), 2465–2472 (2016), doi:10.1002/cphc.201501118, http://doi.wiley.com/10.1002/cphc.201501118.

25. Y. H. Jiang, *et al.*, Ultrafast Extreme Ultraviolet Induced Isomerization of Acetylene Cations. *Phys. Rev. Lett.* **105** (26), 263002 (2010), doi:10.1103/PhysRevLett.105.263002, http://link.aps.org/doi/10.1103/PhysRevLett.105.263002.

26. K. Amini, *et al.*, Alignment, orientation, and Coulomb explosion of difluoroiodobenzene studied with the pixel imaging mass spectrometry (PImMS) camera. *J. Chem. Phys.* **147** (1), 013933 (2017), doi:10.1063/1.4982220.

27. F. Allum, *et al.*, Coulomb explosion imaging of CH_3I and CH_2ClI photodissociation dynamics. *J. Chem. Phys.* **149** (20), 204313 (2018), doi:10.1063/1.5041381.

28. E. Kukk, *et al.*, Coulomb implosion of tetrabromothiophene observed under multiphoton ionization by free-electron-laser soft-x-ray pulses. *Phys. Rev. A* **99** (2), 023411 (2019), doi:10.1103/PhysRevA.99.023411, https://link.aps.org/doi/10.1103/PhysRevA.99.023411.

29. T. Jahnke, *et al.*, Inner-Shell-Ionization-Induced Femtosecond Structural Dynamics of Water Molecules Imaged at an X-ray Free-Electron Laser. *Phys. Rev. X* **11**, 041044 (2021).

30. R. Boll, *et al.*, X-Ray Multiphoton-Induced Coulomb Explosion Images Complex Single Molecules. *Nat. Phys.* **18**, 423–428 (2022), doi:10.1038/s41567-022-01507-0.

31. T. Kitamura, T. Nishide, H. Shiromaru, Y. Achiba, N. Kobayashi, Direct observation of "dynamic" chirality by Coulomb explosion imaging. *J. Chem. Phys.* **115** (1), 5–6 (2001), doi:10.1063/1.1383793.

32. J. Ullrich, *et al.*, Recoil-ion and electron momentum spectroscopy: reaction-microscopes. *Rep. Prog. Phys.* **66** (9), 1463 (2003), http://iopscience.iop.org/0034-4885/66/9/203.

33. T. Jahnke, *et al.*, Multicoincidence studies of photo and Auger electrons from fixed-in-space molecules using the COLTRIMS technique. *J. Electron Spectrosc. Relat. Phenom.* **141** (2), 229–238 (2004), doi:https://doi.org/10.1016/j.elspec.2004.06.010, https://www.sciencedirect.com/science/article/pii/S0368204804003408.

34. C. Vallance, D. Heathcote, J. W. L. Lee, Covariance-Map Imaging: A Powerful Tool for Chemical Dynamics Studies. *J. Phys. Chem. A* **125** (5), 1117–1133 (2021), doi:10.1021/acs.jpca.0c10038.

35. X. Li, *et al.*, Ultrafast Coulomb Explosion Imaging of Molecules and Molecular Clusters. *Chinese Phys. B* **31** (10), 103304 (2022), doi:10.1088/1674-1056/ac89df.

36. Z. Jurek, S.-K. Son, B. Ziaja, R. Santra, XMDYN and XATOM: versatile simulation tools for quantitative modeling of X-ray free-electron laser induced dynamics of matter. *J. Appl. Cryst.* **49** (3), 1048–1056 (2016), doi:10.1107/S1600576716006014, http://scripts.iucr.org/cgi-bin/paper?zd5003.

37. A. C. T. van Duin, S. Dasgupta, F. Lorant, W. A. Goddard, ReaxFF: A Reactive Force Field for Hydrocarbons. *J. Phys. Chem. A* **105** (41), 9396–9409 (2001), doi:10.1021/jp004368u.

38. M. F. Russo, A. C. T. van Duin, Atomistic-Scale Simulations of Chemical Reactions: Bridging from Quantum Chemistry to Engineering. *Nucl. Instrum. Methods Phys. Res., Sect. B* **269** (14), 1549–1554 (2011), doi:10.1016/j.nimb.2010.12.053.

39. H. M. Aktulga, S. A. Pandit, A. C. T. van Duin, A. Y. Grama, Reactive Molecular Dynamics: Numerical Methods and Algorithmic Techniques. *SIAM J. Sci. Comput.* **34** (1), C1–C23 (2012), doi:10.1137/100808599.

40. B. Richard, J. M. Schäfer, Z. Jurek, R. Santra, L. Inhester, Statistical Analysis of Correlations in the X-Ray Induced Coulomb Explosion of Iodopyridine. *J. Phys. B: At. Mol. Opt. Phys.* **54** (19), 194001 (2021), doi:10.1088/1361-6455/ac2f4f.

41. I. T. Jolliffe, J. Cadima, Principal Component Analysis: A Review and Recent Developments. *Phil. trans. R. Soc. A* **374** (2065), 20150202 (2016), doi:10.1098/rsta.2015.0202.

42. S. Huang, *et al.*, Generating Single-Spike Hard X-Ray Pulses with Nonlinear Bunch Compression in Free-Electron Lasers. *Phys. Rev. Lett.* **119**, 154801 (2017), doi:10.1103/PhysRevLett.119.154801, https://link.aps.org/doi/10.1103/PhysRevLett.119.154801.

43. J. Duris, *et al.*, Tunable isolated attosecond X-ray pulses with gigawatt peak power from a free-electron laser. *Nat. Photonics* **14** (1), 30–36 (2020), doi:10.1038/s41566-019-0549-5, https://www.nature.com/articles/s41566-019-0549-5.

44. A. Trebushinin, *et al.*, Experimental Demonstration of Attoseconds-at-Harmonics at the SASE3 Undulator of the European XFEL. *Photonics* **10** (2), 131 (2023), doi:10.3390/photonics10020131, https://www.mdpi.com/2304-6732/10/2/131.

45. R. Boll, Tracing electron rearrangement upon absorption of very intense X-ray pulses, European XFEL Database (2019), doi:https://doi.org/10.22003/xfel.eu-data-002159-00.

46. B. Richard, S. Banerjee, Z. Jurek, R. Santra, L. Inhester, Simulation and reconstruction of the final momenta generated in the Coulomb explosion of iodopyridine, University of Ham-

burg research data repository (FDR-UHH) (2025), doi:https://doi.org/10.25592/uhhfdm.16697, https://doi.org/10.25592/uhhfdm.16697.

47. O. Jagutzki, *et al.*, Multiple hit readout of a microchannel plate detector with a three-layer delay-line anode. *IEEE Trans. Nucl. Sci.* **49** (5), 2477–2483 (2002).

48. D. Toffoli, P. Decleva, Photoelectron angular distributions beyond the dipole approximation: a computational study on the N2 molecule. *J. Phys. B* **39**, 2681–2691 (2006), doi:10.1088/0953-4075/39/12/005.

49. D. R. Miller, G. Scoles, Atomic and molecular beam methods. *Atomic and Molecular Beam Methods* **1**, 14 (1988).

50. M. Hillenkamp, S. Keinan, U. Even, Condensation limited cooling in supersonic expansions. *J. Chem. Phys.* **118** (19), 8699–8705 (2003), doi:10.1063/1.1568331, https://doi.org/10.1063/1.1568331.

51. S. Monti, *et al.*, Exploring the Conformational and Reactive Dynamics of Biomolecules in Solution Using an Extended Version of the Glycine Reactive Force Field. *Phys. Chem. Chem. Phys.* **15** (36), 15062–15077 (2013), doi:10.1039/C3CP51931G.

52. Q. Mao, Y. Ren, K. H. Luo, A. C. T. van Duin, Dynamics and Kinetics of Reversible Homo-Molecular Dimerization of Polycyclic Aromatic Hydrocarbons. *J. Chem. Phys.* **147** (24), 244305 (2017), doi:10.1063/1.5000534.

53. M. V. Fedkin, *et al.*, Development of the ReaxFF Methodology for Electrolyte–Water Systems. *J. Phys. Chem. A* **123** (10), 2125–2141 (2019), doi:10.1021/acs.jpca.8b10453.

54. M. C. Kaymak, *et al.*, JAX-ReaxFF: A Gradient-Based Framework for Fast Optimization of Reactive Force Fields. *J. Chem. Theory Comput.* **18** (9), 5181–5194 (2022), doi:10.1021/acs.jctc.2c00363.

55. V. Sortur, J. Yenagi, J. Tonannavar, Vibrational assignments of 2-iodopyridine. *Spectrochim. Acta A* **69** (2), 604–611 (2008), doi:https://doi.org/10.1016/j.saa.2007.05.010, https://www.sciencedirect.com/science/article/pii/S138614250700251X.

56. Y. Hao, L. Inhester, K. Hanasaki, S.-K. Son, R. Santra, Efficient Electronic Structure Calculation for Molecular Ionization Dynamics at High X-ray Intensity. *Struct. Dyn.* **2** (4), 041707 (2015), doi:http://dx.doi.org/10.1063/1.4919794.

57. J. S. Binkley, J. A. Pople, W. J. Hehre, Self-consistent molecular orbital methods. 21. Small split-valence basis sets for first-row elements. *J. Am. Chem. Soc.* **102**, 939–947 (1980), doi:10.1021/ja00523a008.

58. K. D. Dobbs, W. J. Hehre, Molecular orbital theory of the properties of inorganic and organometallic compounds 4. Extended basis sets for third- and fourth-row, main-group elements. *J. Comput. Chem.* **7**, 359–378 (1986), doi:10.1002/jcc.540070313.

59. K. Carlsson, *et al.*, NearestNeighbors.jl: v0.4.16, Zenodo (2023), doi:10.5281/zenodo.10435686, `https://zenodo.org/records/10435686`.

60. M. Taboga, *Lectures on Probability Theory and Mathematical Statistics* (CreateSpace Independent Publishing Platform), 3 ed. (2017).

61. J. Day, H. Zhou, OnlineStats.Jl: A Julia Package for Statistics on Data Streams. *Journal of Open Source Software* **5** (46), 1816 (2020), doi:10.21105/joss.01816.

62. F. W. Byron, R. W. Fuller, *Mathematics of Classical and Quantum Physics* (Courier Corporation) (2012).

63. P. K. Mogensen, *et al.*, Optim.jl: v1.7.8, Zenodo (2023), doi:10.5281/zenodo.8254057, `https://zenodo.org/records/8418604`.

64. M. Innes, Don't Unroll Adjoint: Differentiating SSA-Form Programs (2019), doi:10.48550/arXiv.1810.07951.

65. F. White, *et al.*, ChainRules.jl: v1.58.0, Zenodo (2023), doi:10.5281/zenodo.8316272, `https://zenodo.org/records/10100624`.

Acknowledgments

L.I. acknowledges discussion with Sebastian Trippel and Ivo Vinklárek about supersonic molecular beams.

Funding: B.R., R.S., M.I., and L.I. acknowledge support from DESY (Hamburg, Germany), a member of the Helmholtz Association HGF. B.R., R.S., M.I., K.K., and L.I. acknowledge support from the Cluster of Excellence 'CUI: Advanced Imaging of Matter' of the Deutsche Forschungsgemeinschaft (DFG) - EXC 2056 - project ID 390715994, and K.K. in addition acknowledges the program "Sachbeihilfe" project ID 497431350 (KU 4184/1-1). F.T. acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project 509471550, Emmy Noether Programme. The reaction microscope and its implementation at SQS were funded by the BMBF. R.D., K.F., S.G., M.H., M.I., and M.S.S. acknowledge support from Deutsche Forschungsgemeinschaft (DFG) via Sonderforschungsbereich 1319 (ELCH). G.K. acknowledges funding by Helmholtz Forschungsakademie Hessen für FAIR (HFHF). We acknowledge European XFEL in Schenefeld, Germany, for the provision of XFEL beam time at the SQS instrument and would like to thank the staff for their assistance. D.R. and A.R. are supported by the Chemical Sciences, Geosciences, and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy, Grant No. DE-FG02-86ER13491. M.I. acknowledges support by the Bundesministerium für Bildung und Forschung (BMBF) within grant 13K22CHA. V.M. acknowledges funding from the Volkswagen foundation for a Peter Paul Ewald-fellowship.

Author contributions: T.J. originated the project's idea and analyzed the experimental data, L.I. initiated the advanced analysis methodology. R.B. was principal investigator of the experiment. B.R. wrote the code, ran the simulations, and performed the fitting of the data with input from J.M.S., Z.J., R.S., and L.I. S.B. and Z.J. performed the adjustments of the force-field parameters. The experiment was carried out by R.B., J.M.S., K.F., M.S.S., S.E., B.E., M.I., K.K., X.L., T.M., C.P., D.E.R., D.R., J.M.S., L.P.H.S., F.T., P.W., M.M., and T.J. after the experiment was previously prepared and commissioned by G.K., N.A., T.M.B., A.D.F., S.G., P.G., M.H., M. Ki., M. Ku., S.M., N.M., J.M., V.M., Y.O., A.P., N.R., J.R., I.S., L.P.H.S., D.T., R.W. and P.Z. The REMI experimental chamber was designed and built by G.K., T.J., M.S.S., R.D., and L.P.H.S., with contributions from

the whole University of Frankfurt group. B.R., T.J., R.B., R.S., and L.I. wrote the manuscript after in-depth discussions with M.M. and D.R. with input from all the authors.

Competing interests: There are no competing interests to declare.

Data and materials availability: The data recorded during the experiment are available in the European XFEL database, under proposal no. 002159 (*45*). The simulation data and the result of the reconstructions are available at the research data repository of the University of Hamburg (*46*).

Supplementary materials

Supplementary Text
Figures S1 to S18
Tables S1 to S2
References *(47-65)*

Supplementary Materials for

Imaging collective quantum fluctuations of the structure of a complex molecule

Benoît Richard, Rebecca Boll[#], *et al.*, Ludger Inhester[$], Till Jahnke[†]

[#]Corresponding author. Email: rebecca.boll@xfel.eu

[$]Corresponding author. Email: ludger.inhester@cfel.de

[†]Corresponding author. Email: till.jahnke@xfel.eu

This PDF file includes:

S1 Experiment

The Coulomb explosion imaging was performed at the Small Quantum Systems (SQS) scientific instrument of the European XFEL. The momenta of the ionic fragments of the exploded molecules were recorded in coincidence employing the COLTRIMS (cold target recoil ion momentum spectroscopy) reaction microscope (*32, 33*) user end-station (SQS-REMI) available at SQS. In the following, we provide details on the properties of the x-rays, the SQS-REMI, and the offline analysis of the measured data.

S1.1 Properties of the x-ray pulses

The pulse pattern available at the EuXFEL consists of pulse trains generated at a repetition rate of 10 Hz with a burst length of several 100 μs. Within a single train, electron pulses are provided at up to 4.5 MHz repetition rate. In the operation mode employed during this experiment, the electron pulses had a repetition rate of 1.1 MHz. In order to comply with the requirements imposed by the flight times of the fragment ions, we used every 6th electron pulse to generate x-rays, yielding a spacing of 5.5 μs between adjacent light pulses. The exact burst length available to the user depends on several operating constraints of the machine. We obtained between 250-570 x-ray pulses per second in our experiment. We employed a photon energy of $h\nu = 2$ keV for triggering the Coulomb explosion. The single-shot pulse energy was 0.8 mJ on target. The focus size at the target was in the range of 1.4 μm (vertical and horizontal) as estimated from calculations considering the beamline's focusing scheme. Based on the electron bunch charge in the accelerator of 250 pC, the x-ray pulse duration was estimated to be shorter than 25 fs.

S1.2 Experimental setup

The target molecules were prepared as a supersonic gas jet consisting of iodopyridine vapor, which was generated by heating a reservoir containing the liquid substance to a temperature of 60 °C and expanding a mixture of the vapor and helium carrier gas (at a stagnation pressure of 500 mbar) through a nozzle with a diameter of 200 μm. The gas jet was collimated by a set of three skimmers and adjustable slits. The adjustable slits were employed to cut down the gas jet such that coincidence-detection conditions were retained. When operating the gas jet, the pressure inside the

main chamber was retained at 1×10^{-11} mbar. The gas jet was crossed with the x-rays at right angle in the focus of the beamline, yielding a well-defined interaction region. Ions that were generated by the x-rays and subsequent Auger-Meitner decay(s) were guided by a strong electric field towards a position- and time-sensitive microchannel plate detector with an active area of 120 mm diameter and hexagonal delay-line position readout (*47*). From the measured flight times and the positions of impact of the ions, we reconstructed the momentum vector of each detected ion in an offline analysis. The flight time also provides information on the ions' mass-over-charge ratios. In more detail, the spectrometer consisted of an acceleration region with a length of 86 mm and an electric-field strength of $E = 423$ V/cm at the interaction spot of the x-rays and the gas jet. A second region with a length of 164 mm connecting the extraction region with the microchannel plate detector was used as a drift region. The spectrometer was designed such that it provided a compromise between achievable momentum resolution and mass-over-charge separation of different ion species.

S1.3 Momentum resolution of the employed spectrometer

To determine an upper limit of the momentum resolution, we did a calibration measurement using K-shell ionization of N_2 molecules under the same experimental conditions as the main measurement. After the absorption of a single photon with $h\nu = 2000$ eV, the molecule undergoes a subsequent Auger decay and fragments in most cases into two singly charged nitrogen ions. We measured these two ions in coincidence and inspected the sum momentum of the two. The measured distributions are shown in Figs. S1A and S1B after correcting for the mean offset velocity of the supersonic gas jet and the linear momentum of the absorbed photon (which basically centers the measured distribution at zero). After these corrections, the sum momentum is dominated by the recoil of the two emitted electrons, namely the photoelectron imposing a recoil of $p_e = 10.8$ a.u. and the Auger electron providing a kick of approximately $p_A = 5$ a.u. While the Auger electron is emitted in a good approximation isotropically in the laboratory frame, the photoelectron has a distinct angular distribution corresponding to an anisotropy factor of $\beta = 2$ (*48*). The latter yields the double-peak structure visible in Figs. S1B, S1D, and S1F, as the polarization vector of the linearly polarized photons pointed along the y direction in our measurement.

We performed a simple modeling of the sum momentum by including the recoil of the two electrons and considering their angular emission distributions. Furthermore, we added a Gaussian-

shaped finite momentum resolution to this simulation. By using the width of that Gaussian function as a fitting parameter, we reproduced the measured sum-momentum distribution by our simulation as depicted in Figs. S1C and S1D. For a quantitative comparison, we provide the distribution of the x and the y components of the measured and simulated sum momenta (both normalized) in Figs. S1E and S1F.

From this procedure, we obtained a momentum resolution of 3.2 a.u. for a measured N^+ ion. The momentum resolution in a COLTRIMS measurement is independent of the energy of a measured particle of a given species. However, it depends in a non-trivial way on the particle's mass m and charge q. For the given spectrometer, the following relations apply: The resolution of the momentum component p_z pointing along the symmetry axis of the spectrometer (i.e., the component that is determined by the time-of-flight measurement) is independent of the mass of the detected particle but scales linearly with q. The resolution of the other two momentum components $p_{x,y}$ (being determined by the measurement of the position of impact) scales with $\sqrt{mq}$. The overall momentum resolution therefore depends on the emission direction of the measured particle in the laboratory frame. An estimation of an upper limit of the resolution, i.e., the worst-case scenario, may be obtained by considering only the component yielding the highest error. In this case, the resolution for protons is 3 a.u., while the momentum resolution for a C^{2+} ion is 6.4 a.u. and an I^{4+} is close to 20 a.u. To put these worst-case values in relation, Fig. S12 shows the unnormalized momenta. The out-of-plane component observed for hydrogens reaches up to almost 100 a.u., and for doubly charged carbon up to 300 a.u., and I^{4+} momenta up to 1000 a.u. are considered in this paper (see subsection S1.6).

S1.4 Estimation of the degree of vibrational excitation

The frequency of the lowest vibrational mode in the molecule is $\nu = 5.3\,\mathrm{THz}$. This frequency corresponds to a temperature of $h\nu/k_B = 256\,\mathrm{K}$ which implies that, at room temperature, there is to some degree thermal excitation in this vibrational mode. The temperature of the molecules in the gas jet is, however, considerably lower than before the expansion. The temperature of the gas can be estimated from (*49*)

$$\frac{T}{T_0} = \left(1 + \frac{\gamma - 1}{2} M^2\right)^{-1}, \tag{S1}$$

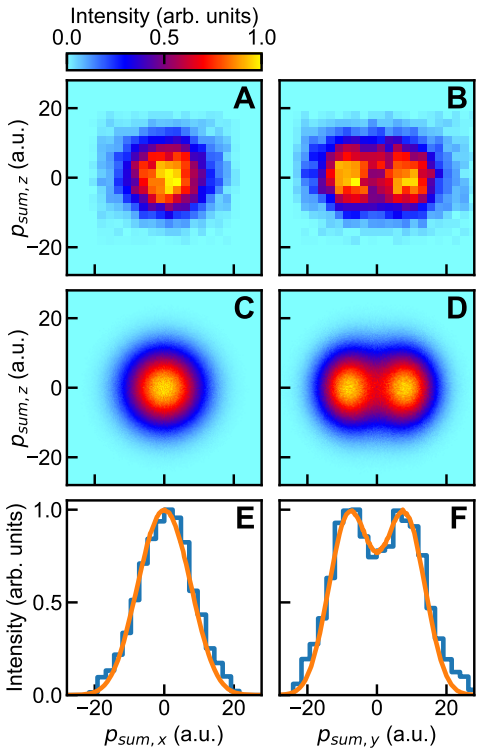

Figure S1: **Measured and simulated sum momenta to determine the experimental momentum resolution.** A, B. Measured sum momentum of the breakup of N_2 into N^+/N^+ ions after K-shell ionization and subsequent Auger decay. C, D. Corresponding results from simple modeling incorporating a momentum resolution of 3.2 a.u. E, F. *x* and *y* components of the sum momentum for quantitative comparison of the measured (blue line) and modeled results (orange line). The polarization vector of the linearly polarized photons is aligned along the *y* axis.

where T/T_0 is the ratio of temperature after and before the nozzle, M is the Mach number after the nozzle, and $\gamma = c_p/c_V$ is the ratio of heat capacities (also known as adiabatic index). Assuming $M = 1$, which is reached at the nozzle exit, $c_p/c_V = 5/3$ (ideal gas), and $T_0 = 333\,\mathrm{K}$, an upper estimate for the gas-jet temperature is $T = 250\,\mathrm{K}$. At this temperature, we estimate the relative occupation of the first excited vibrational state to be $n_1 = e^{-h\nu/k_BT}/(1 + e^{-h\nu/k_BT}) \simeq 0.32$ (assuming a canonical ensemble of two-level entities). Accordingly, we expect that even for this lowest-frequency mode, the fluctuations are dominated by quantum effects. For all the other modes, frequencies are considerably higher and, thus, excited-state populations are below 0.05, showing that any structural variations in the ensemble of molecules are predominantly caused by GSFs and not by thermal fluctuations.

Our estimate is deliberately chosen very conservative in order to show that even under these conditions the target can be expected to be to the largest extent in the vibrational ground state. We would like to note that typical helium-seeded molecular beams of this kind are actually expected to have much smaller internal temperatures (*50*). If we assume a very low (i.e., bad) speed ratio S of, e.g., only $S = 5$, the molecular-beam temperature after the expansion equals approximately $T = 40\,\mathrm{K}$.

S1.5 Estimating the amount of contamination with dimers

In an experiment employing a supersonic molecular beam, contamination with molecular dimers (or even larger clusters) can occur. In the following, we provide several indications that such a contamination with dimers is (if present at all) negligible given the contrast of the features observed in the experiment and presented in Fig. 1.

As a first evidence, we identified in Fig. S2 a faint peak originating from the intact singly charged parent ion, $C_5H_4NI^+$, but no traces of a corresponding dimer peak $[C_5H_4NI]_2^+$ or any larger clusters. While dimers are expected to break up under XFEL irradiation, the fact that we observe singly charged parent ions would suggest that dimer cations could also be created, if dimers are present in the gas jet.

Second, Fig. S2 shows the unfiltered raw data with the measured ion flight time on the x axis and the y component of its position of impact on the detector on the y axis. The confined sharp peak at a flight time of 310 ns belongs to the aforementioned singly charged parent ion; the large

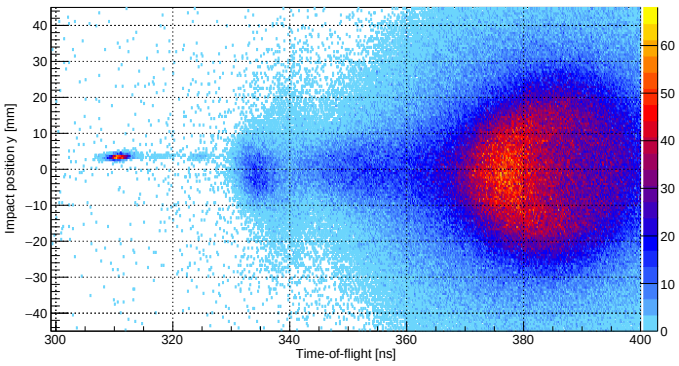

Figure S2: **Raw data showing the region of singly charged parent ions.**

circular feature on the right (centered at at flight time of 383 ns) shows the measured protons. The parent ion appears at these very short flight times due to the pulse train pattern of the EuXFEL. In principle its flight time is approximately 5600 ns, but due to the chosen repetition rate (see section S1.1) this flight time is beyond the recording window such that the parent ion generated in a previous shot appears at short flight times (i.e., the flight times are "wrapping around"). The parent ion is displaced slightly in the y direction. This is caused by the mean velocity of the supersonic jet (it propagates along the y direction) and provides further evidence that the feature found at this expected flight time corresponds to the singly charged parent ion. The parent ion peak is very narrow, which indicates that the ion has a very small momentum spread. A fragmentation of dimers would lead to parent ions with a significantly wider momentum distribution, which we do not observe. In addition, we provide a coincidence map (PIPICO) of the raw data in Fig. S3 covering the same flight-time region as in Fig. S2. The strongest feature can be again attributed to measured protons. The more faint "vertical" stripe belongs to the singly charged parent ion. Cases of coincidences of two singly charged parent ions appear as the very confined spot at the bottom of the vertical feature (located at 310 ns/310 ns), the rest of the stripe corresponds to a coincidence between the parent ion and a proton. Given the narrow width of this stripe, this indicates, as well, that there is almost no breakup energy present. This suggests that we mainly see false coincidences between an intact parent ion and some random proton. In addition, please note the logarithmic color scale indicating how small this contribution is after all.

Third, we investigated coincident detection of two iodine ions, as a Coulomb exploding dimer

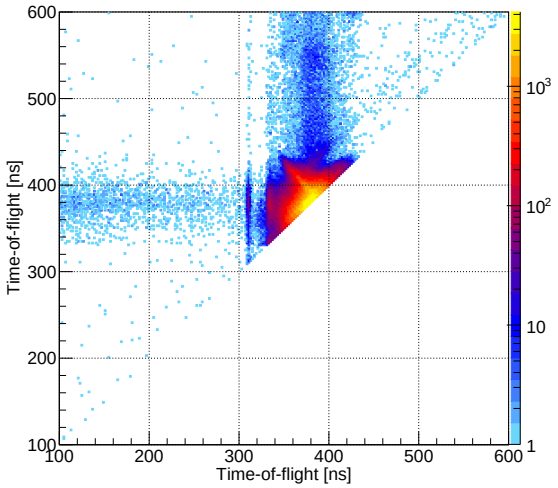

Figure S3: **Coincidence map of the measured flight times.** The histogram depicts the same region of flight times as shown in Fig. S2.

would probably generate these. We extracted a ratio of approximately 3% of I^+/I^+ coincidence pairs as compared to single I^+ ions. In combination with our above findings, we assume that this contribution basically reflects our amount of false coincidences. Either way, it provides a hint for an upper limit of a possible contamination with dimers or larger clusters.

S1.6 Data analysis

In a COLTRIMS measurement, the initial momentum vectors of the detected particles are obtained in an offline analysis from the measured flight times and positions of impact of the particles. If all particles of a molecular fragmentation are detected in coincidence, possible background or false coincidence events can be detected in the data set by inspecting the sum of the momenta of the fragments (*33*). In the present case, this method is not applicable, as only a subset of the up to 11 generated fragment ions were measured. However, even in this case, a certain amount of background events can be filtered out by restricting the plotted data to cases where all measured ions have a momentum within a sensible range. Thus we considered only events for which the individual ion momenta were in the range of $\|\mathbf{p}_\mathrm{H}\| < 130$ a.u., $\|\mathbf{p}_\mathrm{C}\| < 500$ a.u., $\|\mathbf{p}_\mathrm{N}\| < 650$ a.u., and $\|\mathbf{p}_\mathrm{I}\| < 1000$ a.u. In addition, a rather relaxed restriction on the sum of the momenta of the measured fragments of $\|\mathbf{p}_\mathrm{sum}\| < 1500$ a.u. was applied in order to further suppress background

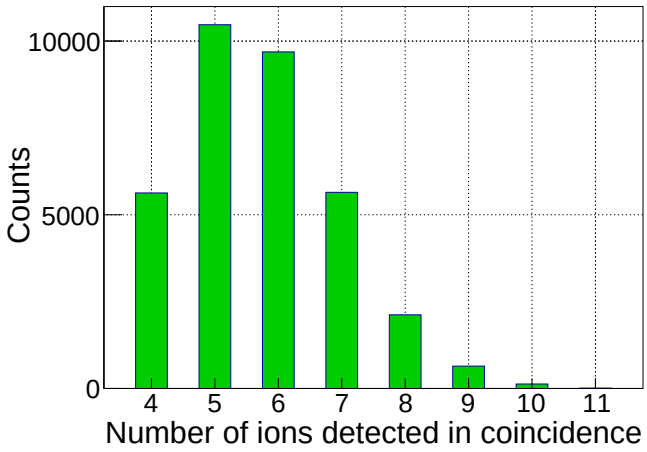

Figure S4: **Ion coincidences.** Histogram of the number of ions measured in coincidence in the experiment for cases where an I^{4+}, an N^{2+}, at least one C^{2+}, and at least one proton were detected.

and false coincidences. Ions with a very similar mass-over-charge-ratio may partially overlap in their time-of-flight distributions. We therefore needed to discard a part of the N^{2+} ion momentum distribution with $\mathbf{p}_{\mathrm{lab,z}} > 100$ a.u., which corresponds to a subset of nitrogen ions being emitted in the laboratory frame towards the ion detector with high kinetic energy.

S1.7 Coincident ion detection

In the experiment, several ions are detected in coincidence. Figure S4 shows a histogram of the number of in-coincidence detected ions where at least an I^{4+}, an N^{2+}, one C^{2+}, and one proton were measured. In this histogram, only those coincidence events are considered that fulfilled the momentum filtering conditions described in subsection S1.6. In total our data set consists of 34336 events. In most cases, 5 or 6 ions were detected in coincidence. For the selected charge states, we only obtained ~ 10 full coincident measurements in which all 11 ions were detected.

S2 Correlation between proton and carbon ion momenta

In Figs. 1G and 1H, we demonstrated the correlated behavior of the H3 proton and the carbon ion momenta by selecting different ranges of the H3 out-of-plane momentum. Similar momentum correlations are also present between other atoms in the measured data. Figure S2 shows an overview of the extracted correlations employing the iodine-nitrogen recoil frame. In each panel, we plotted

the mean value of the emission angle Θ of a specific carbon ion against the normalized out-of-plane momentum of a specific proton as indicated in the panel titles. We obtained the mean value of Θ by fitting a sum of five two-dimensional Gaussian distributions (one for each carbon ion) to the angular emission distribution histogram (similar to the one shown in Fig. 1F) after filtering the data set to the given range of the normalized out-of-plane momentum of the corresponding proton. There are clear trends, such as cases where the out-of-plane emission of the protons and carbon ions occurs in the same direction (e.g., for H6/C6) and other cases where the emission occurs into opposite directions (e.g., H3/C6). The correlated behavior of the carbon ions and the H3 proton indicated in Fig. 1E is depicted in the bottom row of Fig. S2 by the four panels. We omitted the C2 ion from this presentation due to the large spread of its angular emission distribution.

S3 Simulation

S3.1 XMDYN model

In order to simulate the Coulomb explosion, we employ the XMDYN simulation toolkit (revision 0.0-62-g4f4f5c8) (*36*). In XMDYN, the multiphoton multiple-ionization dynamics in the molecule are modelled through a quantum-classical hybrid approach. Neutral atoms, atomic ions, and free electrons are considered classical point particles; bound electrons are assigned to atomic orbitals that are determined via electronic structure calculations for each atom individually via the XATOM toolkit (*36*). When an electron is emitted, via either direct photoionization or Auger-Meitner decay, it is removed from the corresponding orbital, and a new classical free electron is created in the vicinity of the emitting atomic ion. The kinetic energy is computed from the binding energy supplied by XATOM and the classical potential energy relations in the system. The conservation of the total momentum is enforced by imprinting a recoil on the emitting atomic ion. This recoil is the only effect that breaks the planar symmetry of the molecule beyond the GSFs, and causes the out-of-plane momenta shown in Figs. 2E and 2F, where the GSFs are disabled. Atomic cross sections and decay rates enter the ionization dynamics that are modelled by running multiple trajectories in a kinetic Monte-Carlo scheme. Electron transfer processes between different atoms are incorporated via an effective over-the-barrier model (*30*). Along each trajectories, particle positions are propagated using Newton's equations of motion according to the computed forces between the particles.

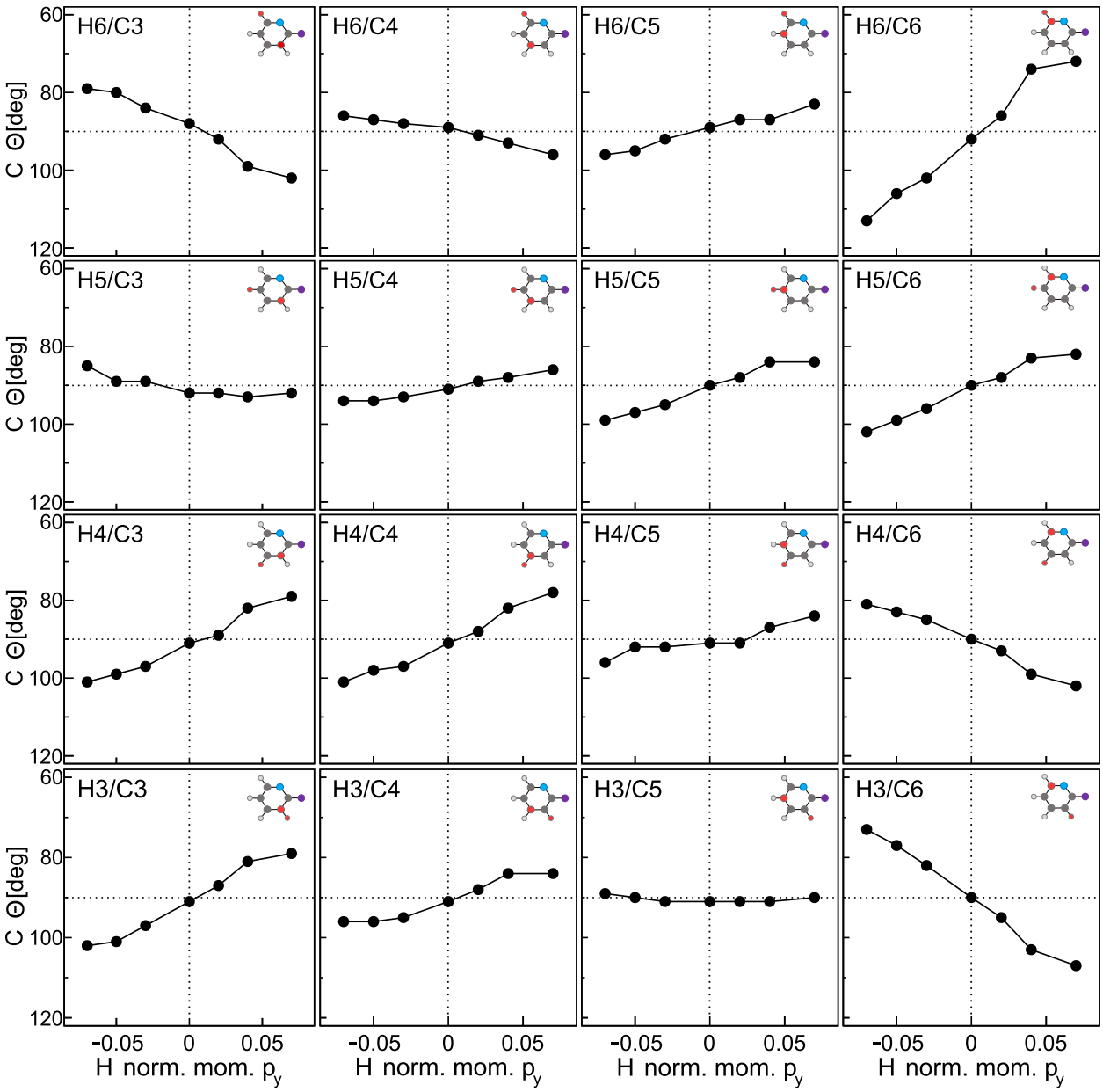

Figure S5: **Correlation among all protons H3 to H6 and carbon ions C3 to C6.** Each panel depicts (similar to Fig. 1H) the dependence between the out-of-plane momentum of a proton normalized relative to the corresponding I^{4+} momentum (x axis) and the emission angle Θ of a carbon ion, as defined in Fig. 1E (y axis). Each row corresponds to one of the protons, from top to bottom H6, H5, H4, and H3. Each column corresponds to one of the carbon ions, from left to right C3, C4, C5, and C6. The small icons indicate, which combination of carbon ion and proton is covered in the corresponding panel. The error of the quantities displayed, as estimated from the fitting, is smaller than the dot size of the data points.

In our earlier modelling of the explosion dynamics for the same system, only Coulomb interactions among the particles were considered (*30*). For the current work, the XMDYN toolkit has been extended to also consider forces due to chemical bonds in the molecule. To that end, the chemical bond potentials are approximated by using the reactive force field (ReaxFF) method (*37, 38*). Specifically, we interface the PuReMD (*39*) implementation of ReaxFF with XMDYN, by integrating the forces provided by the force field into the dynamics computed by XMDYN. During the explosion, the force field is linearly damped proportionally to the charges in the system, reaching zero once 4 electrons have been ionized from the molecule. This simulates the weakening of the chemical bonds as the charging up of the molecule progresses. The choice of this cutoff charge has a minor impact on the results, as discussed in section S3.3.

In order to apply ReaxFF, force-field parameters were determined for the iodopyridine molecule starting from existing parameter tables (*51–53*). With the help of the JAX-ReaxFF framework (*54*), force-field parameters were fitted to existing spectroscopic data (*55*) and to reference data computed with XMOLECULE (*56*), employing the Hartree-Fock method and the 3-21G basis set (*57, 58*). A full table containing the fitted parameters is available in the data repository (*46*).

During the explosion, certain atoms come particularly close to each other. Specifically, this is the case for the iodine atom and the carbon atom C2, as a consequence of the collision dynamics (*40*) (see also the supplementary movie in Ref. (*30*)). To have a more realistic description of the repulsion between these atoms when they come very close to each other, we also incorporated exchange repulsion between carbon and iodine ions using the potential

$$V_{\text{exchange}}(r) = Ae^{-br}, \tag{S2}$$

where r is the distance between the two atoms, $A = 6647.6\,\text{eV}$ and $b = 4.46\,\text{Å}^{-1}$. The values for A and b have been determined by fitting the potential energy curve for an isolated I^+ and C^+ ion computed via XMOLECULE (*56*) employing the Hartree-Fock-Slater approximation.

For the initial states of the trajectories, we draw 61652 samples of positions and momenta from the ground-state Wigner distribution constructed from the ReaxFF force field in the harmonic approximation. To give an impression on the magnitude of the ground-state fluctuations, the real-space distribution of the initial state is visualized in Fig. S6 and the width of the distribution for each atom is reported in Table S1.

Element	Position space	Momentum space
Iodine	0.9 pm	4.4 a.u.
Nitrogen	4.3 pm	6.4 a.u.
Carbon	4.1 pm	6.5 a.u
Hydrogen	10.9 pm	2.8 a.u

Table S1: **Average width (standard deviation) of the Wigner distribution for each element.**

The parameters for the simulations are listed in Table S2.

The number of trajectories, the time step increment, and the number of time steps are numerical parameters that have been chosen such that the simulations converge and behave consistently.

In contrast, the choice of the physical parameters was motivated by various considerations. The photon energy is chosen equal to the experimental value of 2 keV. The pulse is linearly polarized with a random orientation with respect to the molecule, in accordance with the random orientation of the gas-phase molecules in the laboratory frame. The pulse duration is set to 10 fs (full-width at half-maximum) with a Gaussian temporal profile, in agreement with earlier estimations (*30*). We tuned the fluence to 1.5×10^{11} photons/μm^2 to optimize the number of trajectories falling in the I^{4+}/N^{2+} coincidence channel, based on XATOM calculations of single iodine atoms. We note that simulations performed with different pulse durations (1 fs to 30 fs) and different fluences (0.5×10^{11} to 1.5×10^{11} photons/μm^2) show that these two parameters have only a limited impact on the results.

For Figs. 2E and 2F, we generated another set of trajectories without GSFs. To remove the GSFs, all the atoms are initially in the equilibrium ground-state position with zero initial velocity.

S3.2 Charging up of the molecule

Figure S7 shows the simulated average charge as a function of time during the exposure of the molecule to the x-ray pulse, for the I^{4+}/N^{2+} coincidence channel used in the main text. We see that a total charge of more than 11 (on average one charge per atom) is reached about 10 fs after the first ionization. This enormous ionization rate is instrumental for our ability to probe the GSFs and for the validity of our model, as described in Section S3.3.

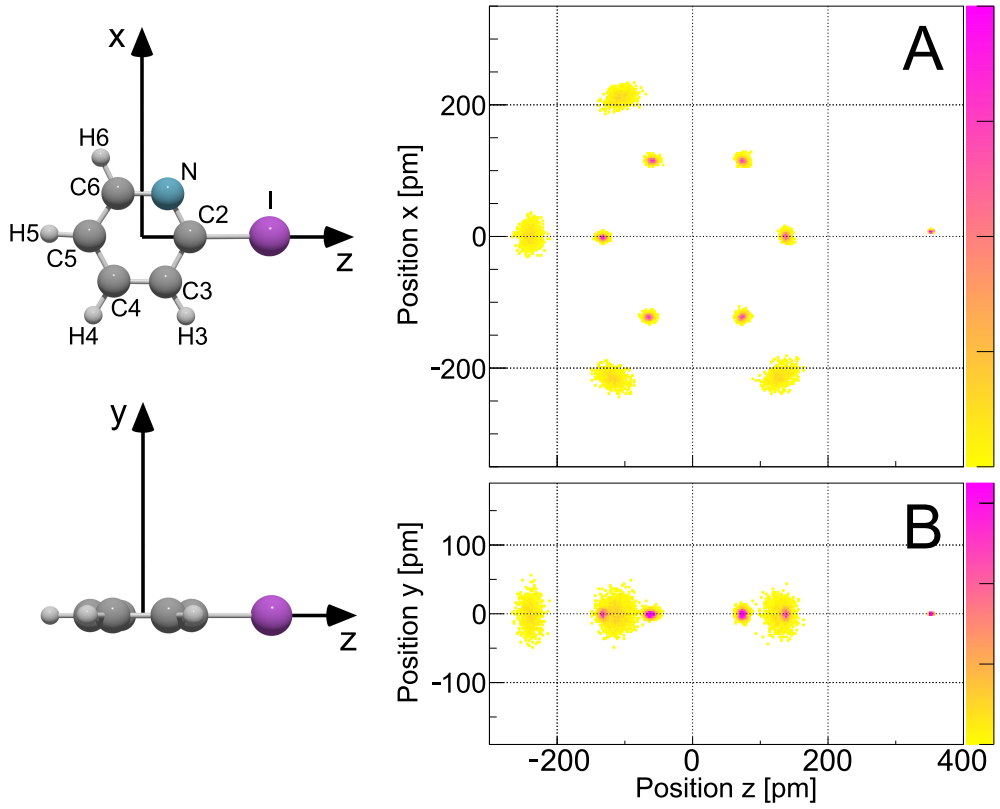

Figure S6: **Visualization of the position-space distribution of the molecule's ground state.** A) View of the molecular plane, B) perpendicular view.

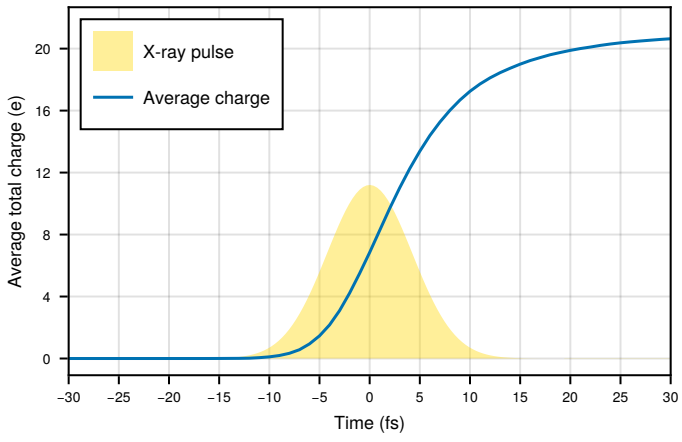

Figure S7: Average total charge of the molecule for trajectories ending up in the I^{4+}/N^{2+} coincidence channel. The yellow area represents the shape of the x-ray pulse.

Simulation software	XMDYN (revision 0.0-62-g4f4f5c8)
Number of trajectories	61 652
Initial state	Geometry and momenta sampled from the ground-state Wigner distribution
Ground-state potential	Harmonic approximation based on the Hessian matrix computed with the reactive force field
Time step	1 as
Simulation duration	1 ps (10^6 time steps)
Photon energy	2 keV
Pulse temporal shape	Gaussian
Photon polarization	Linear, randomly oriented with respect to the molecule
Pulse length	10 fs (FWHM)
Polarization	Random
Fluence	1.5×10^{11} photons/μm^2
Chemical bonds model	ReaxFF (full parameter table in Ref. (*46*))
Force field scaling	$\max\left(0, 1 - \frac{Q_{\text{tot}}}{4e}\right)$
Pauli exchange potential	$6647.6 \exp\left(-4.46 r/\text{Å}\right)$ eV

Table S2: **Parameters of the simulations.**

S3.3 Approximations of the model

The good agreement between the simulation and the experiment demonstrated in Figs. 2 and 3 indicates that all relevant processes are adequately included in our modeling, the most crucial one being the Coulomb repulsion, which amplifies initial molecular out-of-plane deformations and transfers them to the measured asymptotic momenta. In this Section, we discuss why we are able to achieve this high level of agreement, despite employing an approximate model for the chemical bonds of the molecule.

First, we model the disappearance of the bonds by progressively deactivating the force field, completely removing it when a total charge of +4 is reached in the molecule. It turned out that for charge values of 4 and higher, the exact value employed for this cutoff has only a very limited effect on the outcome of our modeling. In order to demonstrate this, we performed simulations in which the force field is turned off for different total charges. Figure S8 shows Newton plots of the C^{2+} and H^{+} ions for different cutoff charges between +3 and +17. The exact cutoff value has barely any effect.

The low impact of the cutoff charge extends to the PCs of the distribution. To demonstrate this fact, we show the PCs for different cutoff charges in Fig. S9. The PCs do not change qualitatively when the cutoff charge is changed. However, the contribution of each PC is slightly different when the cutoff is modified, as quantified by the explained variance (σ^2) of each of them.

These simulations confirm that the choice of the cutoff charge does not have a significant impact on the results and that the choice of a cutoff at a total charge of +4 is justified.

Our second approximation is that the force field that we are using corresponds to the molecule being in its ground state. Therefore, we are neglecting the effect of distinct intermediate cationic or vibrational states and the evolution of the dynamics on the corresponding potential energy surfaces (PESs). While this may seem a drastic approximation, one has to consider that the ultrafast ionization that we consider leaves only little time for any intermediate process to have an effect. To add a more quantitative view, we can estimate on which time scale typical vibrational motion takes place by looking at the period of the normal modes. The highest-frequency out-of-plane normal mode has a period of 27.5 fs, whereas the time at which the molecule is ionized is of the order of 10 fs (see Section S3.3 and Fig. S7). During this short time, several ionizations occur, so there is hardly

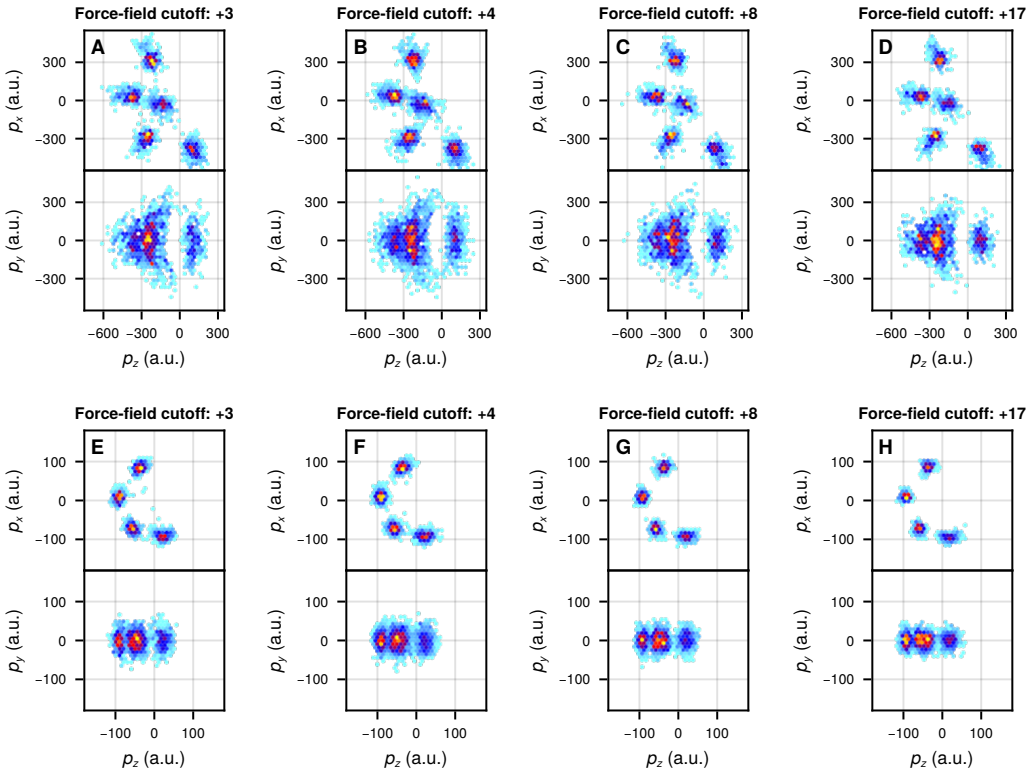

Figure S8: Newton plots of the simulated momenta for different total charge cutoff for the deactivation of the force field during the explosion. A-D. Carbon C^{2+} ions. E-H. Protons.

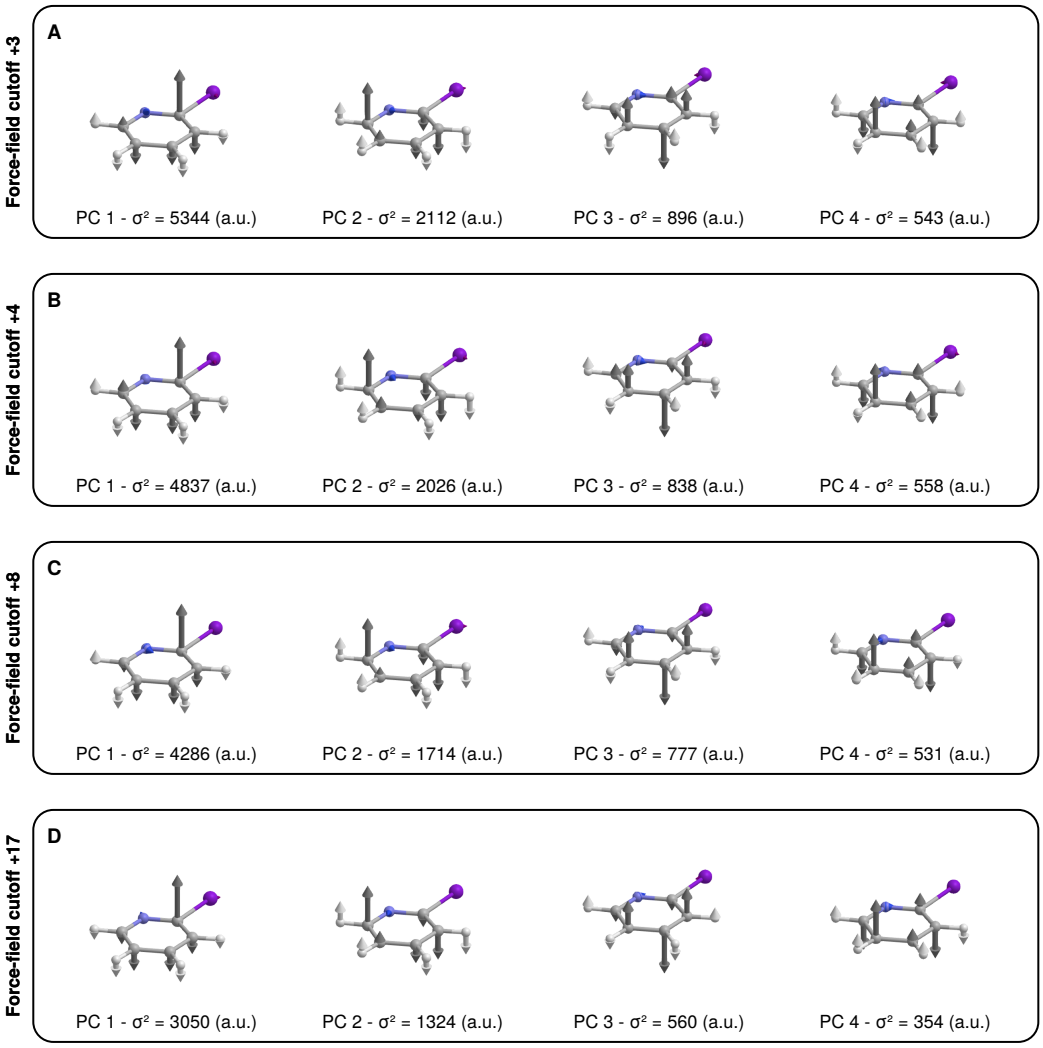

Figure S9: Dominant out-of-plane principal components of the distribution of momenta simulated with different cutoff charges at which the force field is disabled. Under each principal component, the associated explained variance (σ^2) is stated.

any time for the molecule to evolve on specific intermediate cationic-state PESs. The ionization is fast enough for the Coulomb forces to become dominant practically immediately. In fact, as we point out in Section S10, these conditions are only met when we consider coincidence-sets with sufficiently highly charged ions, which in turn correspond to the extreme ionization rate show in Fig. S7.

S3.4 Significance of the effect of the GSFs

We quantify the width of the out-of-plane momentum distributions shown in Figs. 2A-F by computing the standard deviation w of p_y for the hydrogen and carbon ions. For the H^+ ions, w is 4.4 ± 0.1 a.u. for simulations without GSFs, whereas it is 20.3 ± 0.2 a.u. for simulations with GSFs. For the C^{2+} ions, w is 31.5 ± 1.1 a.u. for simulations without GSFs and 127 ± 1.2 a.u. for simulations with GSFs. The uncertainty is the sampling error E_w on w, which is calculated as

$$E_w = \frac{1}{\sqrt{2(N-1)}} \sqrt{\frac{\langle p_y^4 \rangle}{\langle p_y^2 \rangle} - \langle p_y^2 \rangle}, \tag{S3}$$

where N is the number of detected ions of a given species in the I^{4+}/N^{2+} coincidence channel. The corresponding values obtained from the experiment are 38.7 ± 0.1 a.u. and 152.0 ± 0.4 a.u. for the hydrogen and carbon ions, respectively. These values become 22.7 ± 0.1 a.u. and 107.0 ± 0.4 a.u. after applying the filtering procedure described in section S12.

The difference in width between the simulation with and without GSFs is large (w is 4 times larger when the GSFs are considered), while the error E_w is comparatively small (less than 3% of the difference). This shows that the inclusion of the GSFs has a statistically significant impact on the width of the out-of-plane momentum distribution.

S4 Simulations at 250 K

In addition to the simulations presented in the main text, we performed simulations with a temperature of $T = 250$ K, which is a conservative upper bound for the temperature of the molecules in the jet, as described in section S1.4.

We ran 55124 trajectories initiated from the finite-temperature Wigner distribution at $T = 250$ K. The simulation method is otherwise identical to the one used in the main text and presented in

section S3.

Figures S10A and S10B show, respectively, the leading in-plane and out-of-plane PCs for $T = 0\,\mathrm{K}$, while Figs. S10C and S10D show, respectively, the leading in-plane and out-of-plane PCs for $T = 250\,\mathrm{K}$. The components are extremely similar regardless of the temperature. This fact is further quantified by computing their mutual overlap S (the absolute value of their scalar product, $S = 1$ corresponding to identical components), which are reported in Figs. S10E and S10F. Except for the in-plane PC 4 and PC 5, which exhibit some mixing, the change in temperature has no noticeable effect. Finally, Figs. S10G and S10H compare the fraction of explained variance associated with each component. Again, the results are similar regardless of the temperature.

From the simulation, we conclude that a temperature of $T = 250\,\mathrm{K}$ has no significant effect on the correlations between the momenta of the ions. In the experiment, the temperature of the molecules in the jet is likely to be lower than $T = 250\,\mathrm{K}$. Therefore, we conclude that thermal effects only play a negligible role in the explosion, and that the correlations seen in the final momenta are overwhelmingly caused by the GSFs, confirming that we are imaging ground-state properties of the molecule.

S5 Iodine-nitrogen recoil frame

The molecular recoil frame determined by the iodine and nitrogen ion emission is a simple and powerful way to represent the experimental data in a visually accessible way, as demonstrated in Fig. 1. However, this choice of frame has an impact on the variances of the data, as illustrated in Fig. S11. Assuming a hypothetical ion emission pattern where only the iodine ion is emitted out-of-plane (which violates momentum conservation but is useful for illustration purposes), a pattern as in Fig. S11A emerges. By construction, the variance in this example can be described by a single PC, illustrated in Fig. S11B. When these data are rotated into the iodine-nitrogen molecular recoil frame, corresponding to the treatment of our experimental data, it enforces a zero out-of-plane variance for the iodine and nitrogen momenta. Accordingly, most of the other ions now have a significant out-of-plane momentum in this frame of reference (Fig. S11C), similar to what we observe in Fig. 1. The corresponding PC exhibits a collective fluctuation in the out-of-plane momenta (Fig. S11D).

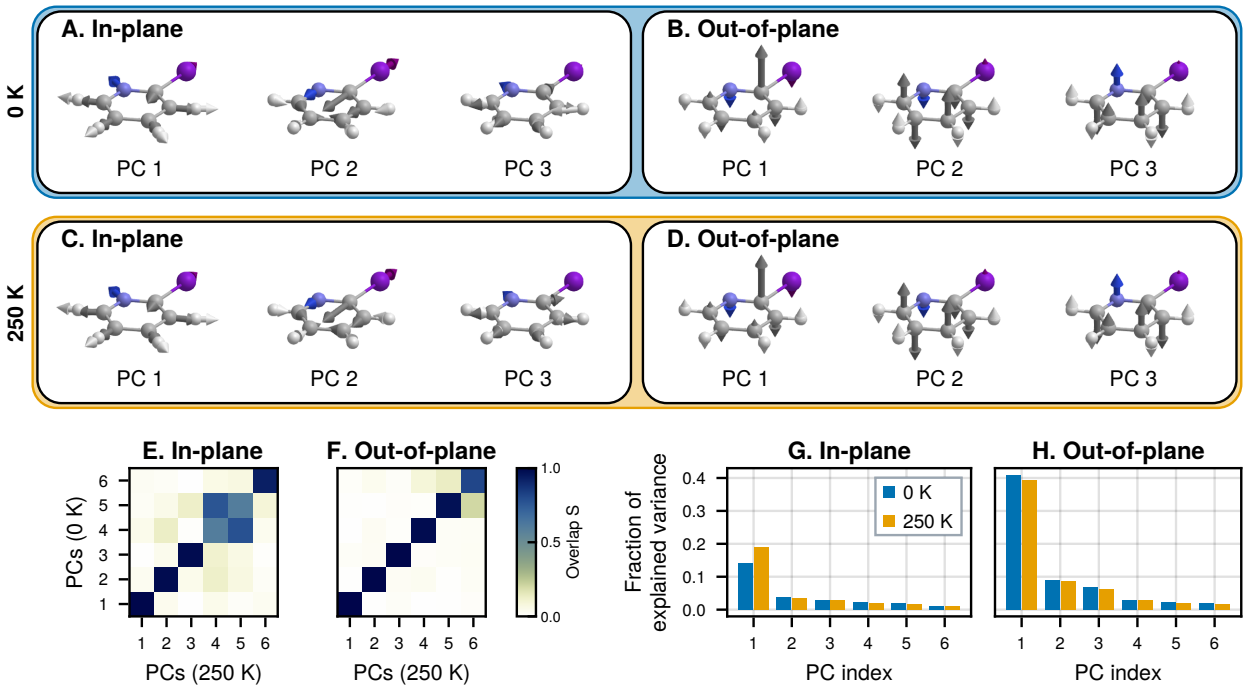

Figure S10: **Comparison of PCs in the simulation at a temperature of 0 K and 250 K.** A-B. Principal components of the momentum distribution from the simulations in the ground state, at 0 K. C-D. Principal components of the momentum distribution from the simulations at 250 K, a conservative upper bound for the temperature of the molecule in the experiment. E-F. Similarity of the principal components obtained at the two temperatures, computed as the scalar product between the components. G-H. Fraction of variance explained by each principal component for both temperatures.

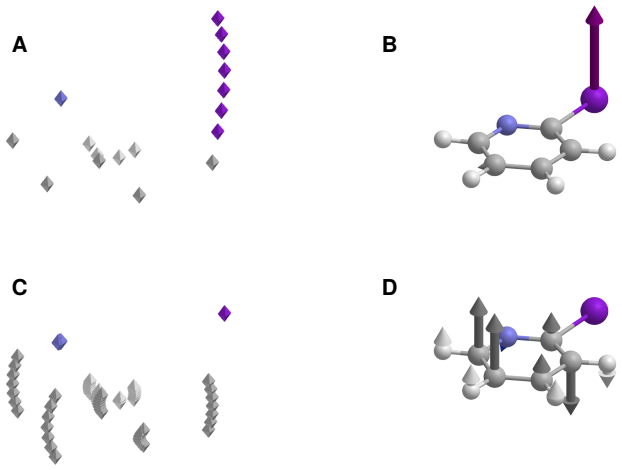

Figure S11: **Effect of the choice of molecular frame on PCs.** Artificial measurements where the iodine ion is emitted out-of-plane, while momenta of all other ions are fixed (relative to the laboratory frame). A. and B. In the laboratory frame. A. Individual measurements in momentum space. Each measurement is represented as a set of dots. B. Corresponding PC. C and D. In the iodine-nitrogen recoil frame. C. Individual measurements in momentum space. Each measurement is represented as a set of dots. D. Corresponding PC.

This example demonstrates that the use of a recoil frame defined by marker ions can significantly change the qualitative picture of the correlations between the ion momenta. In contrast, our approach of fitting a model distribution (see section S6) is based on mutual momentum scalar products as internal coordinates, and therefore the frame around which the fluctuations are considered is fitted directly to the multidimensional data without the need to perform a coordinate transformation to a different frame of reference. Therefore, the fitted model distribution addresses the momentum fluctuations in a frame that is not biased by a choice of marker ions.

S6 Additional data filtering for fitting and reconstruction

The raw data obtained from the experiment have a wide, featureless, low-density contribution underneath the localized spots, see Figs. 1A–D and Fig. S12A. It may originate from false coincidences that were not filtered out by the applied momentum gates (see section S1), or belong to cases where the charge-up induced by the x-rays happens along other fragmentation routes than in the majority of cases (e.g., where some of the direct ionization did not happen on the iodine atom). This contribution was found to reduce the quality of the results of our fitting procedure (see section S6), and we therefore apply an additional filter to remove it. The procedure is described here for the carbon ions, but the analogous procedure was used for protons.

We filter out subsets of ions measured in coincidence that involve the detection of an iodine (I^{4+}), nitrogen (N^{2+}), and any number of carbon (C^{2+}) ions. For each of those sets, we built all possible triplets of the form $(\mathbf{p}_\mathrm{I}, \mathbf{p}_\mathrm{N}, \mathbf{p}_\mathrm{C})$, where $\mathbf{p}_\mathrm{C}$ is the momentum of any of the measured C ions. From these momentum triplets, we generate sextets of internal coordinates containing the three momentum magnitudes and three relative angles,

$$\left(\|\mathbf{p}_\mathrm{I}\|, \|\mathbf{p}_\mathrm{N}\|, \|\mathbf{p}_\mathrm{C}\|, \sphericalangle(\mathbf{p}_\mathrm{I}, \mathbf{p}_\mathrm{N}), \sphericalangle(\mathbf{p}_\mathrm{I}, \mathbf{p}_\mathrm{C}), \sphericalangle(\mathbf{p}_\mathrm{N}, \mathbf{p}_\mathrm{C})\right), \tag{S4}$$

where $\sphericalangle(\cdot,\cdot)$ specifies the angle between two momenta. Since the norms and angles have different units, we whiten the data. That is, we center each internal coordinate around its mean value and divide it by its standard deviation. The whitened internal coordinates are unitless and have a standard deviation equal to 1. We estimate the density at each point in this internal-coordinate space by counting the number of neighbors in a 0.5 radius using nearest neighbor searches (*59*). If a possible sextet has more than a threshold amount of neighbors, we keep it, otherwise we discard

it. Based on visual inspection, we chose the threshold to be 5 for carbon atoms, and 15 for hydrogen atoms, which have a higher density.

We started with a data set with 34336 shots; after filtering the low-density regions, 8060 remain. The effect of the procedure is illustrated in Figs. S12A-B (carbon ions) and Figs. S12D-E (protons). Notably, it removes most of the large halo visible in the raw data (Figs. S12A and S12D) under the carbon ions and protons.

Moreover, when the low-density contribution is removed, the agreement between the simulation and the experiment is particularly good, as shown in in Figs. S12C and S12D (carbon ions) and Figs. S12G-H (protons).

S7 Fitting algorithm

To reconstruct the full dimensionality of the momentum distribution from the experimentally measured data, we fit a 33-dimensional multivariate Gaussian distribution to the recorded momenta. For this, several challenges of the experimental data need to be overcome:

1. Ions of different elements have different masses, thus their asymptotic momenta have very different magnitudes potentially biasing the analysis toward either light or heavy atoms.

2. The orientation of the molecules in the laboratory frame is random and thus unknown for a given shot.

3. Only a subset of all created ionic fragments is measured in coincidence in each x-ray shot, due to the limited detection efficiency of the detector.

4. Ions of the same element are indistinguishable from each other (i.e., the initial atomic site from which a given ion originates is unknown).

To perform a successful fit despite these limitations, we parameterize the distribution of momenta as follows: First, we weight each momentum by the inverse square root of the ion's mass. Hence, we define the weighted momenta as $\mathbf{p}_a = \sqrt{m_a}\mathbf{v}_a$, where $\mathbf{v}_a$ is the velocity of the detected ion a and m_a its mass. This allows a balanced treatment despite the fact that the ions have different masses. Second, we use scalar products between the weighted momenta, as they are independent

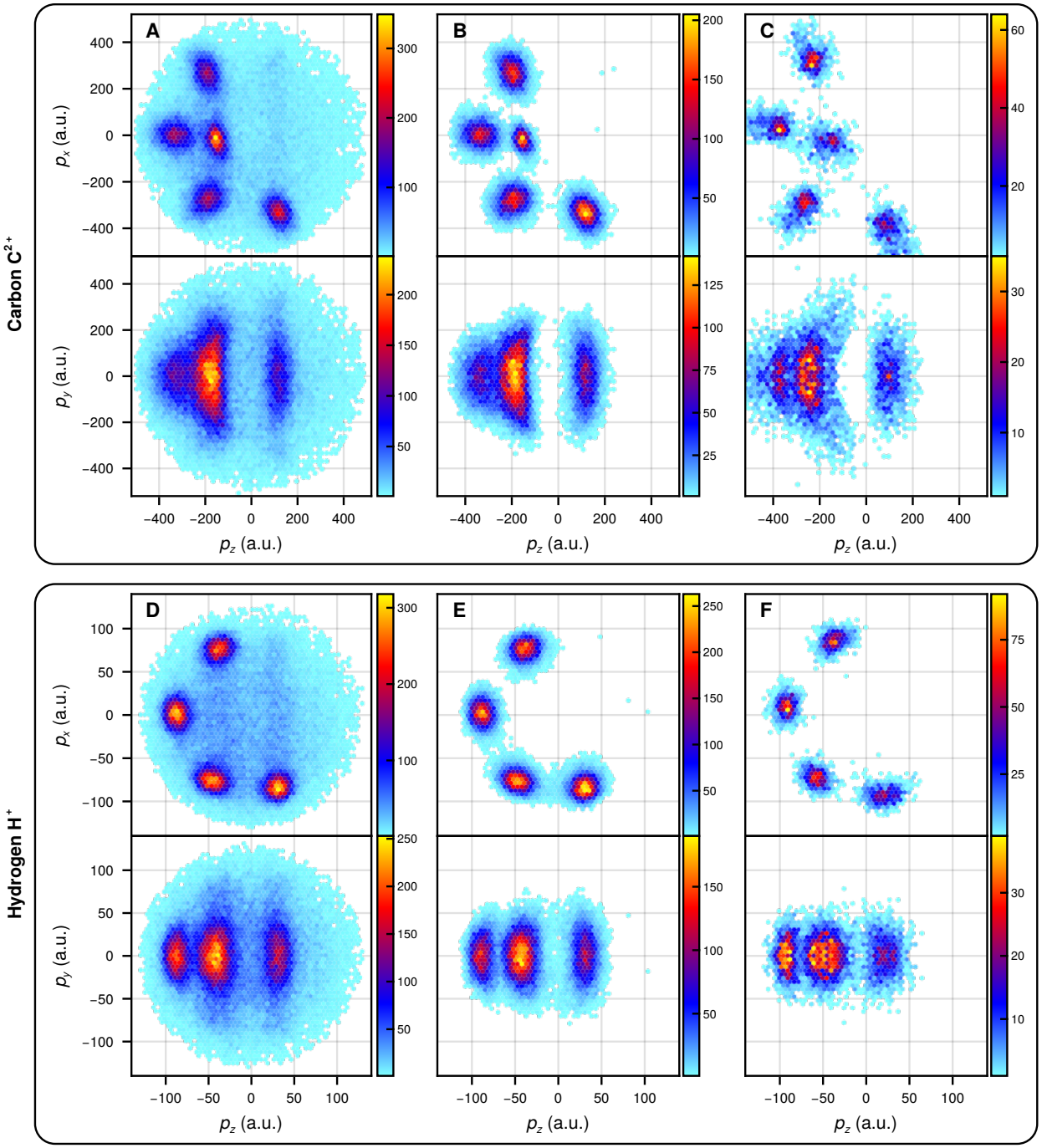

Figure S12: **Filtering of the experimental data.** A-C. Carbon ions. A. Experimental data before filtering. B. Experimental data after filtering. C. Simulation data. D-F. Protons. D. Experimental data before filtering. E. Experimental data after filtering. F. Simulation data.

of the rotation (i.e., laboratory-frame orientation) of the molecule. We consider all possible scalar products $(\mathbf{p}_a \cdot \mathbf{p}_b)$, where the ions a and b are detected in coincidence. Third, we use statistical moments of the distribution of scalar products including mixed moments involving scalar product of different pairs of ion momenta up to a given moment order. An example of such a scalar-product moment of order two is

$$\langle (\mathbf{p}_a \cdot \mathbf{p}_b)\ (\mathbf{p}_c \cdot \mathbf{p}_d) \rangle. \tag{S5}$$

Since the moments only involve a limited number of ions (in the given example, the ions a, b, c, and d), they are computed based on shots in which all the necessary ions were detected and are unaffected by the ions that remain undetected. Finally, we compute the moments *per element*, rather than per atom. Therefore, the indistinguishability of atoms of the same element in the experiment is not a problem. We name these moments the *per-element moments of the scalar products* (PEMSPs). By construction, they avoid all the limitations of the experimental data and can be used to fit a Gaussian model. Below we show how the PEMSPs are computed in practice, and we describe how we compare the PEMSPs computed from the experimental data and from a Gaussian model to perform the fit.

S7.1 Computing the PEMSPs from experimental data

The PEMSPs are calculated in the following steps. We filter the data as described in section S6. From the filtered data set, we compute the scalar products between the weighted momenta $\mathbf{p}_a$ of each pair of ions in a single measurement, and label them by the elements forming the pair. As an example, say we measure three ions in coincidence (I^{4+}, C^{2+}, C^{2+}) in a number of shots. This leads to six scalar products, namely

$$(\mathbf{p}_I \cdot \mathbf{p}_I)\,, (\mathbf{p}_I \cdot \mathbf{p}_C)\,, (\mathbf{p}_I \cdot \mathbf{p}_{C'}) \tag{S6}$$

$$(\mathbf{p}_C \cdot \mathbf{p}_C)\,, (\mathbf{p}_C \cdot \mathbf{p}_{C'}) \tag{S7}$$

$$(\mathbf{p}_{C'} \cdot \mathbf{p}_{C'})\,. \tag{S8}$$

We then compute statistical moments of the scalar-product distribution. In general, a moment $M_{\mathbf{i}}$, of order n, is specified by pairs of measured ions that define the corresponding scalar products,

$$\mathbf{i} = \{(a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)\}. \tag{S9}$$

Note that some scalar products, in the example $\left(\mathbf{p}_C \cdot \mathbf{p}_C\right)$ and $\left(\mathbf{p}_{C'} \cdot \mathbf{p}_{C'}\right)$, are equivalent. We can, however, distinguish between the scalar product of one carbon ion momentum with itself ($\left(\mathbf{p}_C \cdot \mathbf{p}_C\right)$) and the scalar product between the momenta of two different carbon ions ($\left(\mathbf{p}_C \cdot \mathbf{p}_{C'}\right)$), since they are detected in coincidence. To account for the equivalence of atoms of the same elements, we average over all equivalent moments. For the previous example, this means

$$M_{\{(I,C'),(C,C')\}} = \frac{1}{N_{\text{shot}}} \sum_k \frac{1}{2} \left(\left(\mathbf{p}_I \cdot \mathbf{p}_{C'}\right)^{(k)} \left(\mathbf{p}_C \cdot \mathbf{p}_{C'}\right)^{(k)} + \left(\mathbf{p}_I \cdot \mathbf{p}_C\right)^{(k)} \left(\mathbf{p}_{C'} \cdot \mathbf{p}_C\right)^{(k)} \right), \tag{S10}$$

where the sum over k runs over all N_{shot} measurements and $\left(\mathbf{p}_a \cdot \mathbf{p}_b\right)^{(k)}$ is the momentum scalar product from the kth in-coincidence measurement of ions with element types a and b.

In general, a moment is computed as (*60*)

$$M_{\mathbf{i}} = \frac{1}{Z_{\mathbf{i}}} \sum_k \sum_{\substack{\mathbf{j} \text{ in } k \\ \mathbf{j} \sim \mathbf{i}}} \prod_{\{a,b\} \in \mathbf{j}} \left(\mathbf{p}_a \cdot \mathbf{p}_b\right)^{(k)}, \tag{S11}$$

where the product $\prod_{\{(a,b)\} \in \mathbf{j}}$ runs over all ion pairs within label $\mathbf{j}$, the sum over $\mathbf{j}$ runs over all labels that are indistinguishable from $\mathbf{i}$ in measurement k, the sum k runs over all measurements, and $Z_{\mathbf{i}}$ is the total number of contributions to the moment, from all measurements (for example, $Z_{\mathbf{i}} = 2N_{\text{shot}}$ in Eq. (S10)).

The moments are computed using an implementation of online-statistics algorithms (*61*). Due to the high number of shots and the large number of possible ion combinations, it is necessary to limit the required memory to a manageable amount by using such online approach, which only consider one measurement at a time, updating a running tally for all the moments. This procedure yields a set of PEMSPs that characterize the distribution of ion momenta as measured in the experiment.

S7.2 Computing the PEMSPs from a Gaussian model

We consider a $3N$ multivariate Gaussian model, defined by the vector of mean values $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$. The probability density for this model is (*60*)

$$g(\mathbf{p}) = \frac{1}{Z} \exp\left[-\frac{1}{2} (\mathbf{p} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{p} - \boldsymbol{\mu}) \right], \tag{S12}$$

with the normalization constant $Z = \sqrt{(2\pi)^{3N} \det \mathbf{\Sigma}}$. In Eq. (S12) the concatenated weighted-momentum vector,

$$\mathbf{p} = \begin{pmatrix} \mathbf{p}_I \\ \mathbf{p}_N \\ \mathbf{p}_{C2} \\ \vdots \\ \mathbf{p}_{C6} \\ \mathbf{p}_{H3} \\ \vdots \\ \mathbf{p}_{H6} \end{pmatrix}, \tag{S13}$$

is used, where $\mathbf{p}_a$ is the three-dimensional weighted momentum of ion a.

The joint distribution of all scalar products generated from the Gaussian weighted-momentum distribution in Eq. (S12) is

$$g_{\mathrm{SP}}(\mathbf{s}) = \int \mathrm{d}^{3N}p \prod_{a \le b \le N} \delta(s_{ab} - \mathbf{p}_a^T \mathbf{p}_b) g(\mathbf{p}), \tag{S14}$$

where the vector $\mathbf{s}$ is an $N(N+1)/2$-dimensional vector containing scalar-product values s_{ab} for all combinations of ions a and b, $\delta(\cdot)$ is the δ distribution, and the integration is over the full $3N$-dimensional momentum space ($N = 11$). The characteristic function (*60*) of this distribution is

$$\tilde{g}_{\mathrm{SP}}(\mathbf{s}) = \frac{1}{Z} \int \mathrm{d}^{3N}p \ \exp\left[-\frac{1}{2} (\mathbf{p} - \boldsymbol{\mu})^T \mathbf{M}(\mathbf{k}) (\mathbf{p} - \boldsymbol{\mu}) - i \boldsymbol{\mu}^T \mathbf{A}(\mathbf{k}) \mathbf{p} \right], \tag{S15}$$

where $\mathbf{k}$ is an $N(N+1)/2$-dimensional vector in Fourier space with element k_{ab} (one element for each pair of ions). Moreover,

$$\mathbf{M}(\mathbf{k}) = \mathbf{\Sigma}^{-1} - i\mathbf{A}(\mathbf{k}), \tag{S16}$$

and

$$\mathbf{A}(\mathbf{k}) = \begin{pmatrix} 2k_{11}\mathbf{1} & k_{12}\mathbf{1} & \cdots & k_{1N}\mathbf{1} \\ k_{12}\mathbf{1} & 2k_{22}\mathbf{1} & \cdots & k_{2N}\mathbf{1} \\ \vdots & \vdots & \ddots & \vdots \\ k_{1N}\mathbf{1} & k_{2N}\mathbf{1} & \cdots & 2k_{NN}\mathbf{1} \end{pmatrix}, \tag{S17}$$

where $\mathbf{1}$ is the 3×3 identity matrix. The symmetric matrix $\mathbf{A}(\mathbf{k})$ is used to represent a sum of scalar products between 3-dimensional weighted momenta, as

$$\frac{1}{2}\mathbf{p}^T\mathbf{A}(\mathbf{k})\mathbf{p} = \sum_{a\le b\le N} k_{ab}\mathbf{p}_a^T\mathbf{p}_b. \tag{S18}$$

The integral in Eq. (S15) can be performed analytically (*62*), yielding

$$\tilde{g}_{SP}(\mathbf{k}) = [\det\boldsymbol{\Sigma}\,\mathbf{M}(\mathbf{k})]^{-\frac{1}{2}}\exp\left[\frac{i}{2}\boldsymbol{\mu}^T\mathbf{A}(\mathbf{k})\,\boldsymbol{\mu}\right]\exp\left[-\frac{1}{2}\boldsymbol{\mu}^T\mathbf{A}(\mathbf{k})\,\mathbf{M}^{-1}(\mathbf{k})\,\mathbf{A}^T(\mathbf{k})\,\boldsymbol{\mu}\right], \tag{S19}$$

Using Eq. (S19), any statistical moment of the $N(N+1)/2$-dimensional distribution of scalar products is retrieved as (*60*)

$$m_{\{(a_1,b_1),(a_2,b_2),\dots\}} = \langle(\mathbf{p}_{a_1}\cdot\mathbf{p}_{b_1})\,(\mathbf{p}_{a_2}\cdot\mathbf{p}_{b_2})\dots\rangle = \frac{(-i)\,\partial}{\partial k_{a_1,b_1}}\frac{(-i)\,\partial}{\partial k_{a_2,b_2}}\dots\tilde{g}_{\mathrm{SP}}(\mathbf{k})\Big|_{\mathbf{k}=\mathbf{0}}. \tag{S20}$$

To take into account the fact that in the experiment, each ion cannot be identified with a specific atom in the molecule, we average the computed moments over all equivalent labels arising from indistinguishable ions. This yields the PEMSPs,

$$\widehat{M}_{\mathbf{i}} = \langle m_{\mathbf{j}}\rangle_{\mathbf{j}\sim\mathbf{i}}, \tag{S21}$$

where $\langle\cdot\rangle_{\mathbf{j}\sim\mathbf{i}}$ means averaging over equivalent labels. In the following, we consider these PEMSPs as explicit functions of the parameters of the Gaussian weighted-momentum distribution [Eq. (S12)], i.e., $\widehat{M}_{\mathbf{i}}(\boldsymbol{\mu},\boldsymbol{\Sigma})$.

S7.3 Initial guess

Before performing the fitting, an initial guess has to be made. For the experimental data, we fit a mixture of a three-dimensional Gaussian distributions (a sum of Gaussian distribution) to the filtered momenta in the iodine-nitrogen recoil frame (Fig. S12B). We then use the resulting mean values as a guess for $\boldsymbol{\mu}$ and the resulting variances as a guess for the 3×3 block diagonal of the covariance matrix $\boldsymbol{\Sigma}$. With this guess, some diagonal elements would be zero, namely the components corresponding to the variance of the iodine momentum along the x and y axes and the variance of the nitrogen momentum along the y axis, due to the definition of the frame. For these diagonal elements, we take 10% of the total variance of the respective ions along the other directions as initial guess. This prevents the covariance matrix guess $\boldsymbol{\Sigma}$ from being singular.

S7.4 Loss function

In order to fit the PEMSPs modelled by the Gaussian distribution, $\widehat{M}_{\mathbf{i}}$, to the experimental PEMSPs, $M_{\mathbf{i}}$, we need to define a loss function. This loss function must suitably handle moments of widely different magnitude and order in a balanced way. In addition, it has to be considered that the experimentally determined moments can have statistical errors that cover a wide numerical range, because moments involving many different momentum scalar products depend on in-coincidence measurements of the corresponding ions, whose statistics become lower the more ions are involved. We achieve this by employing for each moment the sum of the squared difference of the moments $\left(\widehat{M}_{\mathbf{i}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) - M_{\mathbf{i}}\right)^2$ and the squared standard error for the experimental moment $\delta M_{\mathbf{i}}^2$. We form the ratio of this sum to the squared standard error and take the nth root for moments of order n. For each moment considered, this results in the individual loss contribution

$$D_{\mathbf{i}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \left(\frac{\left(\widehat{M}_{\mathbf{i}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) - M_{\mathbf{i}}\right)^2 + \delta M_{\mathbf{i}}^2}{\delta M_{\mathbf{i}}^2} \right)^{\frac{1}{n_{\mathbf{i}}}} - 1, \tag{S22}$$

where $n_{\mathbf{i}}$ is the order of moment $M_{\mathbf{i}}$. The shift by -1 enforces the convention that perfect agreement ($\widehat{M}_{\mathbf{i}} = M_{\mathbf{i}}$) has a loss of exactly zero.

We combine the loss terms for the individual moments in a way that adaptively reduces weights on moments that are already very well approximated. We consider individual moments computed from the model to be well approximated when the absolute difference to the experimental moment is smaller than twice its standard error. Taken together, we employ the total loss function

$$L(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \left[\frac{1}{N_{\mathrm{PEMSP}}} \sum_{\mathbf{i}} \alpha\left(\widehat{M}_{\mathbf{i}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), M_{\mathbf{i}}\right)\ D_{\mathbf{i}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \right]^{\frac{1}{2}} \tag{S23}$$

where N_{PEMSP} is the number of moments considered, and $\alpha(\widehat{M}_{\mathbf{i}}, M_{\mathbf{i}})$ is the damping term

$$\alpha(\widehat{M}_{\mathbf{i}}, M_{\mathbf{i}}) = \begin{cases} 0.01 & \text{if } |\widehat{M}_{\mathbf{i}} - M_{\mathbf{i}}| < 2\delta M_{\mathbf{i}}, \\ 1 & \text{otherwise.} \end{cases}, \tag{S24}$$

which reduces the effect of terms that already have a low individual loss compared to the statistical error on the experimental moment.

We minimize this loss function using the limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm implemented in the Optim julia package (*63*). This requires the gradient of the loss

function with respect to all of the parameters, namely the mean $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$ of the Gaussian distribution. Using the $\boldsymbol{\Sigma}$ matrix directly causes problems, when the matrix becomes nonsymmetric or not positive definite. Therefore, we employ the Cholesky decomposition of the covariance matrix, $\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^T$ and use the lower triangular matrix $\mathbf{L}$ as parameter, ensuring that $\boldsymbol{\Sigma}$ is always symmetric and positive definite. The gradient is computed using the Zygote julia package (*64*), and implementing custom differentiation rules for most of the internal functions with the ChainRules julia package (*65*).

The number of moments grows extremely rapidly with moment order, making the usage of high moment order unpractical and computationally very expensive. Therefore, for the results shown in the main text, we restrict ourselves to a maximal moment order of 3. Because this involves maximally three mutual scalar products, combined information of up to 6 ions measured in coincidence is considered in the fit.

A high moment order is expected to be necessary when the molecule being imaged contains many atoms of the same species. In this case, the number of different PEMSPs depends mostly on the number of species in the molecules, as the moments are averaged over all indistinguishable atoms. To get more distinct PEMSPs, and thus more information about the distribution, a possibility is therefore to increase the maximal moment order considered in the fitting. In turn, this comes with experimental challenges, as an accurate estimation of higher order moments requires better statistics.

S7.5 Limitations of the fit

The main assumption of the fit is that the momentum distribution follows a Gaussian distribution. This approximation holds well for the data used in this work. However, the Coulomb explosion of some molecules may produce data that has a non-Gaussian distribution. This is expected, for example, if the system is probed while being in a state where its nuclear wave packet is split: in this case the resulting distribution would be bimodal, which is not compatible with the assumption of a Gaussian distribution.

In principle, this limitation can be lifted by considering more general distributions (e.g., a sum of Gaussian distributions to model bimodal data). Fitting to a more complex distribution may then in turn require better statistics or higher level of coincidence in the measurements.

Furthermore, even at the level of a Gaussian model of the data, the fit is uniquely determined only up to the first few leading PCs. This is a standard limitation due to finite sampling: Higher statistics are necessary to reliably reconstruct more PCs.

S8 Application of the fitting to the simulated data

In order to test the validity of the method, we first apply it to the simulated data, for which the input is known, before fitting the experimental data. When applying the algorithm on the simulation data, to obtain the PCs shown in Fig. 3, we reproduce the detection conditions of the experiment. Ions only have a 60% probability to be detected, corresponding to the detector efficiency in the experiment, and we discard all ions that do not have the respective charge (+4 for iodine, +2 for nitrogen and carbon, and +1 for hydrogen). We discuss the effect of this reduction of the simulation data in more detail in section S8.3.

S8.1 Initial guess

As an initial guess for the simulation data, we use the computed mean values and variances as starting points for $\boldsymbol{\mu}$ and the diagonal entries of the covariance matrix $\boldsymbol{\Sigma}$. All off-diagonal entries of the covariance matrix were initially set to zero. This starting point thus assumes that all ion-momentum coordinates are uncorrelated.

S8.2 Convergence

In order to assess the quality of the fitting procedure, we first test it on the simulated data. For the test we consider the momentum distribution data where one I^{4+}, one N^{2+}, and at least one C^{2+} or H^{+} appear; i.e., the charge state of additional carbon and hydrogen ions is ignored (see the discussion in section S8.3 on how this momentum distribution differs from the one obtained in the measurement). Figures S13A and S13B show the dominant in-plane and out-of-plane PCs, respectively, computed from the full simulated data set. We then artificially introduce limitations in the data set, similar to the ones known to affect the experimental data (see above). Specifically, we sampled the momentum sets of 10^6 randomly oriented molecules, considered all possible 6-ion coincidence subsets from it to mimic the coincidence level obtained in the experiment, and removed the knowledge about the initial

site of the ions. The dominant PCs reconstructed from these limited data are shown in Figs. S13C and S13D. The three dominant in-plane and the three dominant out-of-plane components explain more than 80% of the variance in the data (see Figs. S13E and S13F). Figures S13G and S13H quantify the overlap S of the PCs from both analyses, defined as the absolute value of their scalar product. A perfect agreement corresponds to $S = 1.0$. The fit reproduces the original distribution very well for many of the PCs. The in-plane PCs are reproduced with a lower quality, except for in-plane PC 1. This is expected, as the PCs corresponding to the highest explained variance are the easiest to reconstruct and out-of-plane PCs explain more variance than the in-plane PCs, as shown in Figs S13E and S13F. The explained variance of in-plane components 2 and 3 is only less than 0.05. The successful reconstruction of the momentum distribution establishes that it is possible to extract the main characteristics of the full momentum distribution of 11 ions from data including only coincidence information on a much smaller number of ions.

To investigate the convergence of the algorithm, Figs. S14A and S14B show how the loss function varies during the fitting for the experimental and simulated data, respectively. As expected, the value of the loss function converges to zero after a few thousand iterations.

For the simulation data, we can study the convergence of the algorithm in more detail, since we can use the full data set as a reference distribution. In s case, we quantify the dissimilarity between the reference and the reconstructed distribution using the Kullback-Leibler divergence (KLD) D_{KL}, which is defined as (*60*)

$$D_{\mathrm{KL}}(g \| g_{\mathrm{ref}}) = \int \mathrm{d}^{3N} p \, g(\mathbf{p}) \ln\left(\frac{g(\mathbf{p})}{g_{\mathrm{ref}}(\mathbf{p})}\right) \tag{S25}$$

for a momentum probability distribution $g(\mathbf{p})$ and a reference distribution $g_{\mathrm{ref}}(\mathbf{p})$. The KLD is 0 when the probability distribution $g(\mathbf{p})$ matches the reference perfectly; otherwise, it gives a measure for the information difference between the two distributions.

In Fig. S14B, we show the KLD from the reference together with the loss function for the fit to the simulation data. We see that, as expected, the dissimilarity between the fitted model and the reference mostly steadily decreases, similar to the loss function.

We also inspect the KLD to explore the effect of using different restrictions on the maximum coincidence order and maximum moment orders considered. In Fig. S14C, we show how the quality of the fitted model changes when we limit the coincidence order. The plot shows that using

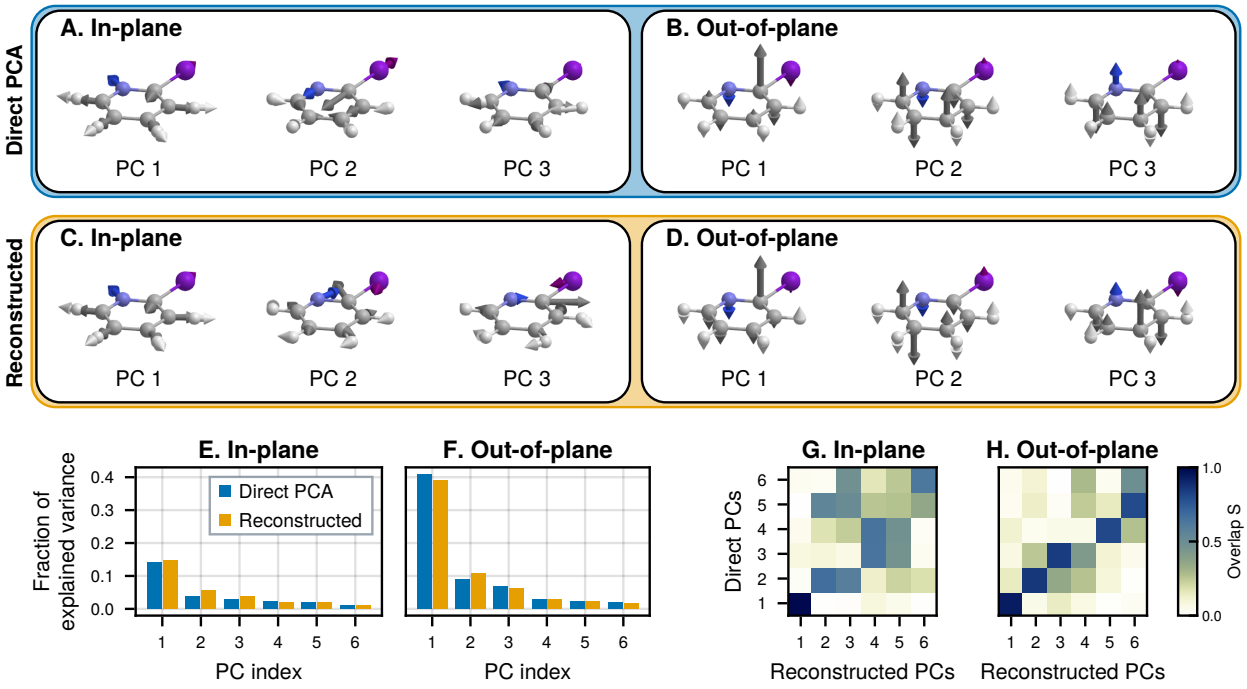

Figure S13: **Comparison of the fits using the full and the limited simulated data sets.** A-B. Principal components of the momentum distribution from the simulations, computed directly from the complete data, where the laboratory-frame rotation and the ion identity are known. C-D. Principal components of the momentum distribution, reconstructed using our algorithm, with information similar to that available in the experiment, i.e., only information on 6 ions is covered, the original ion location inside the molecule is unknown, and the molecule has a random and unknown rotation. E-F. Fraction of variance explained by each principal component for both procedures. G-H. Similarity of the principal components obtained with the two methods, computed as the scalar product between the components.

a coincidence order of at least 3 significantly improves the reconstruction, while using a higher coincidence order has a comparatively minor impact. We note that increasing the coincidence order does not always leads to strictly better KLD, but the changes beyond coincidence order 3 are comparatively small. It is possible that, when a more general distribution (i.e., a non-Gaussian distribution) is fitted, the fit procedure might benefit from a higher coincidence order.

In Fig. S14D, we show the effect of changing the maximum moment order. As can be seen, using a higher maximum moment order slightly worsens the reconstruction. This is expected, as high-order moments are more sensitive to the effect of finite sampling than lower order ones. Taking them into account leads to overfitting that can counterbalance the benefit of including more information in the fit procedure. However, the limited effect on the quality of the reconstruction shows that, with our choice of loss function, this has no significant negative effect. We expect that when considering experiments with more statistics, a higher moment order leads to an overall better fit. In conclusion, we are working with a range of parameters for which the reconstruction yields consistently good and stable results, particularly for the out-of-plane PCs.

S8.3 PCA of the simulation data

In this article, we fit a Gaussian model to the distribution of measured asymptotic momenta of ion tuples that match the selected coincidence channel (i.e., I^{4+}/N^{2+} and any number of ions of type C^{2+}/H^{+}). This procedure can be performed for the experimental data and for the simulation data, yielding the mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ of the respective distributions. The PCs are then computed as the eigenvectors of the covariance matrix $\boldsymbol{\Sigma}$.

In general, each combination of ion charge states may result in a different momentum distribution. Multiple combinations of ion charge states contribute to the measured distribution of momenta since only subsets of ions are detected and the charge states of the undetected ions remain unknown. To explore how the unknown charge states of undetected ions affect the momentum distribution, we tested an alternative procedure and pick simulation data in which all momenta are taken into account if I^{4+}, N^{2+}, and at least one C^{2+} or one H^{+} ion are formed. This generalizes the data set including also ions (and neutral atoms) in the data analysis that do not have the charge state imposed by the given coincidence criterion. This procedure allows us to apply the PCA analysis on the simulation data without performing a fit based on the simulated measurements.

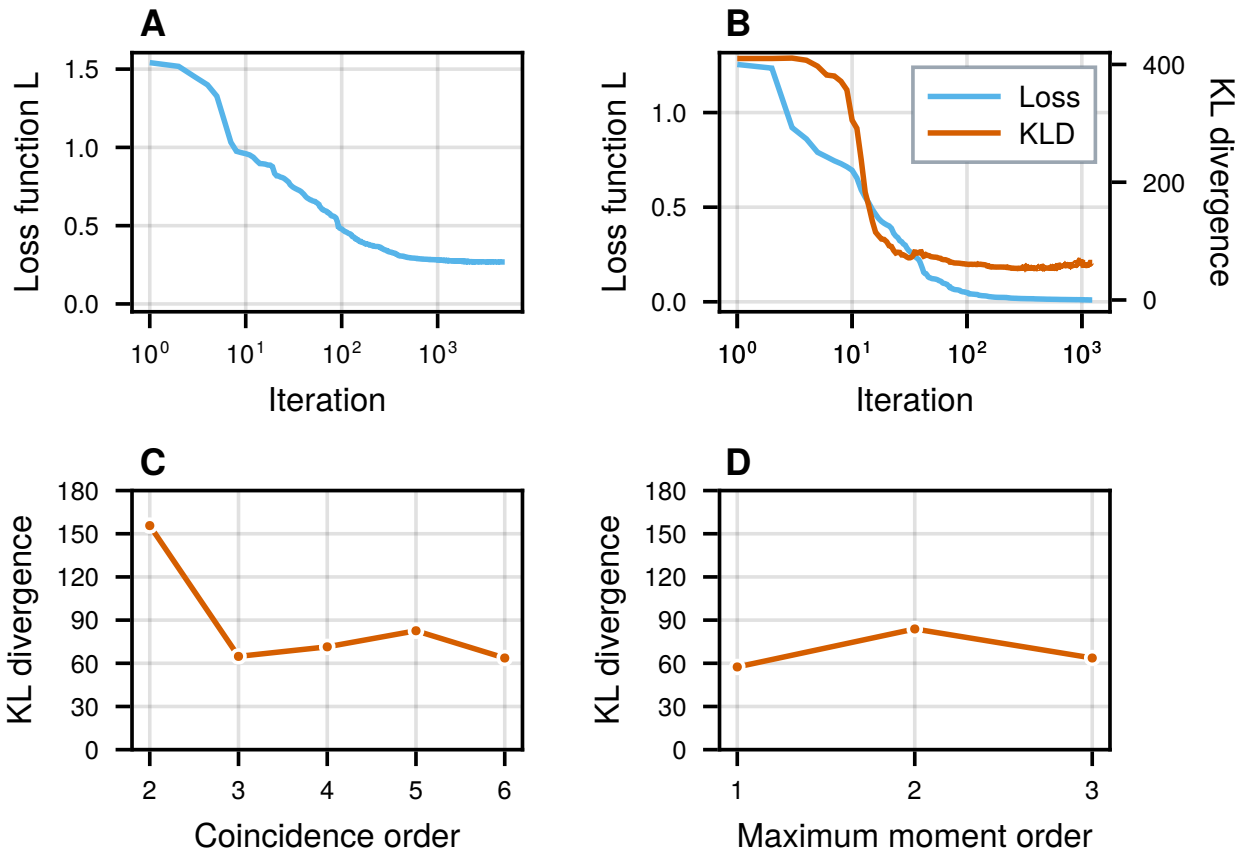

Figure S14: **Convergence of the algorithm.** A. Loss function during the fitting of the experimental data. B. Loss function and KLD from the reference during the fitting of the simulation data. C. KLD of the Gaussian model fitted to the simulation data, for various coincidence orders and moment order 3. D. KLD of the Gaussian model fitted to the simulation data, for three different moment orders and coincidence order 6.

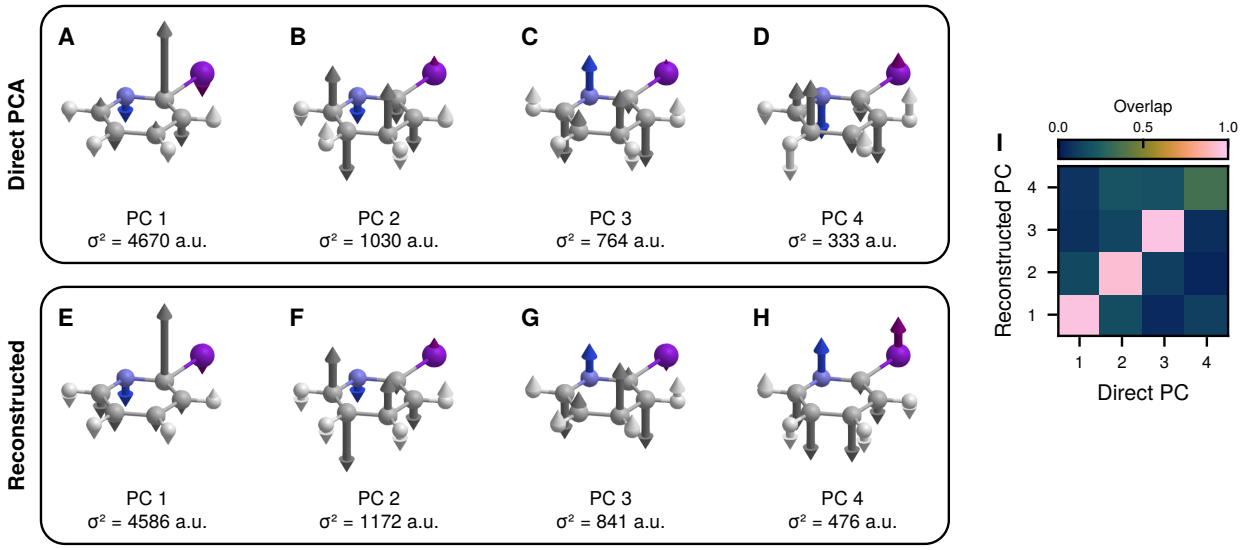

Figure S15: **Principal components of the distribution of asymptotic momenta.** A-D. Computed from the simulation using PCA, disregarding the charge of the atoms. E-H. Reconstructed from the simulation by fitting simulated measurements. σ^2 represents the variance explained by the component. I. Overlap between the computed and reconstructed principal components.

The out-of-plane PCs obtained for the momentum distributions in Figs. S15A-S15D are compared with those PCs obtained from the fit based on the simulated measurements shown in Fig. S15E-S15H. The first three PCs are in close agreement, both in terms of their shape and explained variance, whereas the fourth one is different. This is further illustrated by the overlap matrix shown in Fig. S15I.

The comparison shows that the out-of-plane momentum distribution is only slightly influenced by the fact that the charge state of undetected ions is unknown. We can thus conclude that the 33-dimensional distribution generated by measuring only the specific subset of ions (covering I^{4+}, N^{2+}, and at least one C^{2+} or one H^{+}) is representative of the distribution of ions where the undetected carbon and hydrogen ions may have any charge state. This result is a consequence of the efficient charge redistribution in the molecule during the ionization dynamics and the subsequent Coulomb explosion. For a given charge state, it leads to similar ion momenta regardless of how the charge is distributed among the undetected ions.

In general, however, the unknown charge states of undetected ions can be a challenge for both the theory and the data analysis. The charge of undetected ions might have an impact on the resulting

momentum distributions. Covering such effects would require generalizing the fit procedure to consider all relevant combinations of ion charge states. This poses further challenges also for the simulation, since it is hard to gather sufficient statistics for a specific charge state.

S9 Mode contributions

S9.1 Linear mapping

In Fig. 4, we show the mode contributions of the out-of-plane NMs to the out-of-plane PCs in the simulation. The mode contributions are given by the sum of squared correlation coefficients,

$$r_{ij}^2 = \mathrm{cor}\big[r_i, P_j\big]^2 + \mathrm{cor}\big[p_i, P_j\big]^2 = \frac{\mathrm{cov}\big[r_i, P_j\big]^2}{\mathrm{var}[r_i]\,\mathrm{var}\big[P_j\big]} + \frac{\mathrm{cov}\big[p_i, P_j\big]^2}{\mathrm{var}[p_i]\,\mathrm{var}\big[P_j\big]}, \tag{S26}$$

where $\mathrm{cor}[\cdot,\cdot]$ is the Pearson correlation coefficient, $\mathrm{cov}[\cdot,\cdot]$ denotes the covariance, and $\mathrm{var}[\cdot]$ denotes the variance. We obtain the mode contributions by considering the linear model

$$P_j - \langle P_j \rangle = \sum_i [A_{ji}\,(r_i - \langle r_i \rangle) + B_{ji}\,(p_i - \langle p_i \rangle)], \tag{S27}$$

where $\langle\cdot\rangle$ specifies mean values. The coefficients A_{ji} and B_{ji} are fit parameters of the linear model. Using the statistical independence of the NM coordinates r_i and p_i, they can be expressed through covariances and variances, leading to

$$-\langle P_j \rangle = \sum_i \left\{ \mathrm{cor}\big[r_i, P_j\big]\ \sqrt{\frac{\mathrm{var}\big[P_j\big]}{\mathrm{var}[r_i]}}\ (r_i - \langle r_i \rangle) + \mathrm{cor}\big[p_i, P_j\big]\ \sqrt{\frac{\mathrm{var}\big[P_j\big]}{\mathrm{var}[p_i]}}\ (p_i - \langle p_i \rangle) \right\}. \tag{S28}$$

Taking the variance of both sides of Eq. (S28) yields

$$\mathrm{var}\big[P_j\big] = \sum_i \left\{ \mathrm{cor}\big[r_i, P_j\big]^2\ \mathrm{var}\big[P_j\big] + \mathrm{cor}\big[p_i, P_j\big]^2\ \mathrm{var}\big[P_j\big] \right\}, \tag{S29}$$

which, using Eq. (S26), means that

$$\sum_i r_{ij}^2 = 1. \tag{S30}$$

Additional variations in the momenta that are not caused by the GSFs or that are beyond the linear model of Eq. S28 will cause the right-hand side of Eq. S30 to differ from unity. The individual terms in this sum, r_{ij}^2, can be read as contributions from fluctuations in NM i to the variance along final-momentum PC j.

In Fig. S16, we show mode contributions for all combinations of NM and principal components of the final momentum after the Coulomb explosion. We see that there is no cross-contribution between in-plane and out-of-plane, as one expects from the consideration that initial out-of plane structure fluctuations are mainly responsible for out-of-plane momenta. Most of the in-plane PCs are only weakly correlated with in-plane NM coordinates, with some exceptions, for example, the NM at 7.2 THz. The main feature of this NM is to shift the iodine atom sideways in plane relative to the pyridine ring. This effect is amplified via ion collision during the explosion (*40*), leading to a strong relationship with the PCs that relate to this collision.

Figure S16 shows the total mode contribution of all NMs to each PC. The total contributions to all out-of-plane PCs are close to 1, while it is significantly smaller for the in-place PCs. This indicates that the NMs map linearly to the out-of-plane PCs, however, the in-plane PCs can only partially be attributed to GSFs using a linear model.

S9.2 Limitations for the identification of NM features

In this work, we rely on the simulations to connect the experimental data to the GSFs, but not all relevant interactions are perfectly simulated. Chemical forces are computed with the reactive-force-field method described in section S3, but one cannot expect that they accurately represent non-Coulombic effects in intermediate cationic states of the molecule. Neglecting such non-Coulombic effects is valid as long as the explosion is sufficiently violent, such that Coulomb interaction dominates the dynamics. This is demonstrated in Section S10 where we show that the simulation for lower total charge on the molecule matches the experimental data to a smaller extent.

In the future, this limitation can be overcome by using even more intense and shorter x-ray pulses, resulting in more violent explosions that are easier to simulate accurately within the current framework, or by refinement of the theoretical model itself. Since both of these aspects are currently under active improvement, we are optimistic in our future ability to map measured correlations in the momentum distribution and the ground-state of the molecule.

In fact, thanks to the linear relationship between the out-of-plane PCs and out-of-plane NMs (section S9.1), a quantitative agreement between the simulation and experiment would in principle enable us to uniquely determine the out-of-plane NMs from the measurement.

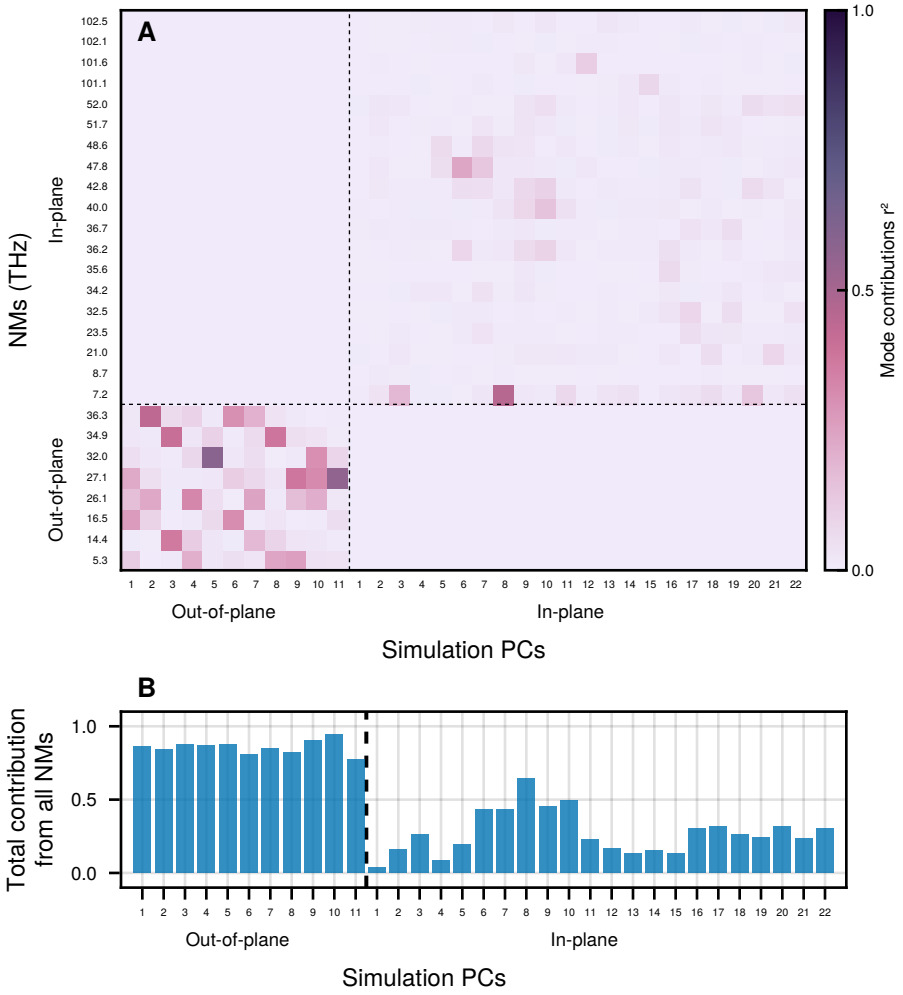

Figure S16: **Contribution of the ground-state NMs to the principal components of the asymptotic momenta (simulation data).** A. Contribution of each NM to each PC. Note that vibrational modes have been grouped into out-of-plane and in-plane modes. B. Total contribution from all NMs to each PC.

S10 Reconstruction with low-charge states

In the main text, we study the $I^{4+}/N^{2+}/C^{2+}/H^+$ coincidence channel. Here, we perform the same analysis on the $I^+/N^+/C^+/H^+$ coincidence channel, and show that this leads to significant discrepancies between experiment and simulation. We applied the fit algorithm to the experimental data with charge states $I^+/N^+/C^+/H^+$, which we analyzed in our previous work (*30*). The resulting leading out-of-plane components are shown in Figs. S17A-S17D.

In order to generate corresponding simulation data, we simulate the explosion using the same parameters as described in Table S2, but with a lower fluence (7.5×10^{10} photons/μm^2), to increase the probability to generate the charge state of interest. The leading out-of-plane PCs that we obtain from the fitting algorithm are shown in Figs. S17E-S17H. We see that, despite using the same data treatment as for the higher charge-state data, the results do not match. This is further demonstrated by the matrix of overlaps, shown in Fig. S17I, which has little in common with the identity matrix that would correspond to perfect agreement with experiment. The only matching PCs are experimental PC 7 and simulation PC 2, stemming from the collision between C2 and I, which was described in Ref. (*40*).

We explain this discrepancy by the presence of non-Coulombic effects that are not included in the simulation, the main one being the impact of the potential energy surfaces of intermediate cationic states, i.e., the residual chemical bonds during the explosion. These residual bonds are not perfectly reproduced and are generally hard to model. In the simulation, we employed a reactive force field that is optimized for a neutral molecule, in combination with Coulomb forces among the ions. The force field is switched off for total molecular charge states of +4 or higher, assuming that Coulomb forces govern the dynamics, and residual chemical bonds do not play a significant role anymore. This approximation is valid if the charge in the molecule is high enough and can be considered the essence of Coulomb explosion imaging.

A sufficiently high final charge in all the atomic ions implies that the charge-up of the molecule occurs fast, such that intermediate charge states in which pure interatomic Coulomb forces do not govern the dynamics are negligible. In this scenario, all electrons relevant to molecular bonding are removed before chemical bonds have time to significantly affect the dynamics of the explosion. Our results indicate that this approximation is valid when highly charged ions, such as those considered

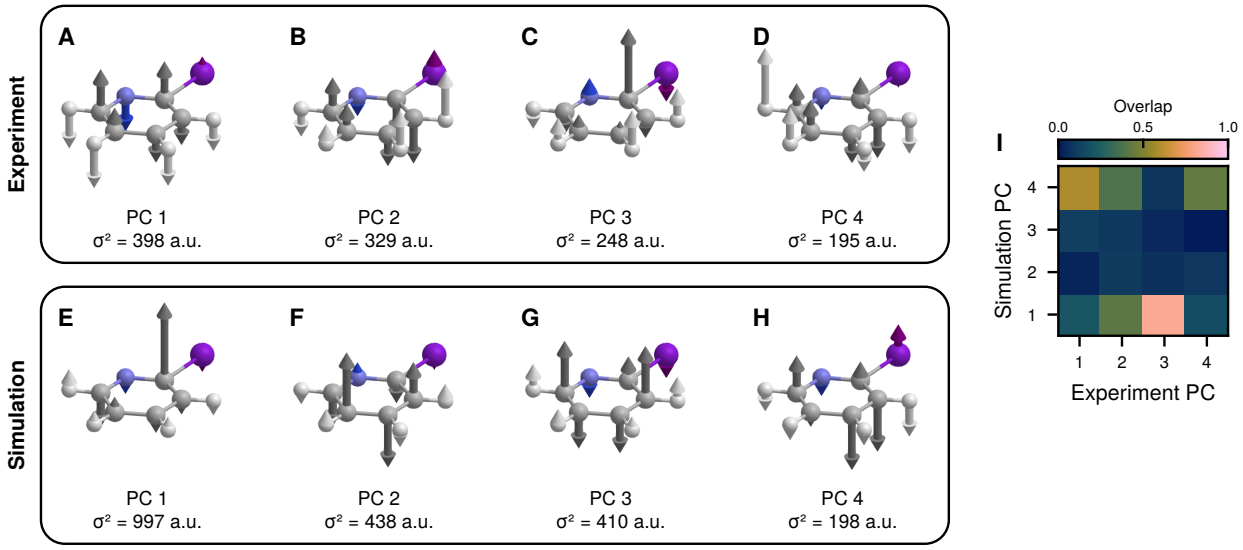

Figure S17: **Principal components of the distribution of asymptotic momenta for lower molecular charge.** A-D. Reconstructed from the experimental data of the coincidence channel $I^+/N^+/C^+/H^+$. E-H. Computed from respective simulated data. σ^2 represents the variance explained by the respective component. I. Overlap between the experiment and simulation principal components. For a perfect agreement, the matrix would be diagonal.

in the main text, are available; but it starts to become invalid for the charge states considered in Fig. S17.

S11 Reproducing the correlation depicted in Fig. 1

In Fig. 1, we show that the H3 proton emission direction is correlated with the emission direction of the carbon ions, when the data is inspected in the iodine-nitrogen recoil frame. We use the model fitted to the experimental data to gain further insight into the source of the observed correlations.

Figure S18A summarizes the dependencies, as found directly by analysing the experimental data, while Fig. S18B shows the correlations according to the fit. To get the latter result, we sampled 100'000 momenta from the fitted Gaussian distribution and rotated them into the iodine-nitrogen frame, before filtering the data based on the emission direction of the H3 proton. The fitted model does not match the experimental data perfectly, in particular the correlation with the C5 carbon ion is exaggerated. Such discrepancies are expected, as the fit is not performed in the recoil frame and cannot be expected to focus on the features that are most visible in this representation of the data. Furthermore, it turns out, the full dimensionality of the fitted model is not required to reproduce the core of the feature shown in Fig. S18A. A projection onto a single dimension, namely along PC 5, already fully describes the overall trends shown in Fig. S18C, meaning that all carbons ions have the correct qualitative behavior. We also look at the correlations of the fitted distribution projected onto the complementary subspace (without PC 5). In this case, as shown in Fig. S18D, the trend is not reproduced, further confirming that the feature observed in Fig. 1H is a consequence of PC 5.

In Figure S18E, we show PC 5 as it is obtained from the fitted model. Note that, since the reconstruction does not rely on the recoil frame, the out-of-plane momenta of the iodine and nitrogen ions are not restricted to the molecular plane. In this representation, the relation to the feature shown in Figs. S18A-C (and Fig. 1H) is not obvious. However, after rotating PC 5 into the recoil frame, the correspondence with the correlated pattern in Fig. S18A becomes clear, as shown in Fig. S18F. The C3 and C4 ion momenta now point towards the same direction as the H3 proton momentum, while the C5 and C6 ion momenta point in the opposite direction. Our analysis reveals that the correlation feature shown in Fig. 1H can be understood from the strong out-of-plane component of the iodine and nitrogen in combination with rotation into the iodine-nitrogen-recoil

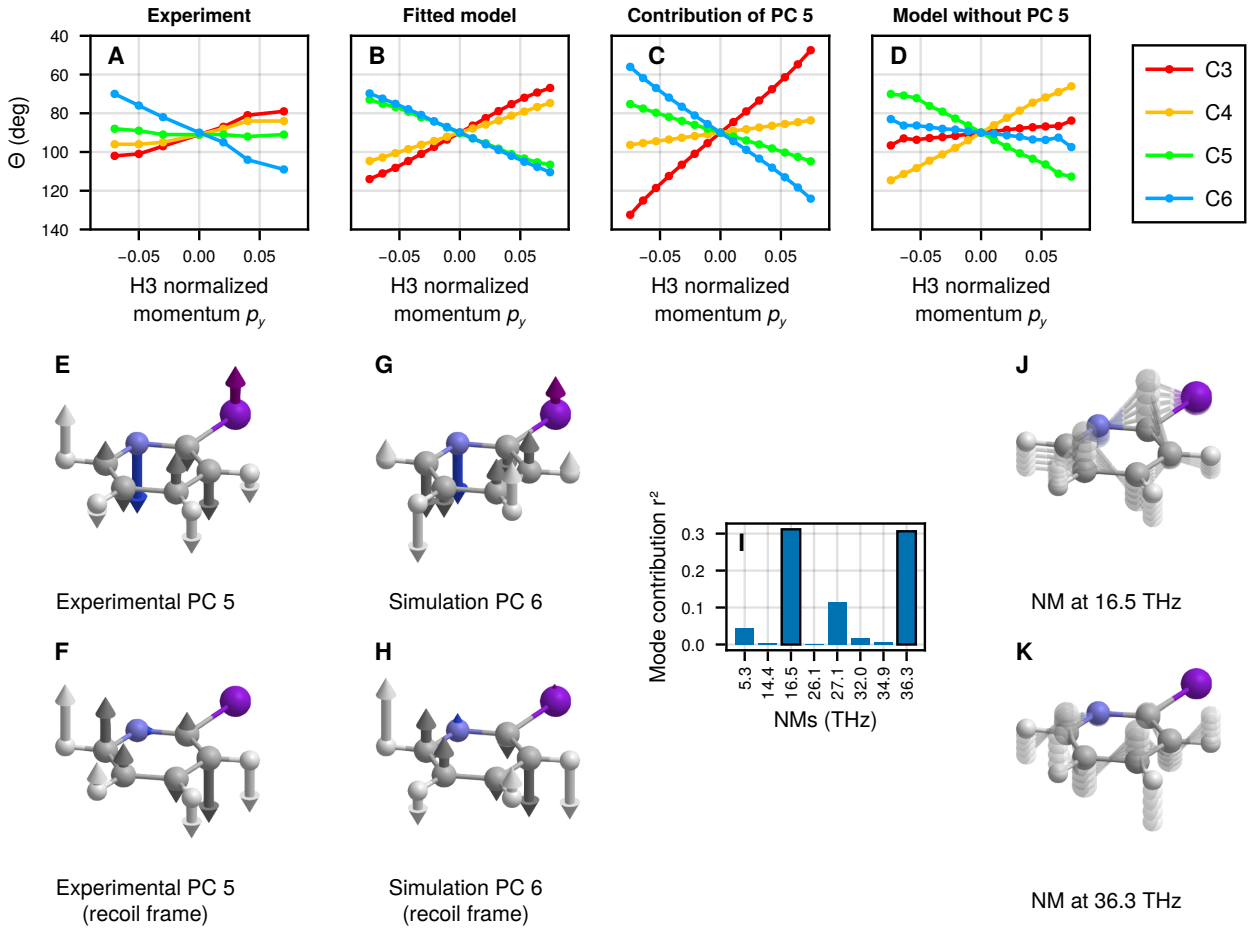

Figure S18: **Explanation of the correlations highlighted in Fig. 1H.** A. Correlations between the emission direction of the H3 ion and the carbon ions from a direct analysis of the experimental data. (reproduced from Fig. 1H). B. Correlations according to the Gaussian model fitted to the experimental data. C. Contribution to the correlations from experimental PC 5. D. Correlations according to the model when PC 5 is removed. E. Experimental PC 5. F. Experimental PC 5 in the iodine-nitrogen recoil frame. G. Simulation PC 6. H. Simulation PC 6 in the iodine-nitrogen recoil frame. I. Contributions of all NMs to simulation PC 6. J-K. Most contributing NMs to simulation PC 6.

frame (see also section S11).

In order to shed further light onto the NMs underlying the correlations observed after the Coulomb explosion in momentum space, we revisit our full Coulomb explosion simulations. We find a single simulation PC with the same strong nitrogen-iodine pattern, namely simulation PC 6, displayed in Fig. S18G. Indeed, experimental PC 5 and simulation PC 6 look very similar when expressed in the recoil frame, as seen in Figs. S18F and S18H. Therefore, we use simulation PC 6 to trace back the possible origin of correlations between the emission direction of the H3 proton and the carbon ions. We compute the contribution of each NM to simulation PC 6, as described in section S9, and plot them in Fig. S18I. The contributions are dominated by two NMs, at 16.5 THz and 36.3 THz, that we show in Figs. S18J and S18K, respectively. This indicates that the correlated features noticed in the experimental data can be rooted back to these two NMs.

We emphasize that arriving at this conclusion relies on the ability to collect multiple ions from the same molecule in coincidence, the strong and fast ionization of the molecules due to the extremely intense x-ray pulses provided by EuXFEL, as well as on the simulation that can satisfactorily reproduce the experiment, and the novel algorithm that we introduce in this work to treat the data.