

Dynamics from noisy data with extreme timing uncertainty

R. Fung^{1*}, A. M. Hanna^{2,3,4}, O. Vendrell^{2,3}, S. Ramakrishna⁵, T. Seideman⁵, R. Santra^{2,3,4,6} & A. Ourmazd^{1*}

Imperfect knowledge of the times at which ‘snapshots’ of a system are recorded degrades our ability to recover dynamical information, and can scramble the sequence of events. In X-ray free-electron lasers, for example, the uncertainty—the so-called timing jitter—between the arrival of an optical trigger (‘pump’) pulse and a probing X-ray pulse can exceed the length of the X-ray pulse by up to two orders of magnitude¹, marring the otherwise precise time-resolution capabilities of this class of instruments. The widespread notion that little dynamical information is available on timescales shorter than the timing uncertainty has led to various hardware schemes to reduce timing uncertainty^{2–4}. These schemes are expensive, tend to be specific to one experimental approach and cannot be used when the record was created under ill-defined or uncontrolled conditions such as during geological events. Here we present a data-analytical approach, based on singular-value decomposition and nonlinear Laplacian spectral analysis^{5–7}, that can recover the history and dynamics of a system from a dense collection of noisy snapshots spanning a sufficiently large multiple of the timing uncertainty. The power of the algorithm is demonstrated by extracting the underlying dynamics on the few-femtosecond timescale from noisy experimental X-ray free-electron laser data recorded with 300-femtosecond timing uncertainty¹. Using a noisy dataset from a pump-probe experiment on the Coulomb explosion of nitrogen molecules, our analysis reveals vibrational wave-packets consisting of components with periods as short as 15 femtoseconds, as well as more rapid changes, which have yet to be fully explored. Our approach can potentially be applied whenever dynamical or historical information is tainted by timing uncertainty.

The fundamental premise of our approach is simple. A series of snapshots concatenated in the order of their inaccurate time stamps will contain some time-evolutionary information (‘a weak arrow of time’), provided that the concatenation window spans a period comparable with, or longer than, the timing uncertainty associated with each individual snapshot. This realization leads one to consider a series of c -fold concatenated snapshots, formed by moving a c -frame-wide window over the raw dataset ordered according to the inaccurate time stamps. The dynamical history can then be extracted from the series of concatenated snapshots using techniques developed to extract signal from noise, such as singular-value decomposition (SVD)⁸. SVD determines a series of statistically significant modes, each consisting of a characteristic pattern (topogram) and its time evolution (chronogram). A topogram can be a characteristic image or spectrum, with the corresponding chronogram showing its change with time. For each mode, a singular value specifies the power contained in that mode⁸.

Consider snapshots, such as images or spectra, that can be represented as vectors by using the pixel values of each snapshot as the components of a vector \mathbf{x} . A snapshot can then be thought of as a point in multidimensional space, and a dataset as a cloud of points in that space.

Similar to principal component analysis, SVD is a linear-algebraic approach, efficiently applicable only when the data cloud defines a flat, low-dimensional hypersurface. Unfortunately, many systems of interest cannot be adequately treated within the framework of linear-algebraic methods such as SVD. Geometrically, data from such systems give rise to intrinsically curved hypersurfaces (manifolds). Fundamental to our approach, therefore, is nonlinear Laplacian spectral analysis (NLSA)⁶, which performs the same analysis as SVD, but on curved manifolds.

For a dataset consisting of a series of N_s time-ordered snapshots, the analysis begins with a ‘time-lagged embedding’^{9–11} to form c -fold concatenated ‘superframes’ (or ‘supervectors’) from the dataset consisting of vectors \mathbf{x} . A typical supervector

$$\mathbf{X}_i = (\mathbf{x}_i; \mathbf{x}_{i-\delta t}; \dots; \mathbf{x}_{i-(c-1)\delta t})$$

is obtained by appending the column vectors $\mathbf{x}_{i-i\delta t}$ ($0 \leq i \leq c-1$) to each other, with $\mathbf{x}_{i-i\delta t}$ representing the i th in the sequence of c snapshots ordered according to time stamps. The time stamp assigned to each of the resulting $(N_s - c)$ supervectors is defined as the mean of the time stamps of its constituent vectors. Unlike averaging, concatenation retains the information content of the dataset⁶; see Supplementary Information section 1.

Next, we use graph-based analysis—specifically the diffusion map algorithm⁵—to identify the nonlinear data manifold formed by the collection of supervectors. The matrix X of supervectors \mathbf{X}_i is then projected (in the sense defined in ref. 6) onto the manifold, to obtain the matrix A

$$A = X\mu\Phi \quad (1)$$

with μ the Riemannian measure of the manifold and Φ the empirical orthogonal eigenfunctions (EOFs)—a truncated set of the eigenfunctions of the Laplace–Beltrami operator on the manifold⁶. This Euclidean description of the nonlinear manifold allows us to analyse the matrix A using standard SVD. The chronograms obtained using SVD are projected from the space defined by Φ back to the time domain, and the topograms corresponding to the superframes are ‘unwrapped’ to obtain individual frames⁶. This approach is able to deal naturally with complex nonlinear dynamics^{6,12}, and to extract conformational information from ultralow-signal snapshots of molecular machines¹³.

Now consider the effect of stochastic timing uncertainty. Recall that the data matrix X is affected by timing uncertainty in two ways: first, the sequence of superframes can differ from the correct, jitter-free case; and, second, the time intervals within the members of a superframe, and those between the superframes themselves, can vary stochastically about a mean. It can be shown that the SVD step in NLSA is immune to jitter-induced changes in the superframe sequence, which are in any case unlikely for large concatenation parameters. As for non-uniformity

¹Department of Physics, University of Wisconsin Milwaukee, 3135 North Maryland Avenue, Milwaukee, Wisconsin 53211, USA. ²Center for Free-Electron Laser Science, DESY, Notkestrasse 85, 22607 Hamburg, Germany. ³The Hamburg Centre for Ultrafast Imaging, Luruper Chausee 149, 22761 Hamburg, Germany. ⁴Department of Chemistry, University of Hamburg, Grindelallee 117, 20146 Hamburg, Germany. ⁵Department of Chemistry, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208, USA. ⁶Department of Physics, University of Hamburg, Jungiusstrasse 9, 20355 Hamburg, Germany.

*These authors contributed equally to this work.

in time sampling, it can be shown analytically and by simulation that, with a sufficient number of snapshots, the outcome of SVD corresponds to time samples that are uniformly spaced to within small oscillations about the mean; see Supplementary Information section 2.

The results of SVD analysis of the matrix A must be projected back into the time domain to reconstruct the dynamics, which involves ‘undoing’ the effect of the projection described in equation (1) by computing $A\Phi^T$ (ref. 6). Because the EOFs represented by Φ are evaluated with our imperfect knowledge of timing, this back-projection re-injects jitter into the results. However, in the limit of large concatenation parameters, the manifold geometry and, hence, the EOFs are biased towards the most stable component of the dynamics⁷, as supported by the reduction in the number of eigenvalues above a spectral gap from five (in the manifold of raw data) to one (after concatenation); see Supplementary Information section 13. One may therefore regard the timing jitter as a form of stochastic forcing, which has been extensively studied¹⁴. In Supplementary Information section 3, we describe how reliable dynamical information can be obtained on timescales substantially shorter than the timing jitter.

For the experimental case analysed below (for which the full-width at half-maximum (FWHM) of the jitter is 300 fs; that is, $\sigma = 120$ fs, $c = 5,800$ and the average time-sample spacing is 50 as), the discussion in Supplementary Information section 3 leads us to expect reliable information on the femtosecond scale. This estimate ignores several important issues, such as the width of the probe pulse and the requirements of Shannon sampling; see Supplementary Information section 3. But it indicates that the approach we have outlined has the potential to reveal dynamics on timescales substantially shorter than the timing uncertainty. The effectiveness of this approach and the guidelines for

its use are outlined in Supplementary Information sections 4–7, with reference to four trial models.

We next demonstrate the capability to obtain dynamical information on timescales much shorter than the timing uncertainty using the noisy experimental data referred to above, which were recorded with substantial jitter stemming from the stochastic nature of the process used to generate ultrashort X-ray pulses in X-ray free-electron lasers¹⁵. The dataset, consisting of 10^5 time-of-flight spectral snapshots spanning a delay time from about -2.7 ps to $+2.3$ ps, was collected in the course of a pump-probe experiment on the Coulomb explosion of nitrogen molecules; see Supplementary Information section 12. During the experiment, an infrared pulse of approximately 60 fs in length either preceded or succeeded an ultrashort (< 10 fs) X-ray pulse generated by the X-ray free-electron laser at the Linac Coherent Light Source¹. Each time-of-flight spectrum (with a signal-to-noise ratio of approximately 0.16; see Supplementary Information section 10) is a record of the dynamics of about 30 molecular ions. The experimental approach and conditions are described in ref. 1.

The data from this experiment span three different regimes: first, that in which the X-ray pulse preceded the infrared pulse (the ‘X-ray-first’ regime); second, that in which the X-ray pulse succeeded the infrared pulse (the ‘infrared-first’ regime); and, third, that in which the ultrashort X-ray pulse arrives while the infrared pulse is active and so the two overlap (the ‘overlap’ regime). In the X-ray-first regime, the X-ray pulse ionizes (and/or dissociates) the N_2 molecules, and the infrared pulse breaks up the quasi-bound molecular ions it encounters, which are then collected by the time-of-flight spectrometer¹. In the infrared-first regime, the infrared pulse induces impulsive alignment¹⁶, and can also ionize the molecules, which are then dissociated by the X-ray

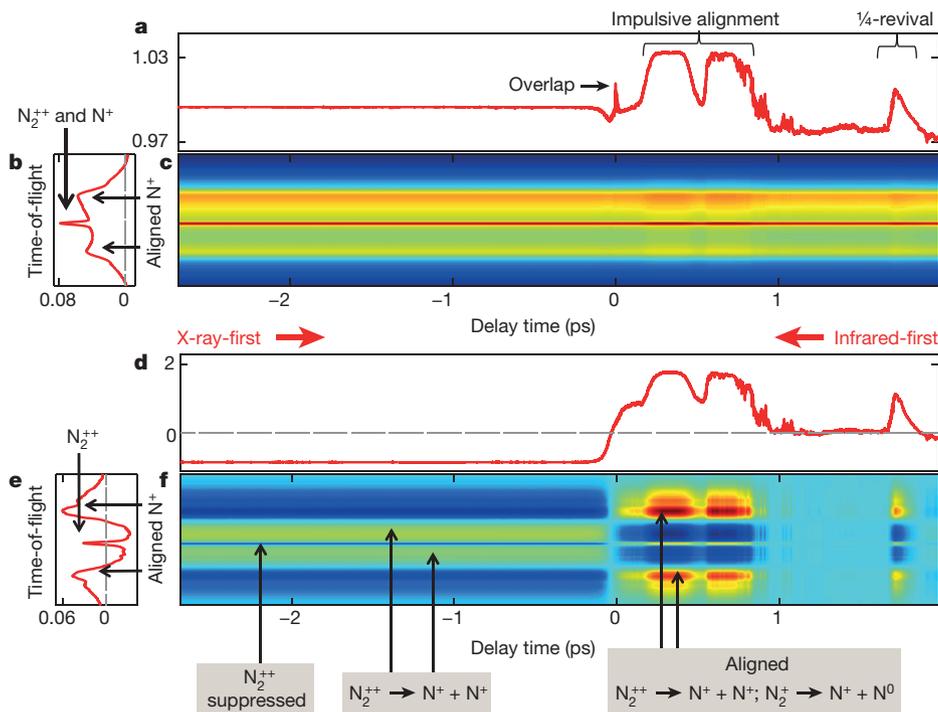


Figure 1 | Coulomb explosion of N_2 , singular modes 1 and 2. **a**, Mode-1 chronogram showing the time evolution of the average spectrum. The vertical axis shows the number of molecular fragments reaching the detector in arbitrary units. Note the three different regimes: X-ray-first, overlap and infrared-first. The effects of impulsive orientational alignment induced by the infrared pulse (including the 1/4-revival) are indicated. A sharp feature marks the overlap region, where the X-ray and infrared pulses overlap. **b**, Mode-1 time-of-flight spectrum showing the region around a mass-to-charge ratio of 14 (the horizontal axis shows the signal strength in arbitrary units). The sharp central peak is due to N_2^{++} dications. The side-peaks stem from N^+ ions ejected towards and away

from the detector. **c**, Mode-1 reconstructed time-of-flight series of spectral frames showing the evolution of the average spectrum with pump-probe delay time (the colour scale indicates the signal strength: lowest signal, dark blue; highest signal, red). **d**, Mode-2 chronogram showing the time evolution of the mode; vertical axis as in **a**. An error-function-like feature spans the region in which the X-ray and infrared pulses overlap. **e**, Mode-2 time-of-flight spectrum showing the region around a mass-to-charge ratio of 14 (horizontal axis as in **b**). **f**, Mode-2 reconstructed time-of-flight series of spectral frames (colour scale as in **c**). In the X-ray-first region, note the strong suppression of the central peak due to N_2^{++} dications. The effect of impulsive alignment is evident in the infrared-first regime.

pulse. This experiment is typical of the class of stroboscopic pump-probe approaches that are widely used to investigate the temporal evolution of ultrafast processes¹⁷. Because such phenomena can be heavily obscured by timing uncertainty, the dataset used here represents a critical test of our data-analytic approach.

Below, we show that our analysis successfully extracts known ultrafast phenomena, including the molecular vibrations of the N₂ system with periods as short as 15 fs. We also pinpoint the start and end of the infrared pulse to about 1 fs, reveal enhanced molecular dissociation during the infrared pulse, and report anticipated, but previously unobserved, wave packets excited by the X-ray pulse. These results do not depend sensitively on the specific values used for the algorithmic parameters; see Supplementary Information section 11. The concatenation window c is the most important parameter; the quality of the results improves steadily with increasing c until the concatenation window is comparable with the FWHM of the timing uncertainty.

Turning to the analysis, each of the 10⁵ spectral snapshots including regions around mass-to-charge ratios of 7 and 14 was represented as a vector, with components consisting of the signal recorded in the pixels of the time-of-flight spectrum. The vectors were ordered according to the jitter-corrupted experimental time stamps, and concatenated to form supervectors. Because the jitter substantially exceeded the 50-as average interval between successive snapshots, the sequence of vectors was strongly compromised, and the concatenation order purely statistical. The results reported below were obtained with a 5,800-fold concatenation window ($c = 5,800$) spanning 290 fs. The diffusion map algorithm⁵ was used to investigate the intrinsic structure of the concatenated data. The resulting manifold was five-dimensional (as determined by the procedure outlined in ref. 18) and nonlinear. This data structure (Supplementary Fig. 9) precludes analysis using standard linear-algebraic means such as SVD; see Supplementary Information section 14.

NLSA of the experimental data reveals the presence of up to six modes with singular values above the noise plateau (Supplementary Fig. 10). As in standard SVD, the first topogram constitutes the mean, with the subsequent modes representing the various deviations from it. Each chronogram shows the time evolution of its respective topogram in the X-ray-first, infrared-first and overlap regimes. Chronograms describing the time evolution of the signal and of the time-of-flight spectra for the first four modes are shown in Figs 1 and 2.

As in standard SVD, unless the physical processes at work are independent and non-degenerate, a single mode need not represent a complete physical process. Therefore, extraction of the individual physical processes requires additional information. However, the measured behaviour of the system is a linear combination of the modes obtained by data analysis. Thus, the features revealed by each mode constitute key elements of the processes at work and provide insight into the behaviour of the system.

With the above caveat in mind, we now discuss the key features of each of the modes obtained by our analysis (Figs 1 and 2). In the infrared-first regime, the modes reveal well-known features associated with the impulsive orientational alignment of N₂ molecules, with successive modes capturing the average and higher moments of the aligned distribution¹⁶.

In all modes, clear features mark the time span during which the infrared and X-ray pulses overlap. The sharp turning points flanking the overlap region in mode 3 (Fig. 3) are separated by 36 ± 2 fs. We associate this time span with the period during which the infrared pulse was sufficiently intense to affect the process reflected in this mode in the overlap regime.

We now turn to specific features in modes 3 and 4 (Fig. 3), which shed light on the behaviour of the system in the overlap regime, where the infrared pulse is active. Mode 3 concerns the detection of N₂⁺⁺, while mode 4 reveals the collection of two N⁺ ions, one ejected towards

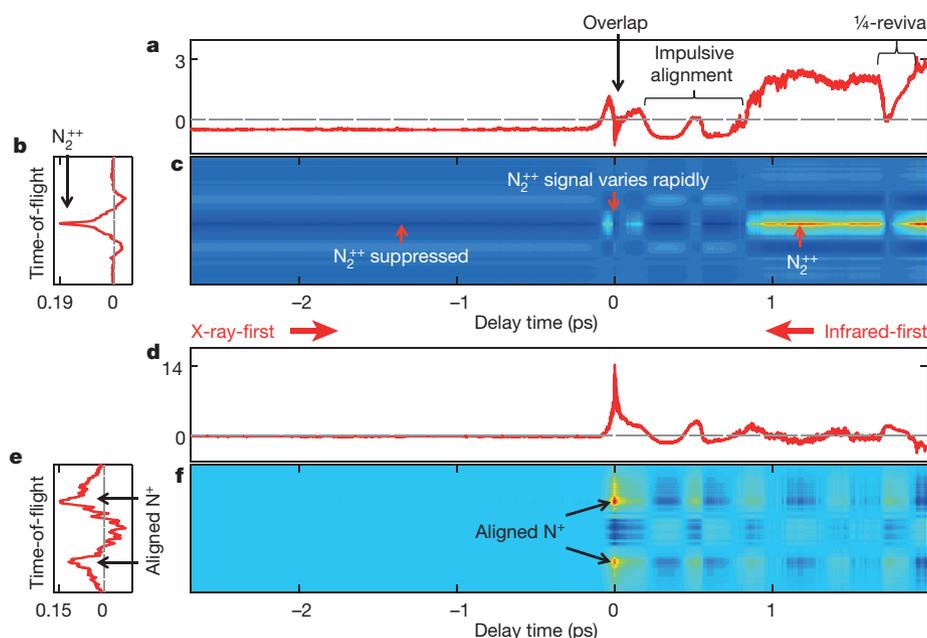


Figure 2 | Coulomb explosion of N₂, singular modes 3 and 4. **a**, Mode-3 chronogram showing the time evolution of the mode. The vertical axis shows the number of molecular fragments reaching the detector in arbitrary units. Note the three different regimes: X-ray-first, overlap and infrared-first. A sharp feature marks the overlap region, where the X-ray and infrared pulses overlap. **b**, Mode-3 time-of-flight spectrum showing the region around a mass-to-charge ratio of 14 (the horizontal axis shows the signal strength in arbitrary units). The sharp central peak stems from N₂⁺⁺. **c**, Mode-3 reconstructed time-of-flight series of spectral frames (the colour scale indicates the signal strength: lowest signal, dark blue; highest

signal, red). The N₂⁺⁺ signal is suppressed in the X-ray-first region, and varies rapidly in the overlap region. **d**, Mode-4 chronogram showing the time evolution of the mode; vertical axis as in **a**. A sharp feature marks the overlap region, in which the X-ray arrives while the infrared pulse is active. **e**, Mode-4 time-of-flight spectrum for mass-to-charge ratio of 14 (horizontal axis as in **b**). **f**, Mode-4 reconstructed time-of-flight series of spectral frames (colour scale as in **c**). A weak signal in the X-ray-first regime is invisible on this scale, but is clearly seen in Fig. 4. In the overlap region, the dissociation of molecules aligned with the polarization vector of the infrared pulse produces the two red peaks.

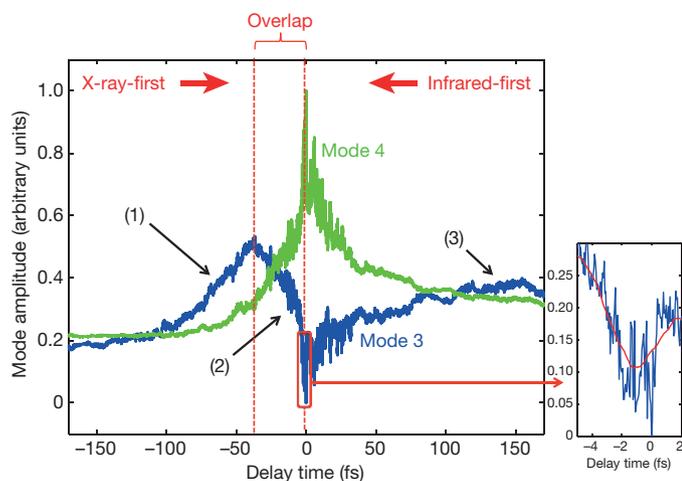


Figure 3 | Infrared/X-ray overlap region as revealed by singular modes 3 and 4. Sharp extrema in mode 3 (N_2^{++} dication signal) mark the start and end of the infrared pulse. The inset demonstrates that the extrema can be located to within ± 1 fs. The pulse width is thus 36 ± 2 fs at the intensity needed to strongly suppress the dication signal. The anticorrelation between modes 3 and 4 in the overlap region indicates that the corresponding N_2^{++} and N^+ channels are competing. The black arrows indicate: (1) dissociation of X-ray-generated N_2^{++} dications by the subsequent infrared pulse; (2) N_2^{++} suppression by the infrared pulse; and (3) build-up of infrared-generated molecular ions.

the detector and one ejected away from it. In the overlap region, modes 3 and 4 are strongly anticorrelated (with a correlation coefficient of -0.9917), indicating they are competing dissociation channels. It is tempting to associate the enhanced ejection of N^+ towards and away from the detector with enhanced dissociation of N_2^{++} along the polarization vector of the infrared pulse (Fig. 2f), as observed in strong-field experiments using optical pump and probe pulses under precisely controlled conditions^{19–22}. But the detector geometry used to obtain our experimental data strongly suppresses the detection of dissociation fragments ejected perpendicular to the detector axis. For this reason, we cannot reach a definitive conclusion on whether the dissociation is preferentially enhanced along the electric-field vector of the infrared pulse. Our results nonetheless highlight the type of detailed information yielded by our approach.

A key finding from our analysis concerns the observation of wave-packet dynamics in the X-ray-first regime (Fig. 4). In the majority of cases, the X-ray pulse creates core holes, which decay rapidly to form molecular dication states²³. In a minority of X-ray absorption events, a valence electron is ejected to produce N_2^+ ions. Abrupt events are expected to launch vibrational^{17,24,25} and/or charge²⁶ wave-packet dynamics. As shown in Fig. 4a, b, our analysis clearly reveals the presence of wave-packet oscillations and their revival in the X-ray-first regime. The periods of these oscillations (Fig. 4c, d) coincide closely with the known vibrations of the N_2^+ and N_2^{++} systems^{25,27}. However, only the 15-fs oscillation has been previously accessed in the time domain²⁵, with the other oscillation periods deduced from spectroscopic measurements²⁷. We also observe oscillations outside the well-studied 40–70-THz range (Supplementary Fig. 7 and Supplementary Information section 8). Initial results from a quantum-mechanical calculation corroborate the spectral features we extract data-analytically in the 10–70-THz frequency range (Supplementary Information section 9).

In the absence of a well-characterized physical process with a temporal knife-edge of sufficient abruptness, it is not straightforward to determine the exact time resolution achieved by our approach. Our results identify the start and end of the optical pulse, each to within about 1 fs (Fig. 3 inset). Given the width of the X-ray pulse (< 10 fs, possibly shorter than 6 fs; ref. 1), it seems surprising that such precise information can be obtained. Other things being equal, the time

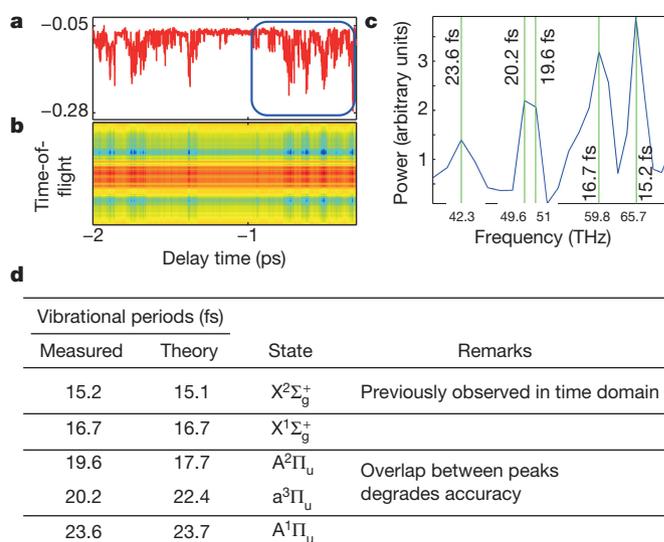


Figure 4 | Vibrational wave packets, singular mode 4, X-ray-first region. **a**, Chronogram in the X-ray-first region, showing the time evolution of a wave packet and its revival (the vertical axis shows the signal strength in the corresponding topogram in arbitrary units). **b**, Reconstructed time-of-flight series of spectral frames in the X-ray-first region (the colour scale indicates the signal strength: lowest signal, dark blue; highest signal, red). **c**, Fourier spectrum (blue) of the chronogram in the 690-fs-long boxed region in **a**, showing the frequency and period (green lines and labels) of each component. **d**, Comparison of observed periods with known vibrational modes of N_2 .

resolution is determined by the signal expected from an infinitely short pulse, convolved with the actual probe envelope. In agreement with previous work on partially coherent optical pulses²⁸, our results suggest that the X-ray pulse contains spikes that are sufficiently narrow to allow extraction of information beyond the nominal pulse envelope. The time resolution also depends on several other experimental parameters, including the timing uncertainty, the magnitude and uniformity of the time interval between snapshots²⁹, their signal-to-noise ratio, and the characteristics of the signal itself. Using modest computing resources (see Supplementary Information section 15), the present demonstration was achieved with 10^5 spectral snapshots with a signal-to-noise ratio of approximately 0.16, covering a time span approximately 17 times the FWHM of the timing uncertainty. Finally, the application of NLSA is essential even when the system under consideration is intrinsically linear. This stems from the large size of the matrix of concatenated data—containing about 10^{12} elements for the experimental data treated here—which greatly exceeds that amenable to standard SVD.

In summary, we have demonstrated a purely data-analytical approach that is capable of extracting the evolution and dynamics of complex systems from noisy snapshots on timescales much shorter than the uncertainty with which the data were recorded. We expect our approach to have a broad impact in many areas of science and technology; examples include geology and climate science, where timing of events can be uncertain, chemistry and biology, where reaction initiation can be non-uniform across a sample, and signal processing, where noise and timing jitter are prevalent.

Received 13 August 2015; accepted 18 February 2016.

- Glowia, J. M. *et al.* Time-resolved pump-probe experiments at the LCLS. *Opt. Express* **18**, 17620–17630 (2010).
- Gahl, C. *et al.* A femtosecond X-ray/optical cross-correlator. *Nature Photon.* **2**, 165–169 (2008).
- Löhl, F. *et al.* Electron bunch timing with femtosecond precision in a superconducting free-electron laser. *Phys. Rev. Lett.* **104**, 144801 (2010).
- Harmand, M. *et al.* Achieving few-femtosecond time-sorting at hard X-ray free-electron lasers. *Nature Photon.* **7**, 215–218 (2013).
- Coifman, R. R. *et al.* Geometric diffusions as a tool for harmonic analysis and structure definition of data: diffusion maps. *Proc. Natl Acad. Sci. USA* **102**, 7426–7431 (2005).

6. Giannakis, D. & Majda, A. J. Nonlinear Laplacian spectral analysis for time series with intermittency and low-frequency variability. *Proc. Natl Acad. Sci. USA* **109**, 2222–2227 (2012).
7. Berry, T., Cressman, R., Gregurić-Ferenček, Z. & Sauer, T. Time-scale separation from diffusion-mapped delay coordinates. *SIAM J. Appl. Dyn. Syst.* **12**, 618–649 (2013).
8. Aubry, N., Guyonnet, R. & Lima, R. Spatiotemporal analysis of complex signals: theory and applications. *J. Stat. Phys.* **64**, 683–739 (1991).
9. Packard, N., Crutchfield, J., Farmer, J. & Shaw, R. Geometry from a time series. *Phys. Rev. Lett.* **45**, 712–716 (1980).
10. Takens, F. in *Dynamical Systems and Turbulence, Warwick 1980* (eds Rand, D. A. & Young, L.-S.) Vol. 898 *Lecture Notes in Mathematics* (eds Dold, A. & Eckmann, B.) 366–381 (Springer, 1981).
11. Sauer, T., Yorke, J. A. & Casdagli, M. Embedology. *J. Stat. Phys.* **65**, 579–616 (1991).
12. Giannakis, D. & Majda, A. J. Comparing low-frequency and intermittent variability in comprehensive climate models through nonlinear Laplacian spectral analysis. *Geophys. Res. Lett.* **39**, L10710 (2012).
13. Dashti, A. *et al.* Trajectories of the ribosome as a Brownian nanomachine. *Proc. Natl Acad. Sci. USA* **111**, 17492–17497 (2014).
14. Stark, J., Broomhead, D. S., Davies, M. E. & Huke, J. Delay embeddings for forces systems. II. Stochastic forcing. *J. Nonlinear Sci.* **13**, 519–577 (2003).
15. Pellegrini, C. & Stohr, J. X-ray free-electron lasers—principles, properties and applications. *Nucl. Instrum. Methods Phys. Res. A* **500**, 33–40 (2003).
16. Stapelfeldt, H. & Seideman, T. *Colloquium*: aligning molecules with strong laser pulses. *Rev. Mod. Phys.* **75**, 543–557 (2003).
17. Rudenko, A. *et al.* Real-time observation of vibrational revival in the fastest molecular system. *Chem. Phys.* **329**, 193–202 (2006).
18. Coifman, R. R., Shkolnisky, Y., Sigworth, F. J. & Singer, A. Graph Laplacian tomography from unknown random projections. *IEEE Trans. Image Process.* **17**, 1891–1899 (2008).
19. Schmidt, M. *et al.* Fragment-emission patterns from the Coulomb explosion of diatomic molecules in intense laser fields. *Phys. Rev. A* **60**, 4706–4714 (1999).
20. Voss, S. *et al.* High resolution kinetic energy release spectra and angular distributions from double ionization of nitrogen and oxygen by short laser pulses. *J. Phys. B* **37**, 4239–4257 (2004).
21. Pavičić, D., Lee, K. F., Rayner, D. M., Corkum, P. B. & Villeneuve, D. M. Direct measurement of the angular dependence of ionization for N₂, O₂, and CO₂ in intense laser fields. *Phys. Rev. Lett.* **98**, 243001 (2007).
22. Guo, W., Zhu, J., Wang, B., Wang, Y. & Wang, L. Angular distributions of fragment ions of N₂ in a femtosecond laser field. *Phys. Rev. A* **77**, 033415 (2008).
23. Eberhardt, W., Stohr, J., Feldhaus, J., Plummer, E. W. & Sette, F. Correlation between electron emission and fragmentation into ions following soft-X-ray excitation of the nitrogen molecule. *Phys. Rev. Lett.* **51**, 2370–2373 (1983).
24. Bocharova, I. A. *et al.* Time-resolved Coulomb-explosion imaging of nuclear wave-packet dynamics induced in diatomic molecules by intense few-cycle laser pulses. *Phys. Rev. A* **83**, 013417 (2011).
25. De, S. *et al.* Following dynamic nuclear wave packets in N₂, O₂, and CO with few-cycle infrared pulses. *Phys. Rev. A* **84**, 043410 (2011).
26. Timmers, H. *et al.* Coherent electron hole dynamics near a conical intersection. *Phys. Rev. Lett.* **113**, 113003 (2014).
27. Dawber, G. *et al.* Threshold photoelectrons coincidence spectroscopy of doubly-charged ions of nitrogen, carbon monoxide, nitric oxide and oxygen. *J. Phys. B* **27**, 2191–2209 (1994).
28. Meyer, K. *et al.* Noisy optical pulses enhance the temporal resolution of pump-probe spectroscopy. *Phys. Rev. Lett.* **108**, 098302 (2012).
29. Jerri, A. J. The Shannon sampling theorem—its various extensions and applications: a tutorial review. *Proc. IEEE* **65**, 1565–1596 (1977).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank P. Bucksbaum, J. M. Glowina and A. Natan for experimental data and comments on an earlier version of the manuscript, and acknowledge discussions with A. Dashti, D. Giannakis, A. Hosseinizadeh, A. Rudenko, M. Schmidt and P. Schwander. The research conducted by A.O. and R.F. was supported by the US Department of Energy, Office of Science, Basic Energy Sciences under award DE-SC0002164 (algorithm design and development, and data analysis), and by the US National Science Foundation under awards STC 1231306 (numerical trial models) and 1551489 (underlying analytical models). The research conducted by T.S. and S.R. was supported by the US Department of Energy, Office of Science, Basic Energy Sciences under award DE-FG02-04ER15612. T.S. thanks the Hamburg Centre for Ultrafast Imaging for a Mildred Dresselhaus Visiting Professorship.

Author Contributions A.O. proposed the approach. R.F. and A.O. developed the algorithm architecture. R.F. implemented the algorithm and, together with A.O., obtained results from experimental data. A.O. and R.S. interpreted the experimental results. S.R. and T.S. performed simulations of impulsive molecular alignment and provided expert advice. A.M.H., O.V. and R.S. performed quantum-mechanical calculations.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.O. (Ourmazd@uwm.edu).