

Two-Dimensional Partial-Covariance Mass Spectrometry of Large Molecules Based on Fragment Correlations

Taran Driver¹, Bridgette Cooper¹, Ruth Ayers¹, Rüdiger Pipkorn², Serguei Patchkovskii³, Vitali Averbukh¹, David R. Klug⁴, Jon P. Marangos¹, Leszek J. Frasinski¹, and Marina Edelson-Averbukh^{1,*}

¹*The Blackett Laboratory Laser Consortium, Department of Physics, Imperial College London, London, SW7 2AZ, United Kingdom*

²*German Cancer Research Centre, Department of Translational Immunology, INF 580, 69120 Heidelberg, Germany*

³*Max Born Institute for Nonlinear Optics and Short Pulse Spectroscopy, Max-Born-Straße 2A, 12489 Berlin, Germany*

⁴*Department of Chemistry, Imperial College London, London, SW7 2AZ, United Kingdom*

 (Received 5 February 2020; revised 5 May 2020; accepted 28 July 2020; published 6 October 2020)

Covariance mapping [L. J. Frasinski, K. Codling, and P. A. Hatherly, *Science* **246**, 1029 (1989)] is a well-established technique used for the study of mechanisms of laser-induced molecular ionization and decomposition. It measures statistical correlations between fluctuating signals of pairs of detected species (ions, fragments, electrons). A positive correlation identifies pairs of products originating from the same dissociation or ionization event. A major challenge for covariance-mapping spectroscopy is accessing decompositions of large polyatomic molecules, where true physical correlations are overwhelmed by spurious signals of no physical significance induced by fluctuations in experimental parameters. As a result, successful applications of covariance mapping have so far been restricted to low-mass systems, e.g., organic molecules of around 50 daltons (Da). Partial-covariance mapping was suggested to tackle the problem of spurious correlations by taking into account the independently measured fluctuations in the experimental conditions. However, its potential has never been realized for the decomposition of large molecules, because in these complex situations, determining and continuously monitoring multiple experimental parameters affecting all the measured signals simultaneously becomes unfeasible. We introduce, through deriving theoretically and confirming experimentally, a conceptually new type of partial-covariance mapping—self-correcting partial-covariance spectroscopy—based on a parameter extracted from the measured spectrum itself. We use the readily available total ion count as the self-correcting partial-covariance parameter, thus eliminating the challenge of determining experimental parameter fluctuations in covariance measurements of large complex systems. The introduced self-correcting partial covariance enables us to successfully resolve correlations of molecules as large as 10^3 – 10^4 Da, 2 orders of magnitude above the state of the art. This opens new opportunities for mechanistic studies of large molecule decompositions through revealing their fragment-fragment correlations. Moreover, we demonstrate that self-correcting partial covariance is applicable to solving the inverse problem: reconstruction of a molecular structure from its fragment spectrum, within two-dimensional partial-covariance mass spectrometry.

DOI: [10.1103/PhysRevX.10.041004](https://doi.org/10.1103/PhysRevX.10.041004)

Subject Areas: Atomic and Molecular Physics,
Chemical Physics,
Interdisciplinary Physics

I. INTRODUCTION

The spectroscopic study of the structure and dynamics of large polyatomic molecules poses a formidable experimental challenge, owing to spectral complexity and spectral congestion. A solution to this challenge comes from multidimensional spectroscopic techniques, which reveal correlations between spectral signals, decongest the spectra, and provide supreme capabilities both for studying the mechanisms of molecular processes and for deducing

*Corresponding author.
m.edelson-averbukh@imperial.ac.uk

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

molecular structure. In our work, we develop a new type of two-dimensional (2D) partial-covariance spectroscopy applicable to fragmentations of large molecules, including polypeptide and full protein molecules well over 10 kilodaltons (kDa), thereby increasing the size of the systems amenable to covariance-mapping spectroscopy by 2 orders of magnitude. Contrary to the current state of the art, the new partial-covariance mapping extracts the spectral correlations using only the measured spectrum itself, without relying on any additional measurement channels.

A. Covariance-mapping spectroscopy

Covariance-mapping spectroscopy was developed by Frasinski *et al.* as a tool for the study of mechanisms of radiation-induced fragmentation of di- and triatomic molecules [1]. The technique is based on measuring the covariance $\text{Cov}(X, Y)$ between the intensities of every pair of spectral signals X and Y across a series of ionic time-of-flight spectra,

$$\text{Cov}(X, Y) = \langle XY \rangle - \langle X \rangle \langle Y \rangle, \quad (1)$$

where angular brackets denote averaging over multiple spectra. If a pair of spectral signals X and Y are characterized by positive covariance, their intensities fluctuate synchronously across different spectra. This correlated fluctuation indicates that the fragments X and Y originate from the same decomposition process of the parent ion Z : $Z \rightarrow X + Y$, e.g., $\text{CO}^{2+} \rightarrow \text{C}^+ + \text{O}^+$. More recently, covariance mapping has found application in x-ray free-electron-laser (XFEL) experiments and has been shown to be effective, for example, in unraveling the decomposition mechanisms of “hollow atoms” (unstable states of matter formed by inner shell ionization by intense x-ray irradiation [2]), study of double-core-hole states in molecules [3], correlating photoelectron emission with the fragmentation of hydrocarbons in intense infrared laser fields [4], and characterizing the degree of alignment of molecules in the gas phase [5].

In many covariance-mapping experiments, the physically informative fragment-fragment correlations are overwhelmed by spurious correlations of no physical significance (“extrinsic correlations”). The extrinsic correlations stem from fluctuations of experimental parameters, such as the laser-pulse intensity in laser-induced decomposition experiments. These fluctuations lead to the simultaneous increase or decrease of all measured signals resulting in extrinsic correlation of every spectral signal with every other one [6]. The solution to such physical situations comes from partial-covariance (pCov) mapping, where these extrinsic correlation signals can be mathematically removed if the fluctuations of the parameters causing them are experimentally recorded. For example, if one is able to continuously monitor a series of independently fluctuating parameters I, J, \dots affecting the correlations, the partial covariance is given by

$$\begin{aligned} \text{pCov}(X, Y; I, J, \dots) = & \text{Cov}(X, Y) - \frac{\text{Cov}(X, I)\text{Cov}(Y, I)}{\text{Cov}(I, I)} \\ & - \frac{\text{Cov}(X, J)\text{Cov}(Y, J)}{\text{Cov}(J, J)} - \dots, \quad (2) \end{aligned}$$

where $\text{Cov}(I, I), \text{Cov}(J, J), \dots$ are the variances of the fluctuating parameters [7]. If the fluctuating parameters I, J, \dots are not fully independent, the partial-covariance formula involves the inverse of their variance-covariance matrix [7]. Since in many laser-induced ionization and fragmentation experiments the overwhelmingly dominant fluctuating parameter is the laser-pulse intensity, Eq. (2) can often be employed with the laser intensity as a single parameter. Indeed, laser-intensity-based partial covariance has enabled the study of x-ray-induced fragmentation at XFELs under significant intensity fluctuations of the ionizing pulse [5,6]. However, application of partial-covariance spectroscopy to laser-induced molecular fragmentation has been restricted to relatively small systems, well within the 100-Da limit.

B. Bioanalytical mass spectrometry

Mass spectrometry (MS) is the method of choice for the structural analysis of biomolecules, such as proteins [8,9], nucleic acids, lipids, and metabolites. The most common bioanalytical application of MS analysis is identification of protein primary structure, i.e., the sequence of protein’s molecular building blocks (amino acids) and chemical groups attached to them, e.g., post-translational modifications (PTMs) [10]. To identify the primary structure of a protein, it is typically first cut into smaller fragments, e.g., using enzymes, to obtain peptides which are subsequently sent to a mass spectrometer via soft ionization techniques [electrospray ionization (ESI) [11] or matrix-assisted laser desorption ionization [12]]. The resulting ionized peptides can be further fragmented to produce detailed structural information by fragmentation methods such as collision-induced dissociation (CID) [13], electron-transfer dissociation (ETD) [14], electron-capture dissociation (ECD) [15], as well as various laser-based techniques [16–18]. The mass-to-charge ratios (m/z) of the intact peptides and the fragment ions are most commonly determined using ion-trap or time-of-flight techniques [19]. Finally, all the spectral information (mass spectra of both peptide ions and their fragments) is pieced together to deduce the sequence of the original protein, using a range of structure-to-spectrum matching methods [20–24].

The success of the analytical MS experiment is determined by the certainty with which it attributes the experimental data to the correct primary structure. Mainstream one-dimensional (1D) MS is striving to increase the fidelity of this structure-to-spectrum assignment through two main avenues: (i) the development of new fragmentation methods [14–18] in order to maximize the variety of

structure-specific molecular fragments generated in an experiment and (ii) improving the accuracy and resolution of the mass-to-charge measurement [25,26] in order to distinguish between peptides and fragments of very similar m/z values. Typically, however, a significant fraction of the measured fragment mass spectra fails to be successfully interpreted and matched to the correct peptide and protein sequences [25].

A two-dimensional MS technique, 2D Fourier-transform ion cyclotron resonance (2D FT ICR), which is based on the physical principle identical to that of the 2D exchange NMR spectroscopy, was previously introduced [27,28]. The method provides an analytical advantage through correlating intact peptides with their fragments to avoid the losses associated with isolation of individual unfragmented peptide ions. However, 2D FT ICR does not generate any new physical information that is not available through the 1D MS analysis. Here we report 2D mass spectrometry based on a different physical principle: partial-covariance mapping, which generates an entirely new channel of physical information—fragment-fragment correlations.

II. SELF-CORRECTING PARTIAL COVARIANCE: THE PHYSICAL CONCEPT

Large polyatomic molecules pose a particular challenge for the state-of-the-art covariance-mapping techniques. The complex methods of delivery of large molecules to the gas phase [11,12], as well as the ion-transfer and -fragmentation methods [13–18] involve a rich manifold of fluctuating experimental parameters (e.g., electrospray intensity, ion-focusing voltages, ion-trap voltages, gas pressure) producing numerous sources of physically meaningless extrinsic correlations. Identifying and continuously monitoring fluctuations of such experimental parameters, as required by the state-of-the-art partial-covariance spectroscopy, becomes unfeasible.

Here we present a conceptually new type of partial-covariance mapping—self-correcting partial covariance—which in contrast to any available technique does not require any separate detection channels. Instead, as we demonstrate theoretically and experimentally, the self-correcting partial covariance removes extrinsic correlations using the information which is already present in the measured fragment spectrum itself. This is done using a single parameter derived from the fragment intensities $I = f(\vec{X})$ [cf. Eq. (2)], where \vec{X} is the vector of the measured intensities of individual fragment ions and f is a finite, nonvanishing function. In this work, we choose the single parameter to be the total integrated intensity of an individual spectrum $f(\vec{X}) = \sum_i X_i$, which in the case of MS measurement, translates into the universally available total ion count (TIC). We call the resulting new type of mass spectrometric measurement two-dimensional partial-covariance mass spectrometry (2D-PC-MS).

The basic principle of 2D-PC-MS is illustrated in Fig. 1. The inherent spectrum-to-spectrum fluctuations in the intensities of fragments born in the same [pair of blue signals in Fig. 1(a)] and consecutive [pair of red signals in Fig. 1(a)] reactions follow each other (correlate positively) in a series of individual spectra [see Fig. 1(b)]. 2D-PC-MS identifies such fragment pairs by calculating the TIC partial covariance between the intensities of each pair of fragments (X, Y) across multiple fragment spectra. The identified correlating fragment pairs (X, Y) are connected by positive islands on the 2D-PC-MS map, while the extrinsic correlations are suppressed; see Fig. 1(c). The map is symmetric with respect to the $x = y$ autocorrelation diagonal [gray dashed line in Fig. 1(c)], since if ion X correlates with ion Y , then ion Y necessarily correlates with ion X . Peaks falling on the autocorrelation diagonal represent the trivial correlations of each spectral signal with itself. These peaks are removed for clarity from the experimental plots below.

The physical idea behind the TIC-based self-correcting partial-covariance mapping is that the compound effect of all fluctuations in the experimental conditions is reflected in the total number of fragment ions in an individual fragment spectrum. Since all the fragment signals correlate with the intense fluctuations of the TIC, the resulting strong extrinsic correlations overwhelm the intrinsic fragment-fragment correlations. Once the correlations of fragment intensities with TIC are removed, the physical intrinsic correlations of the fragments with each other are recovered.

III. TOTAL ION COUNT FOR SELF-CORRECTING PARTIAL-COVARIANCE MAPPING

The TIC-based self-correcting partial covariance between each pair of fragment signals X and Y is given by Eq. (3):

$$\begin{aligned} \text{pCov}(X, Y, \text{TIC}) \\ = \text{Cov}(X, Y) - \text{Cov}(X, \text{TIC})\text{Cov}(Y, \text{TIC})/\text{Cov}(\text{TIC}, \text{TIC}). \end{aligned} \quad (3)$$

In order to provide a theoretical basis for the use of TIC as the single parameter of self-correcting partial-covariance mapping, we extend the statistical theory of covariance for noisy Poissonian processes [29]. The theoretical framework of Ref. [29] is valid when the number of ions trapped in each scan follows a noise-augmented Poisson distribution, which we verify for our ESI experiments; see Fig. S1 in the Supplemental Material [30]. We find theoretically (see Appendix A for details) that the value of the TIC-based self-correcting partial covariance between two fragment ions X and Y is

$$\text{pCov}(X, Y, \text{TIC}) = \beta^2 \nu_0 \alpha_1 - \omega \nu_0, \quad (4)$$

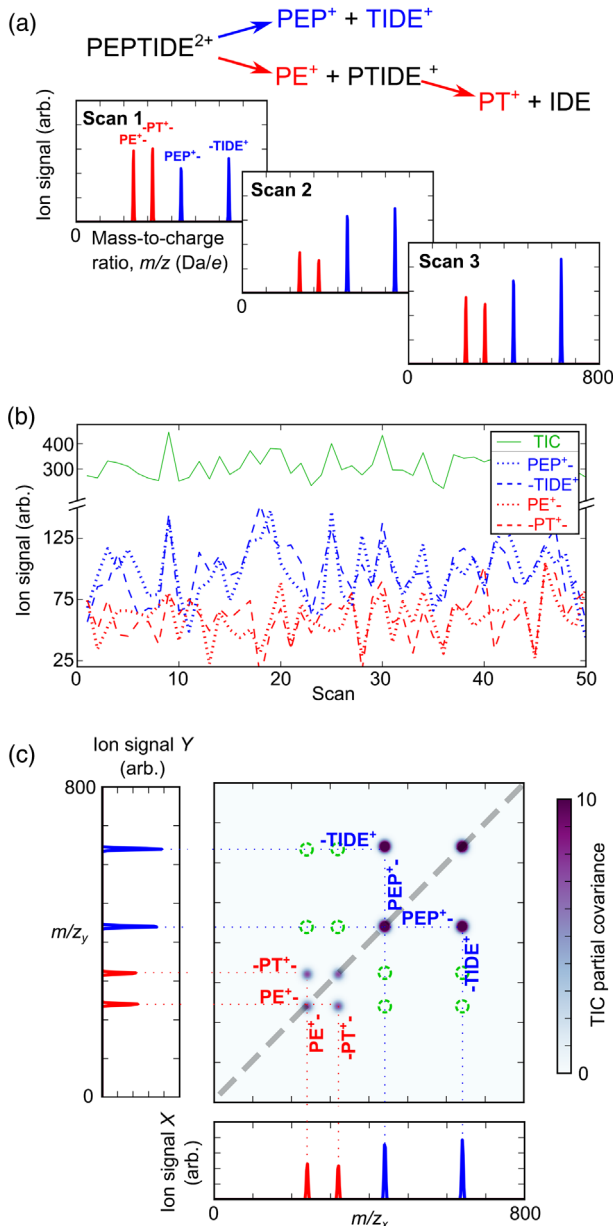


FIG. 1. The principle of 2D-PC-MS. (a) A computer records a series of fragment mass spectra (“scans”) of a biomolecular ion “PEPTIDE²⁺,” which dissociates along two pathways: blue and red. (b) Because of the statistical nature of the dissociation processes, the signals of each kind of fragment (red and blue) fluctuate randomly from scan to scan. These “intrinsic” fluctuations are superimposed with “extrinsic” ones caused by fluctuations in experimental parameters, which affect all the fragments simultaneously and are reflected in the TIC fluctuations (green). Signals formed in the same (blue) or consecutive (red) dissociation processes follow each other and are positively correlated. However, blue fragments are also correlated with red fragments via the TIC fluctuations. (c) Calculating the TIC-based self-correcting partial covariance [Eq. (3)] between each pair of fragment signals produces a map of physical fragment correlations: intrinsic correlations between pairs of blue or red fragments are revealed (black islands), while extrinsic correlations are suppressed (green circles).

where ν_0 is the average number of parent ions fragmented at each scan, α_1 is the probability that a single parent ion will dissociate to produce fragments X and Y, and β is the instrumental fragmentation-to-detection efficiency (e.g., in a mass spectrometer, the product of fragment ion-transfer and -detection efficiencies) typically required to be approximately 30% or higher for a successful covariance analysis [34]. ω is a small positive constant dependent on the instrumental fragmentation-to-detection efficiency and branching ratios of all fragmentation pathways producing two products (see Appendix A). Therefore, the small additional term $-\omega\nu_0$, which arises due to the imperfection of the TIC-based partial-covariance correction, is guaranteed to be negative and cannot result in positive covariance between two unrelated signals. The important consequence of this is that TIC-based self-correcting partial covariance is false-positive-free: It may artificially suppress a very weak true correlation, but it can never result in a positive correlation between two fragments which did not originate from the same or consecutive decompositions.

In Fig. 2, we demonstrate experimentally the suppression of extrinsic correlations by TIC-based self-correcting partial covariance. The figure displays a representative region of the simple covariance [panel (a), Eq. (1)] and the self-correcting partial-covariance [panel (b), Eq. (3)] maps of CID fragmentation of the triply protonated peptide **P1** (VTIMPK_{Ac}DIQLAR³⁺; each letter designates one amino acid [35], K_{Ac} = acetyllysine). In the simple covariance map obtained using Eq. (1), every fragment of **P1** is correlated with every other one [see red islands in Fig. 2(a)]. As we discuss in Sec. IA, this is due to the extrinsic correlations arising from the simultaneous fluctuations of all the fragment intensities caused by multiple fluctuating experimental parameters. On the other hand, the self-correcting partial covariance [Fig. 2(b)] obtained using Eq. (3), suppresses these extrinsic correlations [see blue islands in Fig. 2(b)], exposing the physical correlations between ions born in the same or consecutive dissociation processes. For example, the fragments -PK_{Ac}DIQLAR²⁺ and -QLAR⁺ correlated at the top of Figs. 2(a) and 2(b) contain the same part of the original peptide sequence (QLAR) and therefore cannot be intrinsically correlated products in a decomposition of **P1**; see Eq. (5a). The suppression of extrinsic correlations by TIC partial covariance reveals true physical (intrinsic) correlations between fragments that are indeed formed in the same (VTIM⁺ and -PK_{Ac}DIQLAR²⁺) or in consecutive (e.g., [VTIM- - CO]⁺ and -PK_{Ac}DIQLAR²⁺) decompositions [see Eqs. (5b) and (5c)]



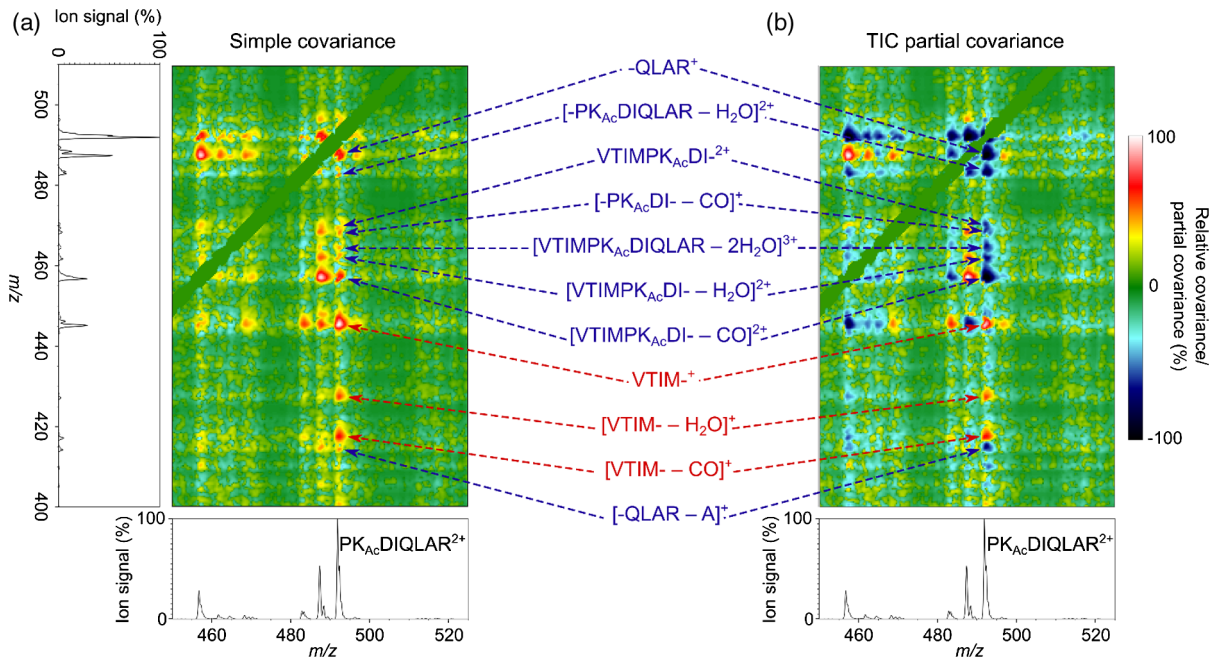
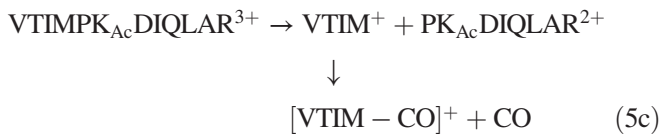


FIG. 2. Suppression of extrinsic correlations in TIC-based self-correcting partial-covariance mapping. In the CID fragmentation of triply protonated peptide **P1** (VTIMPK_{Ac}DIQLAR³⁺), a particular fragment labeled as PK_{Ac}DIQLAR²⁺ (shown on 1D spectrum in the bottom) correlates with several other fragments (shown by blue and red dashed arrows). In simple covariance mapping [panel (a)], fluctuations in experimental conditions induce extrinsic correlations between fragments originating from unrelated decomposition processes of the ionized peptide (the fragment correlates with every other one). Subtracting from the simple map a correction that involves covariance of the ion spectra with the TIC [see Eq. (3)] suppresses these extrinsic correlations (blue arrows). The TIC self-correcting partial-covariance map [panel (b)] correctly identifies fragments that are born in the same or consecutive decomposition processes (red arrows) of the given peptide ion.



IV. 2D-PC-MS: DEMONSTRATION OF PRINCIPLE

We now illustrate a full 2D-PC-MS measurement using the triply protonated peptide **P1**. Fragmentation of the amino acid sequence of **P1** under collision-induced dissociation employed in this work, generates a wide range of fragments of diverse and well-understood types (“terminal,” containing one of the two ends of the peptide sequence; “internal,” not containing either of the two ends of the peptide sequence; “neutral loss,” resulting from elimination of small neutral molecules, e.g., H₂O or CO from a fragment, etc.) providing ample opportunity to examine fragment-fragment correlations of the triply charged peptide. Figure 3 presents the 2D-PC-MS map of **P1** plotted as a 3D view of the relative TIC partial-covariance between every pair (X, Y) of the peptide fragment signals. The signals of related fragments generated in the same or consecutive decompositions of the

peptide appear as off-diagonal peaks. Such are, for example, correlations between fragments VTIMPK_{Ac}DIQ²⁺ (b_9^{2+}) and -LAR⁺ (y_3^+) produced by a cleavage of a single peptide bond of **P1** or those involving product(s) of a loss of a small neutral molecule, such as between [VTI- - H₂O]⁺ ($[b_3^+ - \text{H}_2\text{O}]^+$) and -MPK_{Ac}DIQLAR²⁺ (y_9^{2+}), as well as correlations involving internal fragments—products of consecutive breaking of two peptide bonds [36], e.g., -PK_{Ac}DIQ⁺ ($b_{i(5-9)}^+$) and -LAR⁺ (y_3^+). The standard nomenclature [37] used for peptide fragments is given in the inset of Fig. 3. See the classification of 2D-PC-MS correlation patterns in Appendix B.

In order to rank the strength of the observed correlations between the pairs of fragments, as well as distinguish low-intensity intrinsic correlations from statistical noise due to finite summations in Eq. (3), we introduce a correlation score based on correlation stability toward resampling for each feature on a 2D-PC-MS map; see Appendix C. We find that approximately 10³ individual fragment spectra are required to identify the intrinsic correlations by their correlation score; see Fig. S3 in the Supplemental Material [30]. We verify that 2D-PC-MS is capable of correlating fragments spanning more than 3 orders of magnitude of intensity, including signals of such low relative abundances as 0.1% and below (see

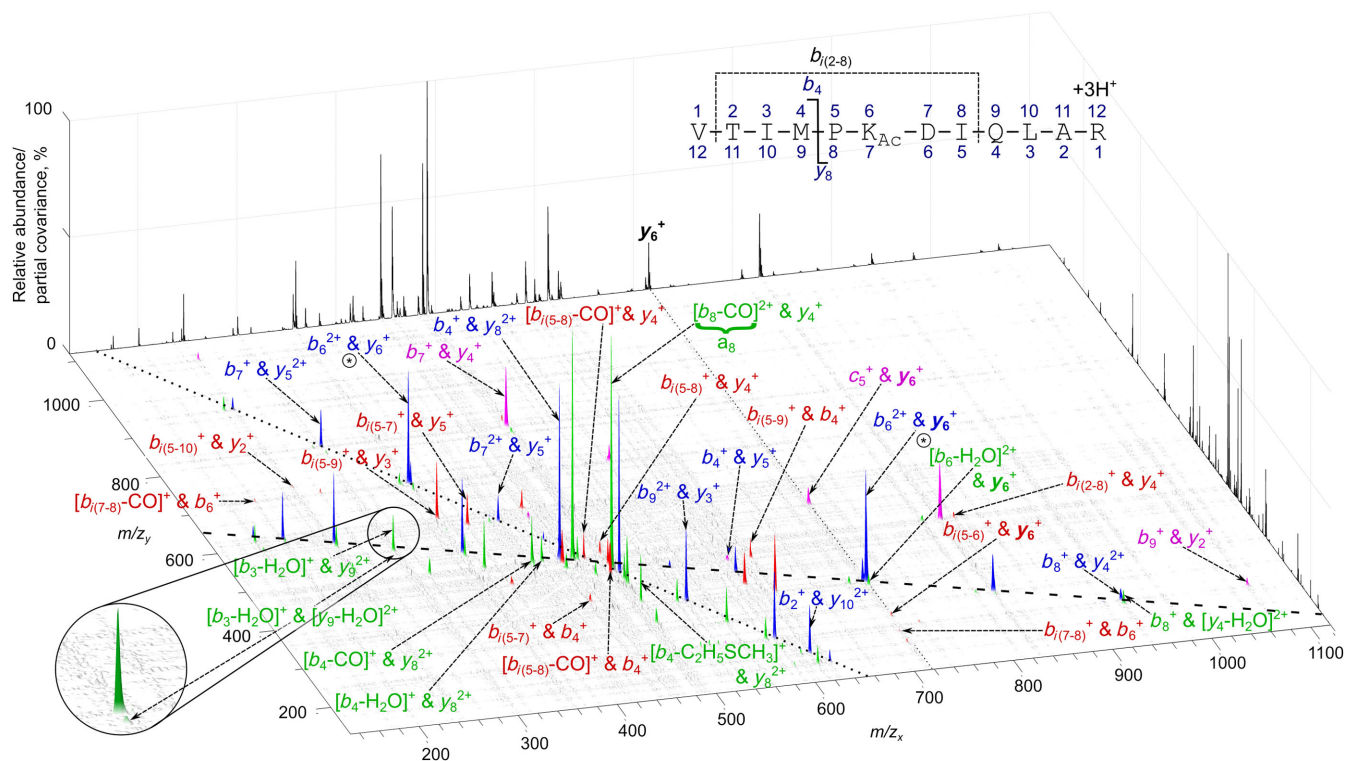


FIG. 3. Fragment-fragment correlations of the peptide ion **P1** (VTIMPK_{Ac}DIQLAR³⁺) revealed on the TIC partial-covariance map. The correlation signals are labeled as (X and Y) pairs of fragments corresponding to the 1D spectra plotted along the back edges of the map. The standard peptide fragment nomenclature used [37] is given in the inset. A single peak on the 1D spectrum belonging to the y_6^+ (-DIQLAR⁺) fragment is resolved into four fragment-fragment correlation peaks along the black dotted line on the map, showing four different fragmentation processes by which it is produced. Apart from the y_6^+ and b_6^{2+} signal (marked by ⊗), the 2D-PC-MS peaks are annotated only on one side of the autocorrelation diagonal for clarity. The blue dotted and dashed mass conservation lines follow correlations (blue peaks) where a single bond breaks in the chain of 12 amino acids of the parent ion. Green correlations mark a loss of small neutral molecules (e.g., water) in the fragmentation process. The red and magenta correlations reveal processes where more than one peptide bond is broken. The map provides much more information on the fragmentation mechanisms and for identifying the parent peptide than the 1D spectrum. Selected correlation signal annotations using single-letter amino acid code [35] are shown in Fig. S2 in the Supplemental Material [30].

Fig. S6 in the Supplemental Material [30]). The ranked 2D-PC-MS data of **P1** reveal that all 50 top scoring correlations correspond to physical decomposition channels of the peptide, i.e., the pathways producing fragments generated in the same or consecutive standard CID decompositions (30 of the 50 top scoring correlations are labeled in Fig. 3). The presence of the predicted intrinsic correlations according to the known amino acid sequence and the peptide fragmentation mechanisms, alongside the observed suppression of spurious signals corresponding to unphysical fragmentation pathways, experimentally validate the TIC-based self-correcting partial covariance. We identify the extrinsic correlations with low correlation scores as violating mass conservation, charge conservation, or covariance-mapping correlation patterns (see Appendix B).

The 2D-PC-MS map of **P1** (Fig. 3) enables one to detect the origin of the correlation signals simply on the basis of their geometric position. There are three regions in the map, each containing peaks of different origin.

First are the so-called mass conservation lines (plotted as dashed and dotted lines in Fig. 3) containing correlation peaks of pairs of complementary terminal fragments generated by fission of a single peptide bond and making up the whole sequence once combined (see signals shown in blue). Second are the horizontal and vertical peak series relative to the complementary fragment correlations containing the correlation signals of fragments originating from consecutive decompositions through loss of neutral molecules (e.g., NH₃, H₂O, CO) from the complementary ions; see green signals in Fig. 3. Third is the manifold of scattered signals revealing correlations involving fragments originating from breakage of more than one bond between the peptide amino acids; these signals are marked in red in Fig. 3. Breaking two or more bonds between amino acids can also lead to noncomplementary terminal fragment correlations excluding the internal parts of the sequence, such as VTIMPK_{Ac}D⁺ (b_7^+) and QLAR⁺ (y_4^+); these signals are marked in magenta in Fig. 3.

TABLE I. List of modified and unmodified peptides used for validation of 2D-PC-MS.

Sequence ^a	Molecular mass (Da)	Charge state(s)
TWR(Me ₂)GGEEK	989.5	2+, 3+
MLGIIAGKNSG	1059.6	2+, 3+
GSNKGAIIGLM	1059.6	2+, 3+
SAPtPEALAFVR	1226.6	2+
RDpYTGW – Nle – DF–NH ₂	1250.5	2+, 2–
TWpTLCGTVEY	1251.5	2+
GGNFSGR(Me)GGFGGSR	1325.6	2+, 3+
GK(Ac)GGKGLGK(Ac)GGAKR	1353.8	2+, 3+
GKGGK(Ac)GLGKGGAK(Ac)R	1353.8	2+, 3+
GK(Ac)GGKGLGKGGAK(Ac)R	1353.8	2+, 3+
GKGGK(Ac)GLGK(Ac)GGAKR	1353.8	2+, 3+
VTIMPKDIQLAR	1383.8	2+, 3+
VTIMPK(Me ₃)DIQLAR	1425.8	2+, 3+
VTIMPK(Ac)DIQLAR	1425.8	2+, 3+
GDFEIIPEEYLQ	1467.6	2–
EQFDDpYGHMRF–NH ₂	1522.6	2+, 3+
LGEY(nitro)GFQNAILVR	1523.8	2+, 3+
GWGR(Me ₂)EENLFSWK	1535.8	2+, 3+, 2–
GDFEIIPEEpYLQ	1547.6	2–, 3–
YGGFLRRIRPKLK	1603.0	2+, 3+

^aEach letter designates one unmodified amino acid [35]. Modified amino acids: R(Me₂) = dimethylarginine; pT = phosphothreonine; pY = phosphotyrosine; R(Me) = methylarginine; K(Ac) = acetyllysine; K(Me₃) = trimethyllysine; Y(nitro) = nitrotyrosine.

We further test 2D-PC-MS on a series of protonated and deprotonated peptide ions with unmodified and PTM-containing sequences of molecular masses ranging from 989.5 to 1603.0 Da (38 in total) listed in Table I. For each examined peptide ion, the 2D-PC-MS map is generated (see examples in Fig. 4), the scored list of correlating fragments is produced, and the highest-scoring correlations (typically numbering at least 40) are successfully confirmed to be free of any spurious signals. Moreover, we test 2D-PC-MS on ribonucleic acid and DNA oligonucleotides as well as on small proteins in the 10⁴-Da mass range (see Figs. S4 and S5 in the Supplemental Material [30]). These data further validate the TIC-based self-correcting partial covariance and demonstrate that it allows one to apply covariance-mapping spectroscopy to molecules of mass greater than 10⁴ Da, without the need for continuous monitoring of the multiple experimental parameters in a complex multistage ESI-CID measurement. The revealed fragment ion correlations provide direct experimental evidence for decomposition pathways of the multiply charged biomolecular ions.

V. 2D-PC-MS: STRUCTURAL SPECIFICITY

2D-PC-MS provides a new kind of structural information which is missing in standard 1D mass spectrometry for

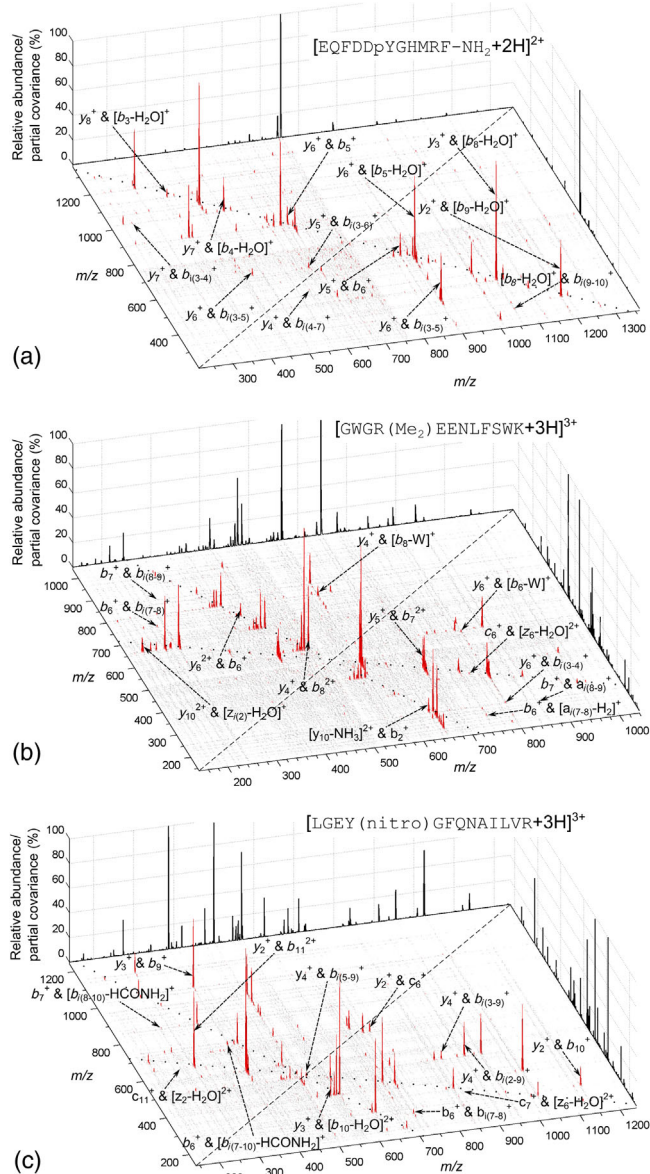


FIG. 4. Examples of 2D-PC-MS maps of protonated peptides. A nonexhaustive selection of top scoring 2D-PC-MS fragment-fragment correlation signals are labeled on each map, including very weak signals corresponding to the true (intrinsic) fragment correlations revealed by the 2D-PC-MS scoring; see Eq. (C1). Signal annotation is as in Fig. 3. It is a general property of 2D-PC-MS that it can reveal structurally informative signals even when the corresponding 1D spectrum is dominated by parent ion neutral loss [see panel (a) and also Fig. S7 in the Supplemental Material [30]]. Mass conservation lines are shown by dotted lines. Autocorrelation signals are removed and the autocorrelation diagonals are shown by dashed lines.

sequence analysis: the fragment-fragment connectivity. This connectivity can be used to solve the peptide sequence puzzle by piecing the fragments together, as is the primary aim of analytical mass spectrometry. Here we investigate and quantify the advantage presented by fragment connectivity for deducing a peptide sequence by numerical simulation of

the number of possibilities to assign fragment-fragment correlations of a given peptide to amino acid sequence of a *different* (“erroneous”) peptide of the same or nearly the same m/z . This number expressed as a fragment ambiguity rate (FAR) (see Appendix D) is compared to the number of possible false-positive assignments based on individual fragment ions measured by standard 1D MS. The lower the FAR of a spectral signal (fragment-fragment correlation for 2D-PC-MS or individual fragment for 1D MS), the lower the number of different erroneous peptides that can produce this signal, and thus, the more reliable the signal is for the purpose of identifying the correct peptide sequence. Reducing this FAR is the fundamental driver for the development of higher-mass resolution measurement instrumentation [25,26].

We calculate the relative FARs of fragment-fragment correlations for a sample of 50 000 peptide sequences from protein database UniProt/Swiss-Prot [38] based on typical fragmentation of protonated peptides under CID [39]. These fragment ambiguity rates are calculated by matching the theoretical correlations of each peptide from the randomly selected sample pool against the theoretical correlations of the full set of isobaric (within 7-ppm mass tolerance) peptides of the database. The resulting relative FAR values are compared to the relative FARs calculated for matching individual (uncorrelated) fragments produced by the same sample of the 50 000 peptides to simulate 1D MS analysis. Further details of the calculations are given in Appendix D.

The results of the FAR comparison between 2D-PC-MS and 1D MS are presented in Fig. 5. The FAR data reveal that even at the relatively large 0.8-Da mass tolerance for fragment ion matching, the 2D fragment-fragment correlations involving internal fragments are significantly more sequence specific than the absolute lower limit of what 1D MS can achieve in principle at the theoretical *infinite* mass accuracy and resolution, for any fragment ion type. This superior sequence specificity is shown by the red bar at mass tolerance 0.8 Da in Fig. 5(b) vs the black bar at zero mass tolerance in Fig. 5(a), whose height is marked by dashed line to guide the eye. The abundance of internal fragments in peptide mass spectra relative to terminal ones is determined by two counteracting trends: (i) the number of theoretically possible internal fragments scales as the square of the sequence length (i.e., much stronger than the number of complementary terminal fragments which scales linearly), and (ii) their production requires more bond cleavages, i.e., is less probable at lower activation energies. As a result, the actual fraction of the internal fragments in a spectrum weakly depends on the peptide sequence length and can reach about 50% of TIC depending on the activation mode [40]. Thus, our simulations of fragment ambiguity rate indicate that for up to half of the fragments, the fragment-fragment correlations provide sequence specificity that cannot be matched by 1D MS

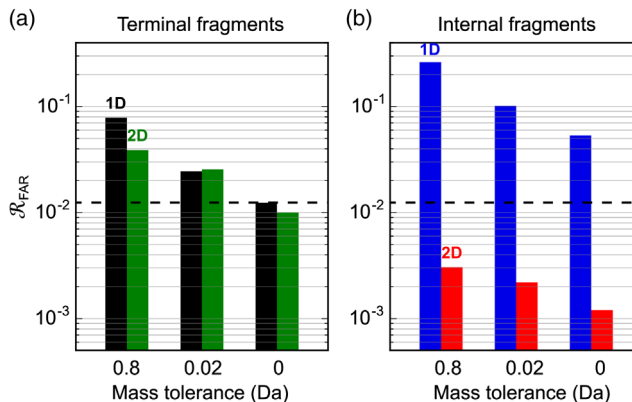


FIG. 5. Ambiguity of peptide fragment identification in the 1D and 2D techniques compared by a numerical simulation. Peptides of a given mass are selected from a protein database within the 7-ppm mass tolerance and fragmented according to known fragmentation rules. The resulting individual fragment ion m/z (in 1D) and correlated pairs of m/z (in 2D-PC-MS) are tested for uniqueness among the database peptides within the given fragment mass-matching tolerance, and any ambiguity is counted toward the average relative FAR of the given type of fragments (terminal or internal). While the 2D and 1D MS perform comparably for the terminal fragments [panel (a)], there is a dramatic reduction in identification ambiguity (note the logarithmic scale) when looking at the internal fragments [panel (b)]. In fact, the ambiguity rate for 2D signals involving internal fragments, even at a modest 0.8-Da tolerance, drops well below the theoretical lowest limit (0-Da tolerance) of the 1D technique for any fragment type (horizontal dashed lines).

even theoretically in an “ideal” (infinite mass-resolution and accuracy) measurement. Noteworthy, the internal-ion correlations can sometimes be the only available key to discriminating between similar isomeric structures [41].

The data presented in Fig. 5 show a large difference between the FARs for correlations between complementary terminal fragments vs those involving internal ions. While the complementary terminal-ion relative FARs in 2D-PC-MS [green bars in Fig. 5(a)] are found to be similar in magnitude to those of the individual terminal ions in 1D MS [black bars in Fig. 5(a)], at all mass tolerances, the relative FARs of 2D-PC-MS correlations involving an internal ion [red bars in Fig. 5(b)] are more than an order of magnitude below the 1D MS values [blue bars in Fig. 5(b)]. This profound difference is explained by the following mass conservation argument. In the case of complementary fragments, the masses of which add up to the total peptide mass, a chance matching of one of the fragments ensures the mass matching of its complementary counterpart and therefore also the random matching of the full complementary ion pair. On the contrary, for correlations involving noncomplementary fragment ions, e.g., an internal fragment with a terminal one, the above mass conservation argument does not apply, because a part of the peptide sequence is missing, and matching a random fragment at m/z close (within a given tolerance) to m/z

of one of the true correlating ions does not guarantee matching another fragment at m/z similar to its correlating 2D-PC-MS partner. In other words, the random matching of a terminal-ion–internal-ion correlation requires the simultaneous occurrence of two approximately independent random events, making the random match of a correlation much less probable than a match of a single ion. This reduced probability is reflected in the low 2D-PC-MS FARs for internal-ion correlations [see Fig. 5(b)].

VI. CONCLUSIONS AND OUTLOOK

In conclusion, we introduce a new method for the analysis of decompositions of large ionized molecules: two-dimensional partial-covariance mass spectrometry. 2D-PC-MS is based on a conceptually new idea of removing the nonphysical extrinsic correlations using the information contained in the acquired spectrum itself. We experimentally and theoretically validate the use of the total ion count as a single spectrum-derived parameter for self-correcting partial-covariance mapping. The TIC parameter derived from the measured spectrum itself embodies the combined effect of all the known and unknown experimental parameters on each spectral signal simultaneously and serves as a key to unlocking the physical correlations between the biomolecular fragments within the self-correcting partial-covariance paradigm. Our method not only enables but in fact makes it straightforward to obtain 2D maps of fragment-fragment correlations of species 2 orders of magnitude larger than has been possible so far: peptides, oligonucleotides, and proteins in the mass region of 10^3 – 10^4 Da.

The question as to the dynamical mechanisms of radiation damage to biomolecules is only now becoming possible to answer with high-brightness ultrafast x-ray sources. This question is important to our understanding of life on Earth as well as on other planets, medical physics, and potential new routes to structure determination. Large molecule fragmentation experiments have already been conducted [42] and planned at x-ray free-electron lasers, with the objective of furthering the time-resolved understanding of x-ray fragmentation pathways that may be important in radiation damage mechanisms in biology and the prospects for single-molecule imaging via the “diffract-before-destroy” concept. While coupling of the linear ion-trap mass spectrometer to an x-ray beam line has recently been achieved [43], the present development paves the way for detailed fragmentation studies within an x-ray facility setting by demonstrating the full compatibility of the self-correcting partial-covariance mapping with electrospray ionization [11] and ion-trapping [19] technologies that are core to biomolecular fragmentation research. Therefore, we expect TIC-based self-correcting partial-covariance mapping to provide a direct route to unraveling the complex biomolecular XFEL-induced fragmentation pathways.

In addition to their mechanistic significance, the fragment-fragment correlations revealed by the self-correcting

partial-covariance mapping open the way for applying covariance-mapping spectroscopy to solving the inverse problem, namely, the reconstruction of molecular structure from the spectrum. This inverse problem is central to analytical mass spectrometry. Our numerical simulations show that fragment-fragment correlations involving internal fragments of the biomolecular sequence provide a structural fingerprint of unparalleled specificity. The proposed method is developed and demonstrated using a commercial benchtop mass spectrometer that requires no hardware modifications, which we believe will facilitate the wide utilization of 2D-PC-MS as a practical tool. In general, the applicability of 2D-PC-MS depends on sufficiently high and uniform fragmentation-to-detection efficiency. Once this is ensured, 2D-PC-MS can be applied to any chemical species, that is, not only to peptide, proteins, and oligonucleotides considered in this work but also to oligosaccharides, lipids, and metabolites. For the same reason, 2D-PC-MS can be applied in combination with any ionization and activation method, including ultraviolet photodissociation, infrared multiphoton dissociation, femtosecond-laser ionization and dissociation, as well as ETD and ECD.

ACKNOWLEDGMENTS

M. E. A. acknowledges support of the Wellcome Trust through the research fellowship Grant No. WT100093MA. Support is provided by EPSRC program Grants No. EP/I032517/1 and EPSRC/DSTL MURI No. EP/N018680/1, Pathways to Impact Grant No. EP/K503733/1, and ERC ASTEX Project No. 290467. Fruitful discussions with Professor Charles Coombes and Professor Simak Ali are gratefully acknowledged.

M. E. A. conceived, designed, and guided the research. T. D. performed the research. R. A. contributed to the oligonucleotide part of the research. B. C. and S. P. constructed the theoretical model of TIC-based self-correcting partial covariance. V. A. and L. J. F. guided T. D. in the numerical calculations. L. J. F. guided T. D. in the development of the peak scoring procedure. R. P. carried out the peptide synthesis. All the authors contributed to the research development, data analysis, and manuscript preparation.

APPENDIX A: FORMAL THEORY OF TIC-BASED SELF-CORRECTING PARTIAL-COVARIANCE MAPPING FOR NOISY POISSON PROCESSES

Mikosch and Patchkovskii [29] developed a formal theory of covariance for noisy Poisson processes. For such processes, the event rate follows a so-called “augmented Poisson distribution,” which is a Poisson distribution for which the mean event rate $\nu = \gamma\nu_0$ is itself sampled from a normal distribution, with constant ν_0 the central event rate and γ normally distributed and centered on $\gamma = 1$. The

statistics are analyzed in terms of the ‘‘compound outcomes’’ of the decomposition of a molecular ensemble [29]. This framework can be used to analyze a covariance measurement on a system with known outcome probabilities for the mechanism being probed and known finite fragmentation-to-detection efficiency. By systematically considering the respective probabilities of each relevant spectral measurement, an analytical expression for the simple covariance between any number of spectral channels is derived.

Here we extend the theoretical approach of Ref. [29] to derive an expression for the partial covariance between two spectral signals where the total number of measured ions (TIC) is used as a single partial-covariance parameter [Eq. (3)]. We are interested in the partial covariance between two fragment ions X and Y for the case of the induced fragmentation of a biomolecular ion. The total number of parent ions subjected to analysis at each scan is assumed to follow the noise-augmented Poisson distribution, which we experimentally confirm in Fig. S1 in the Supplemental Material [30] for our measurements. We consider the following probabilities for the decomposition of a single parent ion:

- (1) X and Y are created in the parent ion decomposition, with probability $P = \alpha_1$.
- (2) X and another molecular fragment Z that is not Y are created in the parent ion decomposition (e.g., Z is the product of a neutral loss from Y), $P = \alpha_2$.
- (3) Y and another molecular fragment Z that is not X are created in the parent ion decomposition (e.g., Z is the product of a neutral loss from X), $P = \alpha_3$.
- (4) Any two detectable molecular fragments, neither of which are X or Y , are created in the parent ion decomposition (e.g., parent ion dissociates to two other complementary fragment ions $A + B$, or

parent ion dissociates to terminal fragment ion A , internal fragment ion B , and neutral fragment ion C), $P = \alpha_4$

- (5) The parent ion does not fragment, $P = \alpha_5$.

We then account for the finite fragmentation-to-detection efficiency β by considering possible measurement outcomes following an elementary fragmentation event. The following elementary measurement events are relevant for the TIC-based self-correcting partial covariance:

- (1) X and Y are both measured at the detector, $P_1 = \alpha_1\beta^2$.
- (2) Only X is detected (i.e., Y or Z is also produced but are not observed due to the finite fragmentation-to-detection efficiency), $P_2 = (\alpha_1 + \alpha_2)\beta(1 - \beta)$.
- (3) Only Y is detected (i.e., X or Z is also produced but is not observed), $P_3 = (\alpha_1 + \alpha_3)\beta(1 - \beta)$.
- (4) X and another ion Z that is not Y (e.g., Z is the product of a neutral loss from Y) are both measured at the detector, $P_4 = \alpha_2\beta^2$.
- (5) Y and another ion Z that is not X (e.g., Z is the product of a neutral loss from X) are both measured at the detector, $P_5 = \alpha_3\beta^2$.
- (6) Only one ion that is not X or Y is detected (i.e., parent ion dissociates to two different fragment ions $A + B$ and only A or B is observed, a non-fragmented parent ion is detected, or fragmentation leads to $X/Y+Z$, but X/Y is not detected), $P_6 = \beta(\alpha_5 + (1 - \beta)(\alpha_2 + \alpha_3 + 2\alpha_4))$.
- (7) Any two ions, neither of which are X or Y , are measured at the detector (i.e., parent ion dissociates to two different fragment ions $A + B$, both of which are measured at the detector), $P_7 = \alpha_4\beta^2$.

Following the procedure outlined in Ref. [29], we obtain the final expression for the TIC-based self-correcting partial covariance between X and Y :

$$\begin{aligned} \text{pCov}(X, Y, \text{TIC}) = & \beta^2\nu_0\alpha_1 - \nu_0 \frac{\beta(1 + \beta)^2(\alpha_1 + \alpha_2)(\alpha_1 + \alpha_3)}{\alpha_5 + 2(1 + \beta)\alpha_s} \\ & + \nu_0^2\sigma^2\alpha_5^2 \frac{\beta^4(\alpha_1 + \alpha_2)(\alpha_1 + \alpha_3)}{(\alpha_5 + 2(1 + \beta)\alpha_s)(\alpha_5 + 2(1 + \beta)\alpha_s + \beta\nu_0\sigma^2(\alpha_5 + 2\alpha_s)^2)}, \end{aligned} \quad (\text{A1})$$

where $\alpha_s = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4$. We observe complete parent ion fragmentation in the experiments performed in this work, meaning $\alpha_5 \approx 0$ and the expression simplifies to TIC-based self-correcting partial covariance between two fragment ions X and Y as given by

$$\begin{aligned} \text{pCov}(X, Y, \text{TIC}) = & \beta^2\nu_0\alpha_1 - \omega\nu_0, \\ \omega = & \frac{\beta(1 + \beta)(\alpha_1 + \alpha_2)(\alpha_1 + \alpha_3)}{2\alpha_s}, \end{aligned} \quad (\text{A2})$$

where ν_0 is the average number of trapped ions, and the fragmentation-to-detection efficiency β is typically

required to be approximately 30% or higher for successful covariance analysis [34].

Equation (A2) shows that the TIC-based self-correcting partial covariance between X and Y has a positive contribution that is directly proportional to the probability of a parent ion producing the fragments X and Y , and a negative contribution stemming from competing processes.

Given high enough fragmentation-to-detection efficiency and provided the number of dissociation pathways available to the fragmenting molecule is large enough such that $\alpha_4 \gg \alpha_{1,2,3}$ (practically, this is the case for all fragmenting biological polymers), the first term dominates, and

the TIC partial covariance is linearly dependent on both parent ion abundance and the branching ratio of the process that produces X and Y together, meaning 2D-PC-MS is suitable for quantitative analysis. The formal analysis of the TIC-based self-correcting partial covariance also shows that it is false-positive-free: If the TIC partial covariance between a fragment pair is positive, to a desired degree of statistical significance, this pair is known to have been produced in the same fragmentation process, to at least the same degree of statistical significance.

APPENDIX B: 2D-PC-MS FRAGMENT CORRELATION PATTERNS

The fragment-fragment correlation pairs (\mathbf{X} , \mathbf{Y}) revealed by 2D-PC-MS can be classified as *siblings* ($Z \rightarrow \mathbf{X} + \mathbf{Y}$, correlations between products of the same decomposition reaction), *niblings* ($Z \rightarrow \mathbf{X} + \mathbf{Y}_1$, $\mathbf{Y}_1 \rightarrow \mathbf{X}_1 + \mathbf{Y}$, correlations involving a product of further decomposition), *cousins* ($Z \rightarrow \mathbf{X}_1 + \mathbf{Y}_1$, $\mathbf{X}_1 \rightarrow \mathbf{X} + \mathbf{X}_2$, $\mathbf{Y}_1 \rightarrow \mathbf{Y} + \mathbf{Y}_2$, correlations between products of further decomposition), etc. In the peptide fragmentations considered in this work, siblings are the complementary terminal ions formed by breaking of a single peptide bond, such as VTIM^- and $-\text{PK}_{\text{Ac}}\text{DIQLAR}^{2+}$ in Fig. 3; niblings include products of a neutral loss of a small molecule (e.g., H_2O , NH_3) from one of the fragment ions, such as $[\text{VTI}^- - \text{H}_2\text{O}]^+$ and $-\text{MPK}_{\text{Ac}}\text{DIQLAR}^{2+}$ in Fig. 3 and pairs of correlated fragments including an internal ion, such as $-\text{PK}_{\text{Ac}}\text{DIQ}^+$ and $-\text{LAR}^+$ in Fig. 3; cousins are, for example, products of neutral losses from both of the fragment ions formed by peptide bond dissociation, such as $[\text{VTI}^- - \text{H}_2\text{O}]^+$ and $[-\text{MPK}_{\text{Ac}}\text{DIQLAR} - \text{H}_2\text{O}]^+$ in Fig. 3.

APPENDIX C: CORRELATION SCORE

It is useful to distinguish low-intensity intrinsic correlations from the statistical noise present because of the finite summations in Eq. (3). We achieve this discrimination by introducing a correlation score $S(X, Y)$ calculated by normalizing the volume of 2D-PC-MS signals (X , Y) to the standard deviation $\sigma(V)$ of their volume upon jackknife resampling [44]:

$$S(X, Y) = \frac{V[\text{pCov}(X, Y; I)]}{\sigma(V)}. \quad (\text{C1})$$

We rank all the 2D-PC-MS signals according to their correlation scores taken as a percentage of the highest $S(X, Y)$ off-diagonal peak on the map. The $S(X, Y)$ score reflects the stability of the 2D-PC-MS correlation signals and is superior to 2D-PC-MS peak height or volume as a measure of true structural signals; see Fig. S7 in the Supplemental Material [30]. The correlation score can be

used as a single universal parameter for 2D-PC-MS peak selection across the full m/z range.

APPENDIX D: DETAILS OF THE FRAGMENT AMBIGUITY RATE CALCULATION

For every theoretical fragment ion (1D) or fragment ion correlation (2D) x , the relative FAR value is calculated according to the following expression:

$$\mathcal{R}_{\text{FAR}}(x) = \frac{\sum_P N_m(P, \mathcal{T}, x)}{N_P}, \quad (\text{D1})$$

where the summation is over all the N_P database peptide ions P being within 7 ppm of the m/z of the parent ion of the fragment ion (or fragment-fragment correlation) x . $N_m(P, \mathcal{T}, x)$ is the number of m/z matches, within the fragment m/z tolerance, \mathcal{T} , between x and all the possible fragment ions (or fragment ion correlations) of the fragmenting peptide P . The resulting relative FARs are averaged over all x of chain lengths between two and 15 residues within a given category, e.g., individual y -type terminal ions or b -ion or internal-ion correlations. The comparison is performed at \mathcal{T} values of 0.8, 0.02, and 0 Da (i.e., matching isomeric fragments only, “ideal” m/z measurement) for both 1D fragment ion and 2D fragment-fragment correlation matching.

To perform the FAR analysis, we subject the protein sequences of the UniProt/Swiss-Prot database [38] to *in silico* tryptic digestion (using the most popular trypsin protease, no missed cleavages are considered). We ignore the 0.47% of UniProt/Swiss-Prot protein sequences with ambiguous amino acid residue coding and limit the digestion products to a minimum length of five residues and a maximum molecular weight of 2000 Da. From the resulting tryptic digest, we select five test sets of 10 000 peptides at random. For each such set, we fragment *in silico* the 3+ parent ion of each peptide within the set to produce every possible b^+ , internal b^+ and y^+ ion triplet from the parent ion. First, we extract all ion-trap measurable b^+ ion and internal b^+ ion and internal b^+ ion and y^+ ion correlations from these triplets (ion-trap measurable meaning that the m/z value of each of the correlated fragments is greater than or equal to 1/3 of the parent ion m/z [45]). Each of these 2D-PC-MS fragment correlations is next compared (m/z pair to m/z pair at a series of fragment m/z tolerances) to every b^+ ion and internal b^+ ion and internal b^+ ion and y^+ ion correlation from all triply charged peptide sequences of the full database tryptic digest within 7-ppm m/z tolerance of the parent peptide ion.

For the analogous 1D relative FARs, the m/z value of each ion-trap measurable terminal b^+ and y^+ fragment ions extracted from each triplet is compared individually to that of every possible terminal b^+ , b^{2+} , y^+ , and y^{2+} ion from all triply charged peptide sequences of the full tryptic digest

with parent peptide ion m/z within 7 ppm, to find all potential 1D fragment ion matches. For the 1D internal b^+ fragment ions, the m/z value of each ion-trap-measurable fragment ion is additionally compared to every possible internal b^+ and b^{2+} ion from the same set of peptide sequences from the tryptic digest, now modeling a 1D matching scheme that also considers internal ions. For comparison, we also calculate the relative FAR for matching each 2D correlation of every b^+ and y^+ terminal ion with its complementary terminal ion (y^{2+} and b^{2+} , respectively) against every b^+ ion and y^{2+} ion and b^{2+} ion and y^+ ion correlation, for all triply charged peptide sequences of the tryptic digest within 7-ppm m/z tolerance of the parent peptide ion.

-
- [1] L. J. Frasinski, K. Codling, and P. A. Hatherly, *Covariance Mapping: A Correlation Method Applied to Multiphoton Multiple Ionization*, *Science* **246**, 1029 (1989).
- [2] L. J. Frasinski *et al.*, *Dynamics of Hollow Atom Formation in Intense X-Ray Pulses Probed by Partial-Covariance Mapping*, *Phys. Rev. Lett.* **111**, 073002 (2013).
- [3] M. Mucke *et al.*, *Covariance Mapping of Two-Photon Double Core Hole States in C_2H_2 and C_2H_6 Produced by an X-Ray Free Electron Laser*, *New J. Phys.* **17**, 073002 (2015).
- [4] A. E. Boguslavskiy, J. Mikosch, A. Gijbsbertsen, M. Spanner, S. Patchkovskii, N. Gador, M. J. J. Vrakking, and A. Stolow, *The Multielectron Ionization Dynamics Underlying Attosecond Strong-Field Spectroscopies*, *Science* **335**, 1336 (2012).
- [5] M. Di Fraia *et al.*, *Impulsive Laser-Induced Alignment of OCS Molecules at FERMI*, *Phys. Chem. Chem. Phys.* **19**, 19733 (2017).
- [6] O. Kornilov *et al.*, *Coulomb Explosion of Diatomic Molecules in Intense XUV Fields Mapped by Partial-Covariance*, *J. Phys. B* **46**, 164028 (2013).
- [7] V. Zhaunerchyk, L. J. Frasinski, J. H. D. Eland, and R. Feifel, *Theory and Simulations of Covariance Mapping in Multiple Dimensions for Data Analysis in High-Event-Rate Experiments*, *Phys. Rev. A* **89**, 053418 (2014).
- [8] R. Aebersold and M. Mann, *Mass-Spectrometric Exploration of Proteome Structure and Function*, *Nature (London)* **537**, 347 (2016).
- [9] B. F. Cravatt, G. M. Simon, and J. R. Yates III, *The Biological Impact of Mass-Spectrometry-Based Proteomics*, *Nature (London)* **450**, 991 (2007).
- [10] M. Mann and O. N. Jensen, *Proteomic Analysis of Post-Translational Modifications*, *Nat. Biotechnol.* **21**, 255 (2003).
- [11] B. Fenn, M. Mann, C. K. Meng, S. F. Wong, and C. M. Whitehouse, *Electrospray Ionization for Mass Spectrometry of Large Biomolecules*, *Science* **246**, 64 (1989).
- [12] F. Hillenkamp, M. Karas, R. C. Beavis, and B. T. Chait, *Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry of Biopolymers*, *Anal. Chem.* **63**, 1193A (1991).
- [13] B. K. Shukla and J. H. Futrell, *Tandem Mass Spectrometry: Dissociation of Ions by Collisional Activation*, *J. Mass Spectrom.* **35**, 1069 (2000).
- [14] J. E. Syka, J. J. Coon, M. J. Schroeder, J. Shabanowitz, and D. F. Hunt, *Peptide and Protein Sequence Analysis by Electron Transfer Dissociation Mass Spectrometry*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 9528 (2004).
- [15] R. A. Zubarev, N. L. Kelleher, and F. W. McLafferty, *Electron Capture Dissociation of Multiply Charged Protein Cations*, *J. Am. Chem. Soc.* **120**, 3265 (1998).
- [16] J. S. Brodbelt, *Photodissociation Mass Spectrometry: New Tools for Characterization of Biological Molecules*, *Chem. Soc. Rev.* **43**, 2757 (2014).
- [17] C. L. Kalcic, T. C. Gunaratne, A. D. Jones, M. Dantus, and G. E. Reid, *Femtosecond Laser-Induced Ionization/Dissociation of Protonated Peptides*, *J. Am. Chem. Soc.* **131**, 940 (2009).
- [18] J. S. Brodbelt and J. J. Wilson, *Infrared Multiphoton Dissociation in Quadrupole Ion Traps*, *Mass Spectrom. Rev.* **28**, 390 (2009).
- [19] J. R. Yates, C. I. Ruse, and A. Nakorchevsky, *Proteomics by Mass Spectrometry: Approaches, Advances, and Applications*, *Annu. Rev. Biomed. Eng.* **11**, 49 (2009).
- [20] R. G. Sadygov, D. Cociorva, and J. R. Yates III, *Large-Scale Database Searching Using Tandem Mass Spectra: Looking Up the Answer in the Back of the Book*, *Nat. Methods* **1**, 195 (2004).
- [21] D. F. Hunt, J. R. Yates III, J. Shabanowitz, S. Winston, and C. R. Hauer, *Protein Sequencing by Tandem Mass Spectrometry*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 6233 (1986).
- [22] V. H. Wysocki, K. A. Resing, Q. Zhang, and G. Cheng, *Mass Spectrometry of Peptides, and Proteins*, *Methods* **35**, 211 (2005).
- [23] H. Lam, E. W. Deutsch, J. S. Eddes, J. K. Eng, N. King, S. E. Stein, and R. Aebersold, *Development and Validation of a Spectral Library Searching Method for Peptide Identification from MS/MS*, *Proteomics* **7**, 655 (2007).
- [24] K. F. Medzihradzky and R. J. Chalkley, *Lessons in De Novo Peptide Sequencing by Tandem Mass Spectrometry*, *Mass Spectrom. Rev.* **34**, 43 (2015).
- [25] A. Michalski, E. Damoc, J.-P. Hauschild, O. Lange, A. Wiegand, A. Makarov, N. Nagaraj, J. Cox, M. Mann, and S. Horning, *Mass Spectrometry-Based Proteomics Using Q Exactive, a High-Performance Benchtop Quadrupole Orbitrap Mass Spectrometer*, *Mol. Cell Proteom.* **10**, M111.011015 (2011).
- [26] C. R. Weisbrod, M. R. Hoopmann, M. W. Senko, and J. E. Bruce, *Performance Evaluation of a Dual Linear Ion Trap-Fourier Transform Ion Cyclotron Resonance Mass Spectrometer for Proteomics Research*, *J. Proteom.* **88**, 109 (2013).
- [27] P. Pfändler, G. Bodenhausen, J. Rapin, R. Houriet, and T. Gäumann, *Two-Dimensional Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*, *Chem. Phys. Lett.* **138**, 195 (1987).
- [28] M. A. van Agthoven, M.-A. Delsuc, G. Bodenhausen, and C. Rolando, *Towards Analytically Useful Two-Dimensional*

- Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*, *Anal. Bioanal. Chem.* **405**, 51 (2013).
- [29] J. Mikosch and S. Patchkovskii, *Coincidence and Covariance Data Acquisition in Photoelectron and Ion Spectroscopy. I. Formal Theory*, *J. Mod. Opt.* **60**, 1426 (2013).
- [30] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevX.10.041004> for materials and peptide synthesis; MS analysis and data processing; Figs. S1–S7, which includes Refs. [31–33].
- [31] R. B. Merrifield, *Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide*, *J. Am. Chem. Soc.* **85**, 2149 (1963).
- [32] T. Pekar Second, J. D. Blethrow, J. C. Schwartz, G. E. Merrihew, M. J. MacCoss, D. L. Swaney, J. D. Russell, J. J. Coon, and V. Zabrouskov, *Dual-Pressure Linear Ion Trap Mass Spectrometer. Improving the Analysis of Complex Protein Mixtures*, *Anal. Chem.* **81**, 7757 (2009).
- [33] S. A. McLuckey, G. J. Van Berkel, and G. L. Glish, *Tandem Mass Spectrometry of Small, Multiply Charged Oligonucleotides*, *J. Am. Soc. Mass Spectrom.* **3**, 60 (1992).
- [34] L. J. Frasinski, *Covariance Mapping Techniques*, *J. Phys. B* **49**, 152004 (2016).
- [35] IUPAC and IUB Joint Commission on Biochemical Nomenclature, *Pure & Appl. Chem.* **56**, 595 (1984), <http://publications.iupac.org/pac/1984/pdf/5605x0595.pdf>.
- [36] K. D. Ballard and S. J. Gaskell, *Sequential Mass Spectrometry Applied to the Study of the Formation of “Internal” Fragment Ions of Protonated Peptides*, *Int. J. Mass Spectrom. Ion Process.* **111**, 173 (1991).
- [37] P. Roepstorff and J. Fohlmann, *Proposal for a Common Nomenclature for Sequence Ions in Mass Spectra of Peptides*, *Biomedical Mass Spectrometry* **11**, 601 (1984).
- [38] <http://www.uniprot.org>.
- [39] K. F. Medzihradsky, *Peptide Sequence Analysis*, *Methods Enzymol.* **402**, 209 (2005).
- [40] Y. A. Lyon, D. Riggs, L. Fornelli, P. D. Compton, and R. R. Julian, *The Ups and Downs of Repeated Cleavage and Internal Fragment Production in Top-Down Proteomics*, *J. Am. Soc. Mass Spectrom.* **29**, 150 (2018).
- [41] T. Driver, R. Pipkorn, V. Averbukh, L. Frasinski, J. P. Marangos, and M. Edelson-Averbukh, *Breaking the Histone Code with Two-Dimensional Partial Covariance Mass Spectrometry*, *ChemRxiv*, <https://doi.org/10.26434/chemrxiv.12743669.v2>.
- [42] N. Berrah *et al.*, *Femtosecond Resolved Observation of Fragmentation of Buckminsterfullerene Following X-Ray Multiphoton Ionization*, *Nat. Phys.* **15**, 1279 (2019).
- [43] A. R. Miloslavljjevic, C. Nicolas, J.-F. Gil, F. Canon, M. Réfrégiers, L. Nahon, and A. Giuliani, *VUV Synchrotron Radiation: A New Activation Technique for Tandem Mass Spectrometry*, *J. Synchrotron Radiat.* **19**, 174 (2012).
- [44] B. Efron and C. Stein, *The Jackknife Estimate of Variance*, *Ann. Stat.* **9**, 586 (1981).
- [45] R. E. March, *An Introduction to Quadrupole Ion Trap Mass Spectrometry*, *J. Mass Spectrom.* **32**, 351 (1997).