PAPER • OPEN ACCESS

Photon energy-resolved velocity map imaging from spectral domain ghost imaging

View the article online for updates and enhancements.

New Journal of Physics
The open access journal at the forefront of physics

PAPER

CrossMark

Photon energy-resolved velocity map imaging from spectral domain ghost imaging

Jun Wang[1,2,*] , Taran Driver[1,3,*] , Felix Allum[1] , Christina C Papadopoulou[4] , Christopher Passow[4] , Günter Brenner[4] , Siqi Li[3] , Stefan Düsterer[4] , Atia Tul Noor[4] , Sonu Kumar[4] , Philip H Bucksbaum[1,2,3] , Benjamin Erk[4] , Ruaridh Forbes[1,3] and James P Cryan[1,3,*]

1   Stanford PULSE Institute, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, United States of America
2   Department of Applied Physics, Stanford University, Stanford, CA 94305, United States of America
3   SLAC National Accelerator Laboratory, Menlo Park, CA 94025, United States of America
4   Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany
*   Authors to whom any correspondence should be addressed.

E-mail: junwang9@stanford.edu, tdriver@stanford.edu and jcryan@slac.stanford.edu

Abstract

We present an approach that combines photon spectrum correlation analysis with the reconstruction of three-dimensional momentum distribution from velocity map images in an efficient, single-step procedure. We demonstrate its efficacy with the results from the photoionization of the $2p$-shell of argon using the Free-electron LASer in Hamburg free-electron laser (FEL). Distinct spectral features due to the spin-orbit splitting of $Ar^+(2p^{-1})$ are resolved, despite the large average bandwidth of the ionizing pulses from the FEL. This demonstrates a clear advantage over the conventional analysis method, and it will be broadly beneficial for velocity map imaging experiments with FEL sources. The retrieved linewidth of the binding energy spectrum approaches the resolution limitation 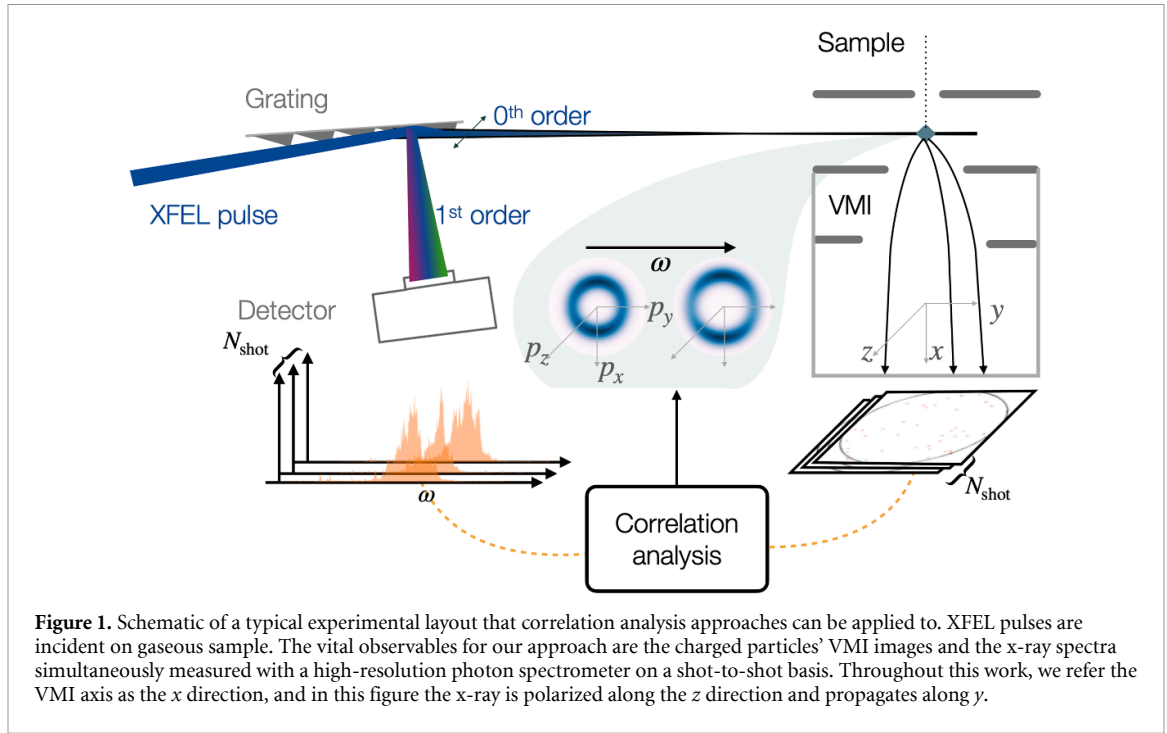prescribed by the spectrometers used to collect the data. Our approach presents a path to extend spectral-domain ghost imaging to the case where the photoproduct observable is high-dimensional.

1. Introduction

Modern x-ray spectroscopy provides a sensitive probe of local electronic density in molecular systems with atomic-site specificity [1–3]. As a result, x-ray free-electron lasers (XFELs) with continuous wavelength tunability throughout the soft x-ray region, unparalleled peak brightness, and short temporal duration, have enabled significant advances in time-resolved measurements of molecular dynamics [4–7]. Oftentimes the intrinsic photon energy jitter of a self-amplified spontaneous emission (SASE) XFEL, along with the strongly fluctuating sub-structure, is thought to be the limiting factor for the achievable energy resolution in time-resolved spectroscopy [6–8]. However, the spectral fluctuations inherent to SASE operation can, in fact, be exploited as a notable advantage by correlating x-ray observables with properties of the incident pulse on a shot-to-shot basis. This is a powerful approach to improving the resolution of experiments at XFEL facilities, beyond the bandwidth limit. In particular, the application of the spectral-domain ghost imaging (SDGI) technique has demonstrated sub-bandwidth resolution for spectroscopies employing XFELs [9–12].

Thanks to their high throughput, 4π collection solid-angle, and ability to provide angle-resolved photoproduct yields, velocity map imaging (VMI) spectrometers [13] have emerged as a popular instrument for time-resolved XFEL studies of dynamics in gas phase systems [3, 6, 7, 13–17]. A VMI spectrometer typically measures the two-dimensional (2D) projection of the three-dimensional (3D) momentum distribution of charged particles. In most cases, however, the quantity of interest is the underlying 3D distribution or another related quantity such as the kinetic energy (KE) spectrum, access to which requires post-processing of the 2D projection which inverts the Abel transform [18–22]. In trying to apply SDGI to

Figure 1. Schematic of a typical experimental layout that correlation analysis approaches can be applied to. XFEL pulses are incident on gaseous sample. The vital observables for our approach are the charged particles' VMI images and the x-ray spectra simultaneously measured with a high-resolution photon spectrometer on a shot-to-shot basis. Throughout this work, we refer the VMI axis as the *x* direction, and in this figure the x-ray is polarized along the *z* direction and propagates along *y*.

retrieve the photon energy-resolved KE spectrum of charged particles from a VMI data set, the momentum projection, inherent to VMI operation, could complicate the procedure. A feature with a specific KE is mapped to a broad range of pixels on the VMI detector, and thus features with distinct KEs will produce heavily overlapping features in the projected distribution. This is in contrast with previous applications of SDGI, where KE is more directly related to the experimental measurement [10]. In this work, we present a single-step regression approach that can simultaneously reconstruct the spectral response and the 3D momentum distribution by exploiting correlation between the single-shot VMI image and the corresponding photon spectrum, demonstrating that the projection inherent in the VMI concept is not detrimental to the SDGI procedure.

The primary components of a typical experiment where correlation analysis can be applied are illustrated in figure 1. Briefly, a stochastic light source (here the SASE XFEL [23]) produces pulses that intercept a grating. The reflected beam (zeroth order) is directed toward the interaction point of a VMI spectrometer, which records the single-shot photoelectron momentum distribution. The incident spectrum is measured from the first order diffraction of the grating [24].

In section 2, we describe our single-step regression approach. In section 3, we demonstrate it with an experiment at the Free-electron LASer in Hamburg (FLASH), in which spectral resolution below the average bandwidth of the incident x-ray pulses has been achieved. We point out that sub-bandwidth resolution is a common advantage of photon spectrum correlation analysis approaches. In section 4, we describe the advantage we observe in regularizing with a single-step reconstruction compared to other possible implementations. Our approach promises to enhance the energy resolution of XFEL experiments using VMI, opening up opportunities to study time-dependent phenomena through the variation in finely resolved energy structures [7, 25].

2. Method

2.1. Model

The primary quantity of interest in VMI measurements is the 3D momentum distribution of charged particles, $f(\boldsymbol{p})$, which depends on the spectral profile of the incoming pulse, according to

$$f(\boldsymbol{p}) = \int \chi(\boldsymbol{p},\omega)a(\omega)\mathrm{d}\omega, \tag{1}$$

where $a(\omega)$ is the frequency spectrum of an incident pulse, and $\chi(\boldsymbol{p},\omega)$ is the (linear) response to the spectral intensity at frequency ω. This model applies to light-matter interactions in the linear regime, such as single-photon ionization. The image obtained in a VMI spectrometer is given by the Abel transform of f with a mapping from momentum to position $\boldsymbol{r} = \boldsymbol{p}/\sqrt{2m\alpha}$,

Figure 2. Schematic of the overall analysis procedure. Given a single-shot spectral profile, our model composes a 3D momentum distribution based on parameters *C* and conducts the Abel transform, obtaining an expected 2D distribution to be compared with the corresponding VMI image. Tuning the model parameters arrives at an optimal point of the objective function that consists of the total discrepancy and the regularization terms. The optimal set of parameters is eventually converted to the photon frequency-dependent momentum distribution.

$$b(y,z) = \int \tilde{\chi}(\boldsymbol{r},\omega)a(\omega)\mathrm{d}x\mathrm{d}\omega , \tag{2}$$

where x is the direction of projection, $\alpha \propto mt_{\mathrm{ToF}}^{-2}$ is the magnification factor depending on the time-of-flight t_{ToF} and particle mass m, and $\tilde{\chi}(\boldsymbol{r},\omega) = (2m\alpha)^{3/2}\chi(\boldsymbol{p},\omega)$ is the ω-dependent momentum distribution.

Our analysis approach is summarized graphically in figure 2. The spectral response of the sample, $\chi(\boldsymbol{p},\omega)$, is extracted by fitting a model to a combination of single-shot VMI images and photon spectra. The most likely $\tilde{\chi}$ is the function that minimizes the difference between the measured and predicted electron momentum distributions, subject to regularization terms that favor sparsity and smoothness in the ω-dependent distribution $\tilde{\chi}$. We restrict our considerations to cases with cylindrical symmetry, where the axis of symmetry (z), is oriented perpendicular to the direction of projection (x), such as the case shown in figure 1. Other configurations satisfying the same symmetry requirements are discussed in section 4.4. Such a model is sufficient to describe experiments using either linearly or circularly polarized laser pulses acting on isotropic samples. The Abel transform is uniquely invertible under this symmetry condition, and the quality of the reconstruction will depend on the number of FEL shots used.

2.2. Implementation
In general, the inverse Abel transform is ill-conditioned, being susceptible to experimental noise [26]. Several inversion algorithms [18–22] have been developed to mitigate this issue and to robustly reconstruct the underlying 3D momentum distribution. In our method we adopt the basis functions employed in the well-known pBasex [18] algorithm to represent $\tilde{\chi}(\boldsymbol{r},\omega)$ and combine the inversion procedure with spectral domain ghost imaging [9, 10]. pBasex expands the 3D momentum distribution with a basis set that is the product of radial basis functions and Legendre polynomials:

$$\tilde{\chi}(\boldsymbol{r},\omega) = \frac{1}{2\pi}\sum_{l=0}^{N_L}\sum_{k=1}^{N_K}c_{lk}(\omega)f_k(r)P_l\left(\frac{z}{r}\right) , \tag{3}$$

where $P_l(z/r)$ is the lth order Legendre polynomial, and the sum over l can be truncated at twice the highest order of light–matter interaction N_L (i.e. twice the number of photons involved in the ionization process). $\{f_k(r)\}_{k=1}^{N_K}$ is a set of radial basis functions , and $c_{lk}(\omega)$ are their ω-dependent coefficients.

Combining equation (3) with equation (2), the projected 2D distribution can be written as:

$$b(y,z) = \sum_{l=0}^{N_L}\sum_{k=1}^{N_K}\int c_{lk}(\omega)G_{lk}(y,z)a(\omega)\mathrm{d}\omega , \tag{4}$$

where $G_{lk}(y,z) = \int f_k(r) P_l(z/r) \mathrm{d}x/(2\pi)$, is the projection of each basis function perpendicular to the axis of cylindrical symmetry. It is worth noting, that after discretization on a Cartesian grid $\boldsymbol{R}_q = (y_q, z_q)$, the projections $G_{lk,q} = G_{lk}(y_q, z_q)$ are the same integrals encountered in the standard pBasex inversion method (equation (4) of [18]). Being further uniformly discretized in photon energy ω_w, equation (4) is

$$\hat{B}_{i,q} = \sum_{l=0}^{N_L} \sum_{k=1}^{N_K} \sum_{w=1}^{N_\omega} G_{lk,q} A_{i,w} C_{lk,w},\tag{5}$$

where for each single-shot, i, $\hat{B}_{i,q}$ is the expected intensity at pixel q given the spectral profile $A_{i,w} = A_i(\omega_w)$ and the spectral response of the sample encoded in the coefficients $C_{lk,w} = C_{lk}(\omega_w)$. Solving for the coefficients C from the expected image $\hat{B}$ is straightforward, as the tensor $H \equiv G \otimes A$ can be reshaped into a matrix $H_{iq,lkw} = G_{lk,q} A_{i,w}$, whose pseudo-inverse maps $\hat{B}$ to C. However, solving C from the measured images B by application of the pseudo-inverse, is highly sensitive to experimental noise in both B and photon spectra A. A more robust approach is to minimize an objective function, h consisting of h_0, which quantifies the model-measurement discrepancy, and regularization terms favoring expected qualities of $C_{lk,w}$, such as sparsity and smoothness [10], which is similar to the procedure discussed in [20]. In our case, $h(C)$ is

$$h(C) = h_0(C; A, G, B) + \lambda_{\mathrm{sp}} h_{\mathrm{sp}}(C) + \lambda_{\mathrm{sm},\omega} h_{\mathrm{sm},\omega}(C) + \lambda_{\mathrm{sm,r}} h_{\mathrm{sm,r}}(C)\ ,\tag{6}$$

where

$$h_0(C; A, G, B) = \sum_i \sum_q W_q \left| \hat{B}_{i,q}(C; A, G) - B_{i,q} \right|^2\tag{7}$$

is the Gaussian log-likelihood up to a global factor, with weights W_q over each pixel q in the VMI image. Weights W_q can be chosen to enhance the sensitivity of the reconstruction to certain regions of the VMI image, and henceforth we denote W to be the diagonal matrix constructed by them. $h_{\mathrm{sp}}(C)$ and $h_{\mathrm{sm},\cdot}(C)$ are the sparsity and smoothness regularization terms, respectively.

We separate smoothness into two terms $h_{\mathrm{sm},\omega}(C)$ and $h_{\mathrm{sm,r}}(C)$, to differentiate smoothness along the frequency (ω), and radial (r) directions. The corresponding hyperparameters, $\lambda_{\mathrm{sp}}, \lambda_{\mathrm{sm},\omega}, \lambda_{\mathrm{sm,r}}$ control the degree to which sparsity and smoothness are enforced in the retrieved $\tilde{\chi}$. For each direction in $d = \omega, r$, the smoothness term quantifies the roughness with the second order difference, i.e. $h_{\mathrm{sm,d}}(C) = \|L^{(d)} C\|_2^2$, $d = \omega, r$, with $L^{(d)}$ representing the finite-difference Laplacian operator along direction d. Common choices for the form of h_{sp} include the L_1-norm [27] and L_2-norm squared [28]. We choose the latter for the demonstration in section 3. We discuss strategies for choosing the proper values for the hyperparameters in the supplementary material.

From the point of view of implementation, the discrepancy term, $h_0(C; A, G, B)$, is quadratic in the C coefficients. Conducting the summation over shot i to formulate h_0 into a quadratic form prior to the optimization significantly improves the efficiency, although at the cost of caching a large matrix $A^T A \otimes G W G^T$ in memory.

3. Results

We demonstrate this reconstruction method for a dataset collected from soft x-ray ionization of argon above the $2p$ ionization threshold [29–31]. The experiment is conducted using the CAMP instrument [32] at beamline BL1 of the FLASH [23]. Argon gas is introduced via supersonic expansion to produce a continuous molecular beam. Following collimation through two skimmers and an aperture, the beam is intercepted by focused x-ray pulses produced by the FEL in the interaction region of the spectrometer. The velocities of the photoelectrons are mapped to a position-sensitive microchannel plate/phosphor screen detector and recorded with a CMOS camera at 10 Hz.

The pulses have an average bandwidth of 5 eV full-width at half maximum (FWHM) at 264 eV central photon energy. The estimated pulse duration is 100–150 fs (FWHM) with a mean pulse energy of 6 μJ/pulse at the sample. The incident photon spectrum is recorded on a shot-to-shot basis using an upstream variable line spacing (VLS) spectrometer [24]. The zeroth order beam from the grating is delivered to the interaction point, while the first order is collected by a detector to image the spectrum.

This measurement is performed with linearly polarized light from the FEL, which implies that the resulting electron momentum distribution has reflection symmetry about the xy plane. In this case, we can limit the sum in equation (3) to only even-order Legendre polynomials [33]. The dataset consists of 2.8×10^4 FEL shots, and the representative single-shot raw VMI images and photon spectra are shown in
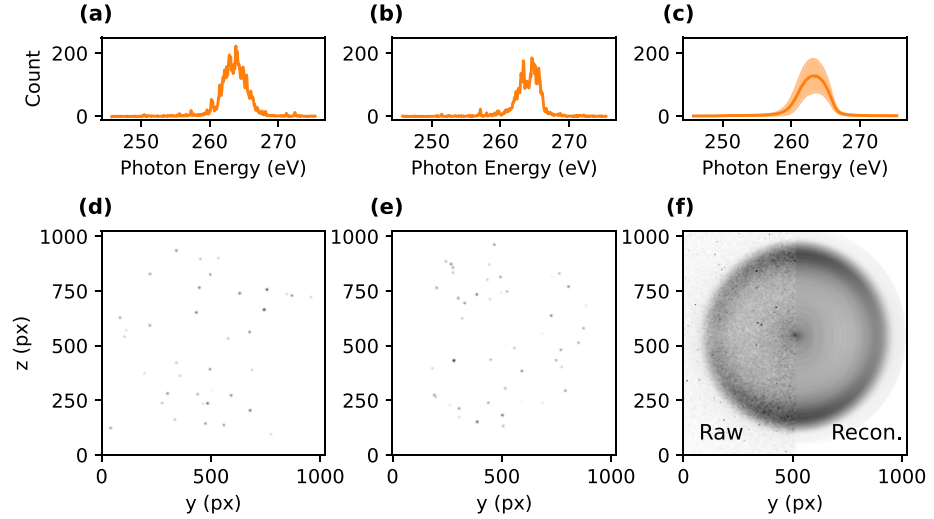
Figure 3. Representative single-shot data and average spectra. (a), (b) Two single-shot photon spectra after a global background subtraction, with (d), (e) showing their corresponding VMI images. For visualization only, we have blurred the raw images with a 3px-wide (standard deviation) gaussian kernel. (c) Average photon spectrum, with the shade representing the ± 1 standard deviation over all FEL shots. (f) Average VMI image, where the left half is the average raw image, and the right half is the reconstruction of the average image (see main text).
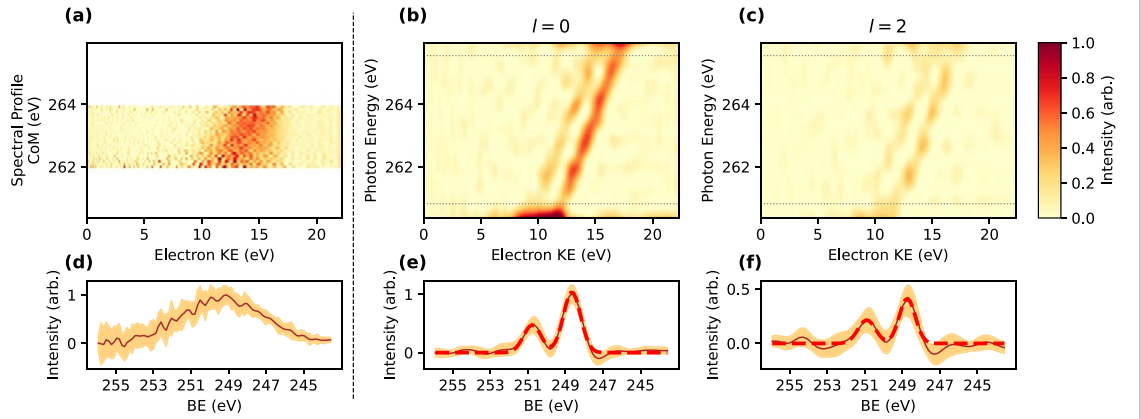

Figure 4. Comparison between the conventional binning-and-averaging method (a)(d) and our method (b)(c)(e)(f). (a) Energy distribution obtained with pBasex applied to the cumulative VMI image, with preliminary grouping of shots by the center-of-mass (CoM) of spectral profile. (b), (c) Photon-energy resolved argon $2p$ photoelectron KE spectra of the $l = 0$ and $l = 2$ Legendre polynomial components of the angular distribution, both being normalized to the maximum intensity in $l = 0$. The grey dotted lines delineate the photon energy range for (e) and (f). (d)–(f) Binding energy (BE) spectra of (a)–(c) extracted by shifting the KE spectra at each photon energy, see main text. The brown solid lines are the mean over the photon energy range, and the shaded regions represent the standard deviation. The red dashed lines in (e) and (f) are the sum of fitted Gaussians.

figure 3, where the average VMI image and photon spectrum are also shown. The single-shot VMI images are quite sparse, containing 40 electrons in each image on average.

From the minimization of equation (6), we obtain the most probable parameters $C_{lk,w}$ given the dataset, which represents the fitted $\chi(\boldsymbol{p}, \omega)$. Using equation (5), the average image is reconstructed by averaging the expected intensity $\hat{B}_{i,q}$ across all FEL shots. As shown in figure 3(f), the reconstructed average image agrees well with the average of raw images except being less noisy. The reconstructed average image, nevertheless, is the projected momentum distribution of the mean photon spectrum and can hardly show the underlying fine structure. With the same set of most-probable parameters $C_{lk,w}$, a better way to reveal the fine structure is to reconstruct the KE spectrum for each lth Legendre polynomial component and each photon energy ω, which is discretized as

$$I_l(E_r, \omega_w) = \sum_k C_{lk,w} f_k\left(\sqrt{\frac{E_r}{\alpha}}\right)\sqrt{\frac{E_r}{4\alpha^3}}\,, \tag{8}$$

with $E_r = \alpha r^2$ denotes the KE at radial grid point r. The ionization from argon $2p$-shell is a single-photon process, so we visualize $I_0(E, \omega)$ and $I_2(E, \omega)$, in figures 4(b) and (c) respectively.

We observe two dispersive features that correspond to the spin-orbit split cationic states of argon, $^2P_{1/2}$ and $^2P_{3/2}$ [31]. For each photon energy in figure 4(b), we shift the reconstructed KE spectrum to binding energy (BE) according to $BE = \hbar\omega - KE$, which is shown in figure 4(e). The spin-orbit splitting is well resolved with our approach. We use a previously reported measurement [31] of this splitting, 2.12 eV, to calibrate the VMI and VLS energy axes. With this calibration, we extract a linewidth of 1.1 eV FWHM from figure 4(e), which is resolved beyond the width of the averaged FEL pulse spectrum, 5 eV FWHM. Furthermore, the extracted linewidth is below the average bandwidth of a single pulse 3.5 eV FWHM. Thus, with the application of spectral domain ghost imaging in this XFEL experiment, we are able to achieve sub-bandwidth resolution.

In figures 4(c) and (f) we plot the projection of the 3D momentum distribution onto the second-order Legendre polynomial $I_2(E, \omega)$, which is related to the photoemission anisotropy parameter, β_2 [34]. We extract a value of 0.46 ± 0.05 and 0.39 ± 0.05 for the anisotropy parameter for the $^2P_{1/2}$ and $^2P_{3/2}$ ionic states, respectively. These values are averaged over a range of photon energies from 261 to 265 eV, which is in good agreement with previous measurements in [30] and calculations in [35]. The results in figures 4(b), (c), (e) and (f) are obtained with a set of $N_K = 125$ radial basis functions and uniform pixel weight W.

In contrast, directly applying pBasex on the cumulative image results in a single spectral feature with a width of 5.4 eV FWHM. Moreover, grouping the FEL shots by the center-of-mass of photon spectrum does not resolve the two spin-orbit features, and the corresponding average BE spectrum has a width of 5.1 eV (FWHM), as shown in figures 4(a) and (d). In order to obtain the resolution observed in figures 4(b) and (c), we need to explicitly consider the correlations between the single-shot spectra and the VMI images.

4. Discussion

4.1. Resolution achieved in our demonstration experiment

To characterize the resolution of this technique, it is useful to define the correlation length of the spectral measurement. The correlation length is a measure of how strongly coupled the intensity fluctuations are between neighbouring pixels in the photon spectrum. Here, we define this metric δ_A to be the average distance between two frequencies where the Pearson's correlation coefficient drops to $1/e$ [9]. The correlation between spectrometer pixels arises from two sources: the resolution of the photon spectrometer, and the intrinsic variation due to spectral fluctuations of the source. In our experiment, the correlation length is measured to be $\delta_A = 22$ pixels, corresponding to 0.99 eV, as shown in figure S1(a) of the supplementary material. This is much larger than the instrumental resolution of the photon spectrometer [24], and thus, we conclude that the dominant contribution to the correlation length is the intrinsic correlation of the FEL source.

The average width of the photoemission features in the BE spectrum in figure 4(e) is determined to be $\sigma_{BE} = 0.48$ eV in standard deviation. This width is the result of the finite VMI energy resolution and the correlation in the spectral measurement, which approximately sum in quadrature $\sigma_{BE}^2 \approx \sigma_{VMI}^2 + \sigma_{\phi Corr}^2$. The contribution from spectral correlation, $\sigma_{\phi Corr}$, is not necessarily the correlation length δ_A, but the two quantities are related by a constant of proportionality. As described in the supplemental material, and shown in figure S1(b), $\sigma_{\phi Corr}^2$ is estimated to be 0.11 eV^2 based on the measured $\delta_A = 22$ px. Having accounted for $\sigma_{\phi Corr}^2$, we find the remaining contribution to σ_{BE}^2 to be $\sigma_{VMI}^2 = 0.10$ eV^2. This corresponds to a VMI energy resolution of 2.3% at KE = 15 eV, which is consistent with the reported value of 2% for this instrument [32]. We note that the average bandwidth of incident x-ray pulses is not a limiting factor for the resolution σ_{BE} of our technique. Instead, it is determined by the KE resolution of the VMI spectrometer and the correlation length of the photon spectrum measurement.

4.2. Extension of SDGI

Spectral domain ghost imaging, in general, exploits spectral fluctuations of the incident source to capture the sample response. The key development in the present work is incorporating the native projection of VMI into the ghost imaging model. This method unravels the projection of different Newton spheres, while simultaneously correlating these unravelled shells with the incident photon spectra. More specifically, our approach models the photoproduct momentum space with a reduced dimensional representation (similar to the pBasex approach) and inverts the tensor product of spectral integration and Abel transform in a single step. Such an approach is a unique way to extend the scope of SDGI. Although under ideal circumstances it is equivalent to conducting the two regression steps sequentially, the single-step approach has several advantages over either sequential approach.

In comparison to the sequential application of SDGI followed by Abel inversion, the single-step approach represents the momentum distribution with a polar basis set, reducing the dimensionality of the problem and coordinating the comprehensive behavior of the momentum distribution over multiple pixels. When

applying SDGI directly to the 2D VMI data, incorporating the smoothness term that penalizes high spatial frequencies in momentum space is inefficient and in most cases intractable. This is because the dimension of the image space is too large. Conceding this smoothness over momentum space, we can implement a sequential approach by applying SDGI to each VMI pixel independently and subsequently applying pBasex to the result. This approach remains efficient, but it is plagued by compromised performance in noise handling, as shown in the supplementary material figure S2. The key difference between this sequential approach and the single-step is that the regression is performed independently for each pixel in the first of the two sequential steps.

Another possible sequential approach is to apply single-shot Abel inversion followed by SDGI. Although the regularization terms can remain in the same form as in the single-step approach, in this approach the Gaussian log-likelihood in the objective function (equation (6)) is replaced by

$$h_{0,\mathrm{p+S}}(C;A,G,B) = \sum_i \left| \sum_w A_{i,w} C_{lk,w} - \sum_q \left(G^+\right)_{q,lk} B_{i,q} \right|^2 , \tag{9}$$

where $G^+ = (GWG^T)^{-1}GW$ is the pseudo-inverse of the Abel transform, which is conducted separately for each photon energy bin.

In our demonstration experiment, both sequential approaches show inferior performance compared to the single-step approach, which is illustrated in the comparison in the supplementary material. Under ideal circumstances, when regularization becomes unnecessary to suppress experimental noise, the optimal point of the original objective function in equation (6) is

$$C_{lk,w}^{(\lambda=0)} \quad = \quad \sum_{i,q} \left((G \otimes A)^+\right)_{wlk,iq} B_{i,q} \tag{10a}$$

$$= \quad \sum_{i,q} \left(A^+\right)_{w,i} \left(G^+\right)_{q,lk} B_{i,q} , \tag{10b}$$

where $A^+ = (A^TA)^{-1}A^T$, and in going from equations (10a) to (10b) we have changed the order of operations between the pseudo-inverse and tensor-product of A and G. Equation (10a) describes the single-step approach and equation (10b) denotes both the sequential approaches described above. Thus it can be seen that in the absence of regularization, both sequential approaches are equivalent to the single-step method. The regularization in equation (6) necessitates the computationally heavy step of inverting an $N_\omega N_L N_K$-dimensional matrix, which is not true for the sequential approaches. However, as described above, the single-step routine outperforms the sequential approaches by virtue of being more robust to noise. Similar results between the single-step and sequential approaches may be achieved in other measurements, and the precise choice of method depends critically on the properties of the measurement; e.g. signal-to-noise ratio, size of the dataset, and number of parameters to fit.

4.3. Connection to covariance approach

We note that there are a number of mathematical techniques to extract correlations in large datasets. While the method presented in this work relies on linear regression, covariance analysis is another common method for extracting correlations [7, 36–39]. The photoproduct yield can be correlated with other intrinsic or extrinsic measurements, to extract signal from noisy data. Both regression and covariance have been employed in ghost imaging experiments in the spectral domain, and these two techniques are closely related.

The connection is shown by regressing the mean-subtracted VMI images $\Delta B_i \equiv B_i - \langle B \rangle$ on the mean-subtracted spectra $\Delta A_i \equiv A_i - \langle A \rangle$, with $\langle \cdot \rangle$ denoting the average across all shots. Given the standard definition of sample covariance $\mathrm{Cov}[X,Y] = \Delta X^T \Delta Y/(N_{\mathrm{shot}} - 1)$, the unregularized spectral regression of ΔB on ΔA gives

$$(\Delta A)^+ \Delta B = (\Delta A^T \Delta A)^{-1}(\Delta A^T \Delta B) = \mathrm{Cov}[A,A]^{-1}\mathrm{Cov}[A,B] , \tag{11}$$

where the right hand equality demonstrates this is identical to applying the inverse of the photon autocovariance matrix to the photon-electron covariance matrix.

The transform from the projected momentum distribution to the coefficients C is a linear operation in the electron momentum space, which is in tensor product with the operations in the photon energy space. Therefore the aforementioned single-step and sequential regression approaches are all applicable to the mean-subtracted data. Comparing to the original approach elaborated in section 2, regressing the

mean-subtracted data is more robust to static background in both the VMI images and photon spectra, but it may suffer from the loss of information due to subtraction of the mean, especially when the spectral fluctuation is limited. Regardless whether the mean is subtracted or not, one can further constrain the parameters with additional prior knowledge about the sample, to solve the most probable parameters in a space with dimension lower than $N_\omega N_L N_K$. Although related, these further restricted approaches require more assumptions and are less generalizable than the main approach we present in this work. We provide an example with further discussion in the supplementary material.

4.4. Applicable apparatus configurations

As for standard inverse Abel transform procedures, a prerequisite of our method is cylindrical symmetry about any axis z that is perpendicular to the projection axis of the VMI x, for which the coordinate system shown in figure 1 is not the only configuration. Here we describe two other configurations, without exhausting all possible cases. For circularly polarized x-rays, such as in [40], the layout in figure 1 is still applicable except that the symmetry axis z is along the axis of beam propagation. For a co-axial VMI [41] with linearly polarized x-rays propagating along the VMI axis x, z falls along the x-ray polarization axis.

5. Conclusion and outlook

We present a regression approach which can achieve VMI measurements at resolution better than the inherent bandwidth of a noisy photon source, by simultaneously reconstructing the spectral response and performing the inverse Abel transform. Our approach demonstrates a clear advantage over the conventional binning-and-averaging method. In our experimental demonstration on the photoionization of argon, the retrieved linewidth is dominated by the resolution limit prescribed by the VMI resolution and the measured spectral correlation. We anticipate that the outlined approach will be of great use in the emerging field of time-resolved inner-shell photoelectron spectroscopy at FELs, which promises to interrogate nuclear and electronic dynamics in a site-selective manner.

   We demonstrate the connection between covariance and regression analysis and we relate and compare our single-step approach to other sequential approaches. Our method extends the scope of SDGI by adopting a reduced-dimensional representation for the properties of the photoproducts, and it can be directly applied to any measurement where VMI images are recorded in coincidence with the incident photon spectrum. This makes the technique broadly applicable to many different light sources.

Author contributions

T D, J P C and R F initiated the research. J W, T D and J P C conceived the method, which is implemented by J W. The experiment was carried out by C C P, C P, G B, S D, A T N, S K, and B E. All authors contributed to collection or analysis of the experimental data and writing of the manuscript.

Data availability statement

The data that support the findings of this study are available from the authors upon reasonable request. An implementation of the proposed approach can be found in our custom package developed for general SDGI [42]. All data that support the findings of this study are included within the article (and any supplementary material).

Acknowledgments

ORCID iDs

Jun Wang ⦿ https://orcid.org/0000-0003-1962-895X
Taran Driver ⦿ https://orcid.org/0000-0002-3761-6883
Felix Allum ⦿ https://orcid.org/0000-0002-8044-8969
Günter Brenner ⦿ https://orcid.org/0000-0002-1755-876X
Siqi Li ⦿ https://orcid.org/0000-0002-1262-6292
Stefan Düsterer ⦿ https://orcid.org/0000-0003-4379-1327
Atia Tul Noor ⦿ https://orcid.org/0000-0001-6614-3256
Benjamin Erk ⦿ https://orcid.org/0000-0001-8413-3588
Ruaridh Forbes ⦿ https://orcid.org/0000-0003-2097-5991
James P Cryan ⦿ https://orcid.org/0000-0002-7776-0919

References

[1] Siegbahn K 1982 *Rev. Mod. Phys.* **54** 709
[2] Picón A *et al* 2016 *Nat. Commun.* **7** 11652
[3] Erk B *et al* 2014 *Science* **345** 288
[4] Wolf T J A *et al* 2017 *Nat. Commun.* **8** 29
[5] Jay R M, Kunnus K, Wernet P and Gaffney K J 2022 *Annu. Rev. Phys. Chem.* **73** 187–208
[6] Brauße F *et al* 2018 *Phys. Rev. A* **97** 043429
[7] Allum F *et al* 2022 *Commun. Chem.* **5** 42
[8] Mayer D *et al* 2022 *Nat. Commun.* **13** 1
[9] Driver T, Li S *et al* 2020 *Phys. Chem. Chem. Phys.* **22** 2704
[10] Li S *et al* 2021 *J. Phys. B: At. Mol. Opt. Phys.* **54** 144005
[11] Kayser Y *et al* 2019 *Nat. Commun.* **10** 1
[12] Klein Y *et al* High-resolution absorption measurements with free-electron lasers using ghost spectroscopy 2022 (arXiv:2203.00688)
[13] Eppink A T J B and Parker D H 1997 *Rev. Sci. Instrum.* **68** 3477
[14] Squibb R J *et al* 2018 *Nat. Commun.* **9** 63
[15] Rouzée A *et al* 2011 *Phys. Rev. A* **83** 031401
[16] Fushitani M *et al* 2021 *J. Chem. Phys.* **154** 144305
[17] Prince K C *et al* 2016 *Nat. Photon.* **10** 176
[18] Garcia G A, Nahon L and Powis I 2004 *Rev. Sci. Instrum.* **75** 4989
[19] Roberts G M, Nixon J L, Lecointre J, Wrede E and Verlet J R R 2009 *Rev. Sci. Instrum.* **80** 053104
[20] Dick B 2014 *Phys. Chem. Chem. Phys.* **16** 570
[21] Montgomery Smith L, Keefer D R and Sudharsanan S 1988 *J. Quant. Spectrosc. Radiat. Transfer* **39** 367
[22] Vrakking M J J 2001 *Rev. Sci. Instrum.* **72** 4084
[23] Feldhaus J 2010 *J. Phys. B: At. Mol. Opt. Phys.* **43** 194002
[24] Brenner G *et al* 2011 *Nucl. Instrum. Methods Phys. Res.* **635** S99
[25] Inhester L, Li Z, Zhu X, Medvedev N and Wolf T J A 2019 *J. Phys. Chem. Lett.* **10** 6536
[26] Pretzier G, Jäger H, Neger T, Philipp H and Woisetschläger J 1992 *Z. Naturforsch.* **47** 955
[27] Tibshirani R 1996 *J. R. Stat. Soc. B* **58** 267
[28] Tikhonov A N, Goncharsky A V, Stepanov V V and Yagola A G 1995 Regularization methods *Numerical Methods for the Solution of Ill-Posed Problems* (Dordrecht: Springer Netherlands) pp 7–63
[29] Heimann P *et al* 1987 *J. Phys. B: At. Mol. Phys.* **20** 5005
[30] Avaldi L, Dawber G, Camilloni R, King G, Roper M, Siggel M, Stefani G and Zitnik M 1994a *J. Phys. B: At. Mol. Opt. Phys.* **27** 3953
[31] Jurvansuu M, Kivimäki A and Aksela S 2001 *Phys. Rev. A* **64** 012502
[32] Erk B *et al* 2018 *J. Synchrotron Radiat.* **25** 1529
[33] Yang S-c and Bersohn R 1974 *J. Chem. Phys.* **61** 4400
[34] Cooper J and Zare R N 1968 *J. Chem. Phys.* **48** 942
[35] Lindle D W *et al* 1988 *Phys. Rev. A* **38** 2371
[36] Frasinski L, Codling K and Hatherly P 1989 *Science* **246** 1029
[37] Frasinski L *et al* 2013 *Phys. Rev. Lett.* **111** 073002
[38] Frasinski L J 2016 *J. Phys. B: At. Mol. Opt. Phys.* **49** 152004
[39] Amiot C, Ryczkowski P, Friberg A T, Dudley J M and Genty G 2018 *Opt. Lett.* **43** 5025
[40] Ilchen M *et al* 2021 *Commun. Chem.* **4** 119
[41] Li S *et al* 2018 *AIP Adv.* **8** 115308
[42] Wang J 2022 Spook - a package of spectral-domain ghost imaging implementations (available at: https://github.com/congzlwag/spook)