xcalib: a focal spot calibrator for intense X-ray free-electron laser pulses based on the charge state distributions of light atoms[1]

Koudai Toyota,[a] Zoltan Jurek,[a] Sang-Kil Son,[a] Hironobu Fukuzawa,[b] Kiyoshi Ueda,[b] Nora Berrah,[c] Benedikt Rudek,[d] Daniel Rolles,[e] Artem Rudenko[e] and Robin Santra[a,f]*

[a]Center for Free-Electron Laser Science, DESY, 22607 Hamburg, Germany, [b]Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, Sendai, Japan, [c]Physics Department, University of Connecticut, Storrs, CT, USA, [d]Physikalisch-Technische Bundesanstalt, Braunschweig, Germany, [e]J. R. Macdonald Laboratory, Department of Physics, Kansas State University, Manhattan, KS, USA, and [f]Department of Physics, University of Hamburg, 20355 Hamburg, Germany. *Correspondence e-mail: robin.santra@cfel.de

The xcalib toolkit has been developed to calibrate the beam profile of an X-ray free-electron laser (XFEL) at the focal spot based on the experimental charge state distributions (CSDs) of light atoms. Characterization of the fluence distribution at the focal spot is essential to perform the volume integrations of physical quantities for a quantitative comparison between theoretical and experimental results, especially for fluence-dependent quantities. The use of the CSDs of light atoms is advantageous because CSDs directly reflect experimental conditions at the focal spot, and the properties of light atoms have been well established in both theory and experiment. Theoretical CSDs are obtained using xatom, a toolkit to calculate atomic electronic structure and to simulate ionization dynamics of atoms exposed to intense XFEL pulses, which involves highly excited multiple core-hole states. Employing a simple function with a few parameters, the spatial profile of an XFEL beam is determined by minimizing the difference between theoretical and experimental results. The optimization procedure employing the reinforcement learning technique can automatize and organize calibration procedures which, before, had been performed manually. xcalib has high flexibility, simultaneously combining different optimization methods, sets of charge states, and a wide range of parameter space. Hence, in combination with xatom, xcalib serves as a comprehensive tool to calibrate the fluence profile of a tightly focused XFEL beam in the interaction region.

1. Introduction

The recent X-ray free-electron laser (XFEL) technologies have enabled us to conduct experiments at ultrashort time scales (∼a few femtoseconds) and ultrahigh intensities (∼10^{20} W cm^{-2}), which are far beyond the domain of conventional synchrotron radiation sources (Schneider, 2010). Theoretical calculations have played a crucial role in revealing new ionization mechanisms of atoms and molecules found in experiments driven by such unprecedented light (Young *et al.*, 2010; Doumy *et al.*, 2011; Rudek *et al.*, 2012, 2013, 2018; Fukuzawa *et al.*, 2013; Motomura *et al.*, 2013; Murphy *et al.*, 2014; Rudenko *et al.*, 2017). In these discoveries, a quantitative comparison between theoretical and experimental results was crucial to elucidate the underlying physics. When an XFEL beam is focused onto a target in experiments, the fluence values of the beam have a non-uniform spatial distribution in the focal spot (Barty *et al.*, 2009). Thus, a range of fluence values covered by the distribution contributes to the yield of
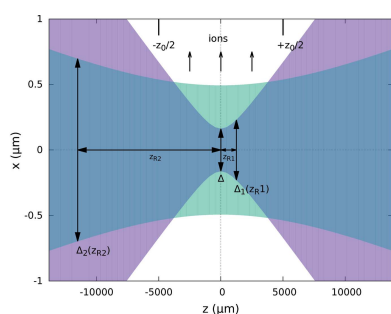
an observable such as ions, electrons and photons. When theoretically computing the yield of an observable, we need to add up all fluence-dependent contributions calculated with the actual fluence values, which is called volume integration (Young *et al.*, 2010). It is thus essential to characterize the spatial profile of the XFEL beam to perform the volume integration.

However, a direct measurement of the focal volume parameters for XFELs represents a significant experimental challenge. The ablation imprint method (Chalupský *et al.*, 2007, 2010, 2011, 2015; Rösner *et al.*, 2017) has been used to characterize spatial properties of a tightly focused X-ray beam, but it cannot be employed *in situ*, which hinders its routine use in user experiments. The Hartmann wavefront sensing method (Le Pape *et al.*, 2002; Schäfer *et al.*, 2006; Flöter *et al.*, 2010; Keitel *et al.*, 2016) can provide an *in situ* single-shot characterization of a spatial beam profile, but it is unsuitable for a high-fluence X-ray beam. Also the Ronchi test (Nilsson *et al.*, 2012; Nagler *et al.*, 2017), grating interferometry (Rutishauser *et al.*, 2012; Kayser *et al.*, 2014), curved grating monitors (Schneider *et al.*, 2016, 2018), diffraction with aerosol spheres (Loh *et al.*, 2013), ptychographic imaging (Schropp *et al.*, 2013), iterative diffractive imaging (Mehrjoo *et al.*, 2017) and speckle visibility spectroscopy (Kobayashi *et al.*, 2018) have been utilized for monitoring the wavefront and focal spot. Even though the above-mentioned methods provide the beam shape and, hence, the shape of the spatial fluence distribution, most of them do not provide the actual fluence values on an absolute scale, which are required for the volume integration and quantitative comparison between experiment and theory.

Here we present a method for calibrating the fluence distribution utilizing experimental and theoretical charge state distributions (CSDs) of light atoms such as neon (Ne) or argon (Ar) atoms. Our aim is not a shot-by-shot characterization of the spatial fluence distribution. Instead, we calibrate the fluence distribution that reflects a property of an ensemble of XFEL pulses, in order to compare pulse-ensemble-averaged quantities such as ion time-of-flight (ToF) spectra accumulated over many shots. In addition, instead of mapping the spatial fluence distribution as a function of position, we assume a specific functional form of the spatial fluence profile at the focal spot depending on a few parameters, so that we calibrate a fluence distribution function (FDF), *i.e.* a histogram of fluence values, which actually governs the volume integration. The technique of measuring CSDs of light rare-gas species can be used for an *in situ* fluence characterization that provides, in principle, real-time information during a beam time. This approach has already been employed in several FEL experiments (Young *et al.*, 2010; Doumy *et al.*, 2011; Rudek *et al.*, 2012, 2013; Fukuzawa *et al.*, 2013; Motomura *et al.*, 2013; Murphy *et al.*, 2014; Rudenko *et al.*, 2017), because it is relatively straightforward and can be used during any experimental run that has ion ToF detection at its disposal. The main obstacle to turning the CSD technique into a routine tool available to users is the difficulty with extracting the fluence distribution from experimental CSDs. Overcoming this

limitation is the motivation behind the development of *xcalib* presented here.

Using CSDs of light atoms has three advantages. First, because of the high non-linearity of the XFEL interactions with atoms, these CSDs are very sensitive to the peak fluence value as well as to the spatial fluence profile in the focal spot. Second, we can utilize the well established atomic properties of light atoms. Third, calculating the CSDs of light atoms is computationally cheap. In the present work, the CSDs of atoms were calculated using the *xatom* toolkit (Son *et al.*, 2018; Toyota *et al.*, 2017). Recently, a calibration procedure at low and intermediate fluences based on fragment ion spectra of Ar clusters has been proposed (Kumagai *et al.*, 2018). The CSDs of Ne and Ar atoms are often used as fast experimental feedback for minimizing the focal spot size when changing the focusing mirror settings (Schorb, 2012).

In previous studies, the focal volume parameters have been determined by minimizing a certain measure by manually exploring the parameter space. However, such manual procedures lack efficient algorithms to obtain an optimized solution, and are insufficient to handle a large number of experimental results in a wide range of parameter space. The situation motivated us to develop a toolkit to automatize the optimization procedures, employing the reinforcement learning technique (Raschka, 2015). A machine-learning technique has been used in single-shot characterization of spectral and temporal profiles of XFEL pulses (Sanchez-Gonzalez *et al.*, 2017). These optimization methods have an advantage in finding the direction to a solution in the parameter space with efficient algorithms. We designed the *xcalib* toolkit to have flexibility simultaneously combining different pulse profiles, parameter ranges, charge states and optimization methods. Therefore, *xcalib* offers a comprehensive tool to calibrate X-ray beam parameters in XFEL experiments. *xcalib* will be made available to the scientific community in the next release of *xraypac* (Jurek *et al.*, 2016a).

The paper is organized as follows. In Section 2, we introduce our numerical method focusing on the volume integration and the optimization method. In Section 3, we revisit the three Ar calibrations of Fukuzawa *et al.* (2013), Murphy *et al.* (2014) and Rudek *et al.* (2018) to show that calibrated results by *xcalib* are consistent with the previous results. We also study the effect of attenuators on the fluence profile used by Rudek *et al.* (2018). In Section 4, we conclude the paper with a summary. We use atomic units throughout the paper unless stated otherwise.

2. Numerical method

2.1. Volume integration

In this section, we first formulate the numerical procedure for the volume integration for ion yield distributions, which is essential to obtain theoretical results that may be compared with experimental results. The ion yield $\mathcal{Y}_{\text{theo}}^{(+q)}$ of charge state q produced at position $\mathbf{r}$ in the focal spot is assumed to be a function of the position-dependent fluence value $F(\mathbf{r}; \mathbf{P})$,

where a set of parameters **P** characterizes the ensemble-averaged spatial profile. The volume-integrated yield is then given by the following three-dimensional integral,

$$Y_{\text{theo}}^{(+q)}(\mathbf{P}) = \int \mathcal{Y}_{\text{theo}}^{(+q)}[F(\mathbf{r}; \mathbf{P})] \, \mathrm{d}^3 r. \qquad (1)$$

We normalize $Y_{\text{theo}}^{(+q)}(\mathbf{P})$ as follows,

$$y_{\text{theo}}^{(+q)}(\mathbf{P}) = \frac{1}{N(\mathbf{P})} \, Y_{\text{theo}}^{(+q)}(\mathbf{P}), \qquad (2)$$

where the constant $N(\mathbf{P})$ is a normalization factor given by

$$N(\mathbf{P}) = \sum_{\text{all } q} Y_{\text{theo}}^{(+q)}(\mathbf{P}). \qquad (3)$$

The summation runs over the set of all available charge states. Hence, the sum over the normalized ion yields is unity,

$$\sum_{\text{all } q} y_{\text{theo}}^{(+q)}(\mathbf{P}) = 1. \qquad (4)$$

In the following, ion yields refer to the normalized ion yield defined in equation (2), unless specified otherwise.

The theoretical ion yields $\mathcal{Y}_{\text{theo}}^{(+q)}$ are calculated by employing the *xatom* toolkit. *xatom* is a set of computer codes to calculate the electronic structure of atoms based on the Hartree–Fock–Slater (HFS) method and to describe multiphoton multiple ionization dynamics during intense XFEL pulses employing a rate-equation approach. *xatom* has been tested with a series of gas-phase atomic experiments and has played a crucial role in many XFEL applications [see Jurek *et al.* (2016b) and references therein]. Note that the theoretical ion yields can also be obtained by other tools, for example, *scfly* (Chung *et al.*, 2007; Ciricosta *et al.*, 2011), *averroès/transpec* (Peyrusse, 2000; Peyrusse *et al.*, 2014), *dlayz* (Xiang *et al.*, 2012; Gao *et al.*, 2013, 2015) and *mcre* (Ho *et al.*, 2014, 2015).

Here we assume that the pulse duration is an experimentally known parameter and that the temporal pulse profile is constant throughout the interaction region. XFELs based on the SASE (self-amplified spontaneous emission) principle produce chaotic pulses (Rohringer & Santra, 2007). In the X-ray beam parameter regime considered in the present study, however, the CSDs are insensitive to temporal pulse shapes, and CSDs can be calculated with an ensemble-averaged temporal pulse shape (Rohringer & Santra, 2007). Thus, we use a Gaussian function for the ensemble-averaged temporal profile, and the nominal pulse duration taken from the experiment is used as the full width at half-maximum (FWHM) in the *xatom* calculations in order to calculate fluence-dependent CSDs. In the frustrated-absorption regime (Young *et al.*, 2010; Hoener *et al.*, 2010), where the pulse duration competes with the lifetime of core-hole states, the CSDs become sensitive to the pulse duration. Under such conditions, the ion yields before and after volume integration depend on the pulse duration τ, *i.e.* $\mathcal{Y}_{\text{theo}}^{(+q)}[F(\mathbf{r}; \mathbf{P}), \tau]$ and $Y_{\text{theo}}^{(+q)}(\mathbf{P}, \tau)$, and then one could treat τ as an additional fit parameter for calibration together with **P**. The experimentally known pulse duration (30 fs FWHM) used in the present work is, however, much longer than the calculated lifetime of core-

hole states; for example, ~1 fs for $Ar^{+}(1s^{-1})$ and ~12 fs for $Ar^{15+}(1s^1 2s^2)$, and is, thus, way beyond the frustrated-absorption regime. Around this experimental pulse duration, the calculated CSDs show almost no dependence on the pulse duration as demonstrated in Appendix A, and therefore we fix the pulse duration in the following results.

2.2. Single Gaussian spatial profile

To perform the volume integration, equation (1), we need to model the spatial fluence profile, $F(\mathbf{r}; \mathbf{P})$, to map a given position $\mathbf{r}$ to a fluence value in the interaction volume. The interaction volume is defined by the intersection between the XFEL beam and the target gas jet (atomic or molecular beam). If the target beam size (typically ~mm) is larger than the XFEL beam size (~µm or less), the XFEL beam determines the shape of the intersection in the direction transverse to the XFEL beam propagation. Since the tightly focused XFEL beam diverges with increasing distance z from the focus, the shape of the interaction volume is similar to an hourglass laid on the z-axis, as shown in purple in Fig. 1. Assuming Gaussian beam optics, the diameter of the cross section transverse to the z-axis is given by

$$\Delta_1(z) = \Delta\left[1 + (z/z_{R1})^2\right]^{1/2}, \qquad (5)$$

where the quantity Δ represents the focal spot size, and z_{R1} represents the Rayleigh range. Let λ be the wavelength of the XFEL beam. The Rayleigh range is then given by
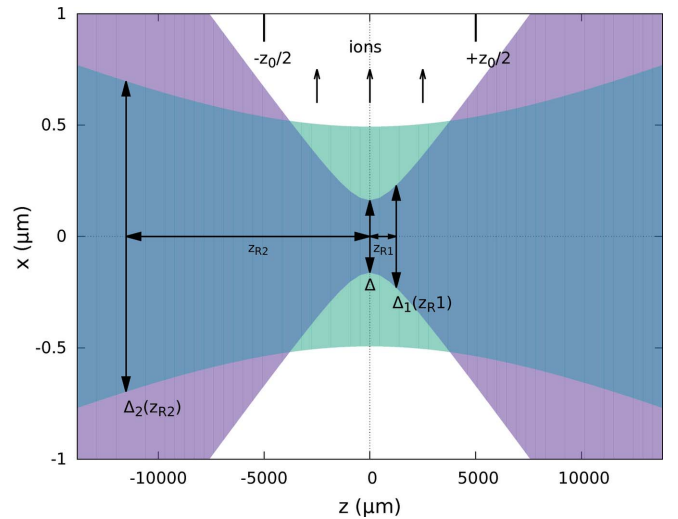

Figure 1
The beam geometry in the xz plane. The X-ray beam direction is along the z axis, and the target gas jet flows along the y axis. When the highest peak fluence is desirable, the target gas is located at $z = 0$. For the single Gaussian spatial profile, the beam shape in the xz plane is shown in purple. For the double Gaussian spatial profile, the purple area is for the first Gaussian and the green area for the second Gaussian. z_{R1} and z_{R2} are the Rayleigh ranges for the first and second Gaussian profiles, respectively. Δ indicates the X-ray beam width or focal spot size, which varies along with z. The spatial fluence distribution is given in the xy plane, and a spatial profile as a function of x at $y = z = 0$ is shown in Fig. 2. The produced ions are collected by a detector with a slit length z_0 shown at the top of the figure.

$$z_{\text{R1}} = \frac{1}{2\ln(2)}\frac{\pi\Delta^2}{\lambda}. \tag{6}$$

The diameter of the cross section at $z = z_{\text{R1}}$ is $\sqrt{2}$ times larger than that at $z = 0$, namely, $\Delta_1(z = z_{\text{R1}}) = \sqrt{2}\Delta$. The beam geometry is depicted in Fig. 1. The Rayleigh range becomes larger as the photon energy increases and smaller as the focal size decreases.

We model the ensemble-averaged spatial fluence distribution transverse to the XFEL beam propagation direction (the z-axis in Fig. 1). In this work, we employ two types of spatial fluence profile: a single Gaussian or a double Gaussian spatial profile, which is consistent with the findings of Barty *et al.* (2009) who concluded that the spatial profile has a shape similar to a Gaussian function. The model allows us to reduce the number of calibration parameters. When we employ the observed yields of N charge states for calibration, the number of degrees of freedom is $(N - 1)$ after normalization. Hence, we can reconstruct typically fewer than N parameters, which we use to parametrize the fluence distribution. The ansatz of using a single Gaussian or a double Gaussian is a convenient parameterization of the fluence distribution function. It has been shown in a series of LCLS and SACLA experiments that these simple Gaussian profiles successfully produce theoretical volume-integrated ion yields in good agreement with experimental results (Young *et al.*, 2010; Doumy *et al.*, 2011; Rudek *et al.*, 2012, 2013, 2018; Fukuzawa *et al.*, 2013; Motomura *et al.*, 2013; Murphy *et al.*, 2014; Rudenko *et al.*, 2017).

The single Gaussian spatial profile (SGSP) is given by

$$F(\mathbf{r}; F_0) = \frac{\Delta^2}{\Delta_1^2(z)} F_0 \exp\left[-\pi a \frac{x^2 + y^2}{\Delta_1^2(z)}\right]. \tag{7}$$

The constant $a = 4\ln(2)/\pi$ is chosen so that $\Delta_1(z)$ becomes the FWHM. The beam size $\Delta_1(z)$ is given by equation (5). The quantity F_0 represents the peak fluence defined by number of photons per unit area. The total number of photons n in the xy-plane at an arbitrary value of z is given by

$$n = \int F(\mathbf{r}; F_0)\, dx\, dy = \frac{F_0\Delta^2}{a}. \tag{8}$$

Here we assume that the decrease of n due to photon absorptions by target atoms or molecules is negligible. Because the number of photon n is a constant, we can only determine either F_0 or Δ. Thus, the SGSP, equation (7), is characterized by only one of them. We use the experimentally determined focal spot size Δ^2 in this work, so F_0 is the parameter to be optimized. One may have the impression that an accurately measured focal area is a prerequisite to perform calibrations using *xcalib*. However, this is not the case. One can easily show that the focal area Δ^2 is factored out by changing integration variables, $x = x'\Delta$ and $y = y'\Delta$, in equation (1). It is then found that the ion yields $Y_{\text{theo}}^{(+q)}(\mathbf{P})$, equation (1), and correspondingly the normalization constant $N(\mathbf{P})$, equation (3), are proportional to the focal area. Hence, the dependency on the focal area for the calibrated CSDs, $y_{\text{theo}}^{(+q)}(\mathbf{P})$ in equation (2), is canceled out after the normalization.

Next, we address the transmission of the X-ray optics. The energy delivered to the focal spot is given by the product of the pulse energy E as an input parameter and the transmission T, namely TE, which is equivalent to the quantity of $n\omega$, where the X-ray photon energy ω is an input parameter. Here, we assume that the pulse energy has no fluctuation, which could be realized in experiment by choosing a narrow interval of pulse energies. Equating them, the transmission T is given by

$$T = \frac{n\omega}{E}. \tag{9}$$

Substituting equation (8) into equation (9), the transmission T for the SGSP is given by

$$T = \frac{\Delta^2\omega}{aE} F_0. \tag{10}$$

If Δ is not known experimentally, then we can only determine the ratio

$$\frac{T}{\Delta^2} = \frac{\omega F_0}{aE}. \tag{11}$$

The length over which ions are collected along the X-ray beam is often, but not always, determined by a slit aperture in the spectrometer (see $\pm z_0/2$ marked at the top of Fig. 1). When the Rayleigh range is much wider than the ion detector slit size or the molecular beam size, the z-dependence of the X-ray beam width is no longer relevant. If this is the case, the volume integration in three dimensions in equation (1) can be approximated to that in two dimensions in the xy-plane, assuming that $\Delta_1(z) = \Delta$ [see equation (22) for this case].

2.3. Double Gaussian spatial profile

Another spatial fluence profile employed in this work is a double Gaussian spatial profile (DGSP) consisting of a narrow, high main peak and a wide, low-fluence tail. This profile consists of two Gaussian profiles. The first Gaussian profile given by equation (7) is characterized by the peak fluence F_0. In addition to the peak fluence F_0, the second Gaussian profile is characterized by two supplementary parameters: a fluence ratio f_r and a width ratio w_r between the first and second Gaussian profiles at $z = 0$. The peak fluence and the beam size of the second Gaussian profile at $z = 0$ are given by $f_r F_0$ and $w_r\Delta$, respectively. Then the spatial profile is given by

$$F(\mathbf{r}; \mathbf{P}) = \frac{\Delta^2}{\Delta_1^2(z)} F_0 \exp\left[-\pi a \frac{x^2 + y^2}{\Delta_1^2(z)}\right] \\ + \frac{(w_r\Delta)^2}{\Delta_2^2(z)} f_r F_0 \exp\left[-\pi a \frac{x^2 + y^2}{\Delta_2^2(z)}\right], \tag{12}$$

where the beam size $\Delta_2(z)$ of the second Gaussian profile is given by

$$\Delta_2(z) = w_r\Delta\left[1 + \left(\frac{z}{z_{\text{R2}}}\right)^2\right]^{1/2}. \tag{13}$$

The quantity z_{R2} represents the Rayleigh range of the second Gaussian profile given by

$$z_{R2} = \frac{1}{2\ln(2)} \frac{\pi(w_r\Delta)^2}{\lambda} = w_r^2 z_{R1}. \quad (14)$$

The first and second Gaussian profiles in the DGSP enable modeling an XFEL beam with a narrow intense hot spot on a broad tail, as demonstrated in previous work (Murphy *et al.*, 2014). It was also reported that an experimental spatial profile of an XFEL pulse may show one ideal peak as well as several additional peaks due to aberrations of the focus (Nagler *et al.*, 2017), which could be modeled as a broad, low-fluence tail. We thus limit the range of the values of w_r and f_r such that $w_r > 1$ and $f_r < 1$ so that the second Gaussian profile corresponds to a wide and low fluence tail. In case the optimization obtains a solution such that $w_r < 1$ and $f_r > 1$, the solution can be converted into an equivalent solution so that the first Gaussian becomes narrower and higher (see Appendix B).

Next, we derive an expression for the transmission for the DGSP. Summing up the maximum amplitudes of the first and second Gaussian profiles, we introduce the global peak fluence F_G given by

$$F_G = (1 + f_r) F_0. \quad (15)$$

Using the global peak fluence F_G as well as w_r and f_r, we characterize the DGSP, equation (12), by the set of three parameters **P** given by

$$\mathbf{P} = (F_G, w_r, f_r). \quad (16)$$

Let n_1 and n_2 be the number of photons in the first and second Gaussian profiles in equation (12), respectively. Then these are given by

$$n_1 = \frac{\Delta^2}{a} \frac{F_G}{1 + f_r}, \quad (17a)$$

$$n_2 = f_r w_r^2 n_1, \quad (17b)$$

The total number of photons n is

$$n = n_1 + n_2 = \frac{\Delta^2}{a} \frac{1 + w_r^2 f_r}{1 + f_r} F_G. \quad (17c)$$

Substituting equation (17c) into equation (9), we obtain a formula for the transmission for the DGSP,

$$T = \frac{\omega\Delta^2}{aE} \frac{1 + w_r^2 f_r}{1 + f_r} F_G. \quad (17d)$$

Again, if Δ^2 is not accurately known, we can only find the ratio

$$\frac{T}{\Delta^2} = \frac{\omega}{aE} \frac{1 + w_r^2 f_r}{1 + f_r} F_G. \quad (18)$$

Because the width ratio w_r of the second Gaussian profile changes at each step of the optimization procedure, we adopted a specific grid scheme to accurately calculate the volume integration (see Appendix C).

A typical DGSP at $y = 0$ and $z = 0$ is shown in Fig. 2. It can be seen in the figure that the first and second Gaussian profiles form the narrow and high main peak and the wide and low fluence background, respectively. The set of three parameters **P** is given by $F_G = 3.85 \times 10^{12}$ photons μm^{-2}, $w_r = 3.04$, and $f_r = 0.234$. The set is obtained by Ar calibration at 6.5 keV and
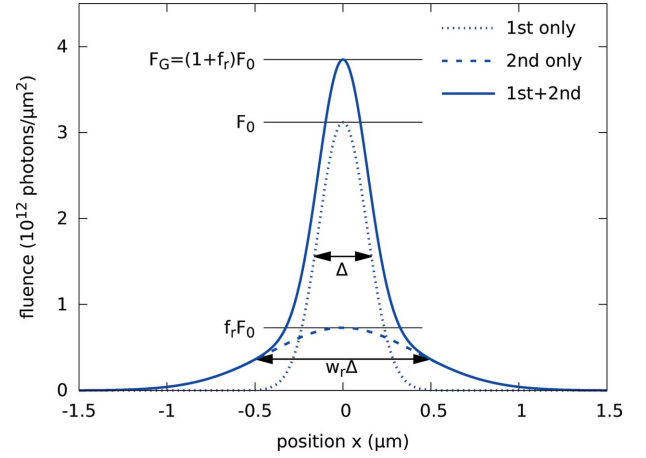

Figure 2
The DGSP, equation (12), as a function of x at $y = z = 0$. The first and second Gaussian profiles are represented by the dotted and dashed lines, respectively. The FWHM values Δ and $w_r\Delta$ and the peak fluence values F_0 and $f_r F_0$, and the global peak fluence F_G are shown. The first Gaussian profile forms the narrow and high main peak, and the second Gaussian profile does the wide and low background. The values of parameters are given by equation (25). These values are obtained in the third example in Section 3.3 calibrating the Ar CSD at 6.5 keV and 4.3 mJ.

4.3 mJ demonstrated in Section 3.3 [see equation (25)]. The beam sizes of the first and second Gaussian profiles, $\Delta_1(z)$ in equation (5) and $\Delta_2(z)$ in equation (13), are shown by purple and green in Fig. 1. The Rayleigh ranges equation (6) and equation (14) and a detector of slit length z_0 are also shown in this figure.

2.4. Ion yields before and after volume integration

Next, we show the calculated yields of Ar ions before and after the volume integration, equation (2), using the DGSP, equation (12), for illustrative purpose. We use *xatom* (Son *et al.*, 2018) to simulate ionization dynamics of isolated atoms interacting with intense XFEL pulses and calculate CSDs for a given fluence (Son *et al.*, 2011; Jurek *et al.*, 2016b). The set of parameters **P** for the DGSP is the same as that used for Figs. 1 and 2.

The photon energy used in the simulations is 6.5 keV, and the pulse duration is 30 fs FWHM. The calculated CSDs are shown in Fig. 3 as a function of the fluence. In Fig. 3(a), the yields of Ar ions $\mathcal{Y}_{theo}^{(+q)}(F)$, the integrand in equation (1), as a function of fluence are shown for several charge states. It is seen in the figure that, except for +18 which is the highest charge state of Ar, the ion yields $\mathcal{Y}_{theo}^{(+q)}(F)$ reach a maximum at a certain fluence value and then they start decreasing due to target depletion, *i.e.* they are saturated. Fig. 3(b) shows the volume-integrated absolute ion yields defined in equation (1) as a function of global peak fluence F_G. A vertical arrow indicates the global fluence value of $F_G = 3.85 \times 10^{12}$ photons μm^{-2} in the set of **P**, equation (25). In Fig. 3(b), the volume-integrated ion yields after reaching saturation do not decrease but rather become flattened as F_G increases, because of low-fluence contributions to their ion yields.
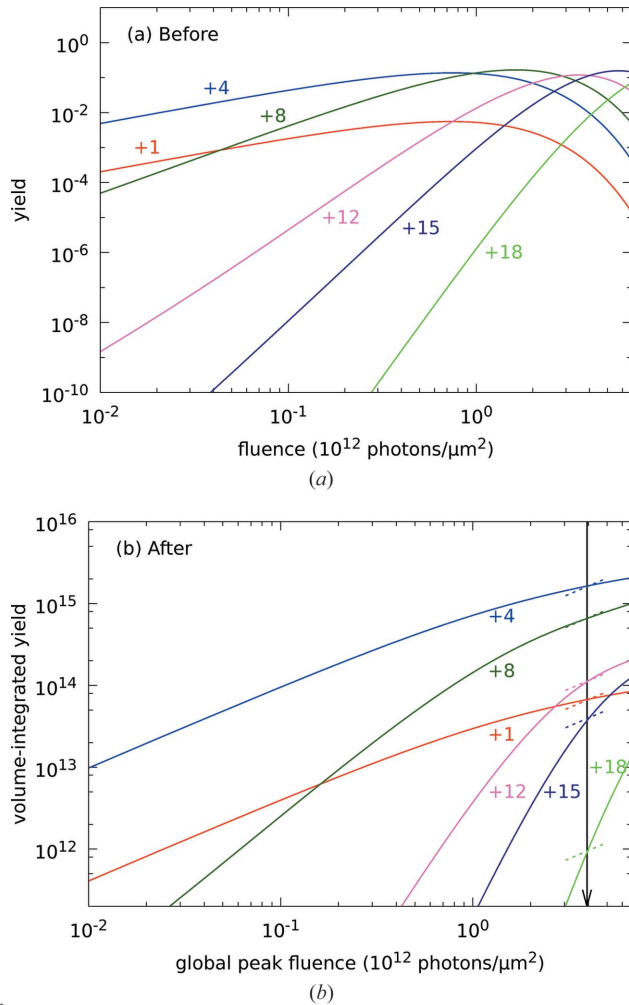
Figure 3
(a) Ion yields for several Ar charge states at 6.5 keV as a function of fluence before the volume integration, equation (1). Note that we employ logarithmic scales for both axes. (b) The same ion yields but after the volume integration using the calibrated parameters in equation (25). The vertical arrow indicates $F_G = 3.85 \times 10^{12}$ photons μm^{-2}, which is the calibrated global peak fluence. The dashed lines representing the slope of unity are marked at the vertical arrow to guide the eye.

We utilize the slope of ion yields $s(F_G)$ as a function of global peak fluence F_G to measure the degree of saturation,

$$s(F_G) = \frac{\ln Y_{theo}^{(+q)}(F_G + \Delta F_G) - \ln Y_{theo}^{(+q)}(F_G)}{\ln(F_G + \Delta F_G) - \ln F_G}, \quad (19)$$

where the quantity $Y_{theo}^{(+q)}(F_G)$ represent the absolute ion yield given by equation (1). The other parameters in $\mathbf{P}$ are not shown for simplicity. Note that we define the slope using a double logarithmic scale so that in the low fluence limit the value corresponds to the number of required photons to produce a certain charge state (see Appendix D). We regard a charge state as saturated at a given fluence F_G if the slope of the associated ion yield is lowered by one or more in comparison with the low-fluence limit. If the slope becomes even less than unity, especially for low charge states, the ion yield might be less sensitive to fluence values around the saturation point, and then it could introduce an unnecessary

ambiguity into the optimization of the F_G value. Thus, it is conceivable to select charge states based on the slope of their volume-integrated ion yields. The pieces of dashed lines in Fig. 3(b) are marked to guide the eye, representing the slope of unity to be compared with the calculated slope of ion yields at the calibrated fluence F_G. It is found that the ion yields for $q \leq +9$ are saturated and their slope is less than unity (+1, +4 and +8 are shown in the figure) before reaching $F_G = 3.85 \times 10^{12}$ photons μm^{-2}. We attempt to remove these charge states in the optimization, which will be demonstrated in Section 3.3. Note that the saturation effect needs to be carefully taken into consideration, especially when calibration is performed at very high fluences. If the majority of charge states are saturated and their ion yields become insensitive to different fluence values, the result of the calibration procedure becomes unreliable.

2.5. Optimization

Here we explain the automated calibration procedure for spatial fluence profiles employing the volume integration scheme established in the previous subsections. The aim of calibration is to find out the best parameter set $\mathbf{P}$ that minimizes the difference between the volume-integrated theoretical result and the experimental result. In previous works (Rudek et al., 2012; Fukuzawa et al., 2013; Murphy et al., 2014; Rudenko et al., 2017), the calibrations were conducted by manually exploring the parameter space. The manual exploration is inefficient because the direction to a solution from an initial guess is not known in general. Hence, the manual procedure becomes impractical when the dimension of $\mathbf{P}$ grows or a number of calibrations must be performed. In xcalib the calibration is performed by reinforcement learning (Raschka, 2015) combined with optimization modules in Python that automatize exploring the parameter space. The concept of the reinforcement learning is sketched in Fig. 4. In this figure, the Agent gives the Environment Action, then the Environment returns the Reward. The Agent repeats the procedure to find the Action that minimizes or maximizes the Reward in a trial-and-error approach. In xcalib, the Agent
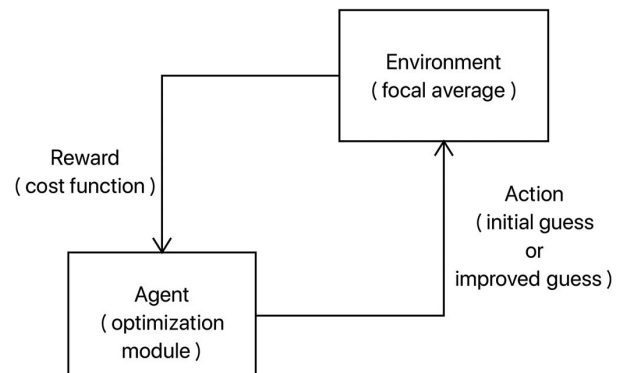

Figure 4
The concept of the reinforce learning in machine learning (Raschka, 2015).

is built up employing optimization modules in Python. The Action and the Environment correspond to an initial (improved) guess and the calculation of the volume integration using the guess. The Reward in Fig. 4 corresponds to a cost function $\delta(\mathbf{P})$ defined by the difference between the volume-integrated theoretical result and the experimental data. Since the next direction in the parameter space from a current guess is determined by established mathematical algorithms, the reinforcement learning approach is much more efficient than the manual procedure. Although we calibrate at most three parameters in this work, using *xcalib* will be critical when calibrating the fluence profiles in high-dimensional parameter spaces.

In this work, we define the cost function by the sum of quadratic logarithmic differences between the theoretical $y_{\text{theo}}^{(+q)}(\mathbf{P})$ and the experimental ion yield $y_{\text{expt}}^{(+q)}$,

$$\delta(\mathbf{P}) = \sum_{\text{selected } q} \left[\log \frac{y_{\text{theo}}^{(+q)}(\mathbf{P})}{y_{\text{expt}}^{(+q)}} \right]^2. \qquad (20)$$

The experimental ion yields are normalized in the same manner as the theoretical result [see equations (2) and (3)]. In equation (20), the summation runs over selected charge states. The set of selected charge states is formed by removing the charge states whose ion yields are not accurate enough in theory and/or experiment. At the end of each iteration in the optimization, the ion yield is normalized according to equations (2) and (3). Although some charge states are excluded in the sum of equation (20), an optimized result depends on them through the normalization, equation (3). We keep them in the normalization condition for compatibility with previous works (Rudek *et al.*, 2012; Fukuzawa *et al.*, 2013; Murphy *et al.*, 2014). The advantage of using a logarithmic function is that the relative weights of the ion yields over the wide range of charge states become comparable in amplitude. Therefore, tiny ion yields for high charge states and large ion yields for other charge states can be treated on equal footing in the optimization.

3. Results

In this section, we demonstrate the *xcalib* toolkit for three examples. The first example is the Ar calibration at 5.5 keV for the xenon (Xe) experiment performed at SACLA (Fukuzawa *et al.*, 2013). We confirm that the results produced by *xcalib* are consistent with those of Fukuzawa *et al.* (2013). The second example is the Ar calibration at 805 eV for the C_{60} experiment at LCLS (Murphy *et al.*, 2014). We demonstrate that the second Gaussian profile in the DGSP, equation (12), is essential to model the low fluence tail of an XFEL pulse, as previously shown by Murphy *et al.* (2014) and Nagler *et al.* (2017). In the third example, we study the effect of an attenuator on the spatial fluence distribution by performing an Ar calibration at 6.5 keV for a recent experiment (Rudek *et al.*, 2018).

3.1. Ar calibration with a single Gaussian spatial profile

In the SACLA experiment (Fukuzawa *et al.*, 2013), the photon energy and the pulse energy were 5.5 keV and 239 µJ, respectively. The pulse duration used in calculations was 30 fs FWHM. The nominal focal spot area was $\Delta^2 = 1$ µm × 1 µm. The molecular beam size was about 2 mm. The peak fluence F_0 of equation (7) was manually calibrated to reproduce the experimental ratio

$$\frac{y_{\text{expt}}^{(+8)}(F_0) + y_{\text{expt}}^{(+9)}(F_0)}{y_{\text{expt}}^{(+3)}(F_0) + y_{\text{expt}}^{(+4)}(F_0)}, \qquad (21)$$

where the numerator represents the sum of ion yields of Ar^{8+} and Ar^{9+} produced by two-photon absorption, and the denominator is that of Ar^{3+} and Ar^{4+} produced by one-photon absorption. The ratio thus gives us the relative contribution between two-photon and one-photon absorption processes. However, *xcalib* does not require such additional physical considerations on the number of required photons to produce a certain charge state in order to define a cost function.

The Rayleigh range calculated via equation (6) with the given photon energy and the focal spot area is $z_{\text{R1}} = 10$ mm, which is five times larger than the molecular beam size. In such a situation, it can be assumed that the atoms were subject to the same fluence value in the z-direction. Under this consideration, the focal size in $\Delta_1(z)$, equation (7), is approximated by $\Delta_1(z) \simeq \Delta$. The SGSP, equation (7), is then simplified to

$$F(x, y; F_0) = F_0 \exp\left(-\pi a \frac{x^2 + y^2}{\Delta^2}\right). \qquad (22)$$

In the following, we revisit the calibration for the SACLA experiment with *xcalib*, employing the two-dimensional (2D) version of the SGSP, equation (22).

The experimental Ar CSD is shown by circles (black) in Fig. 5(*a*). The set of all charge states for equations (2), (3) and (4) consists of the charge states from +1 to +10 observed in the experiment. The set of selected charge states to calculate the cost function, equation (20), is obtained by removing the charges states of +6 and +7 from the set of all charge states whose ion yields are underestimated in theory (Fukuzawa *et al.*, 2013; Murphy *et al.*, 2014; Rudenko *et al.*, 2017). To obtain the optimized value of F_0, 30 *xcalib* runs were submitted with the initial guesses of the peak fluence F_0 equidistantly distributed over $F_0(10^{12}$ photons µm$^{-2}) \in [0.036, 0.36]$. The optimized peak fluence value is $F_0 = 0.056 \times 10^{12}$ photons µm^{-2} corresponding to the lowest cost function value, equation (20). The transmission value of 23.2% calculated using equation (10) is consistent with 22.3% obtained by Fukuzawa *et al.* (2013) by performing a three-dimensional volume integration. Therefore the integration with respect to the z-axis does not affect the result in this case. The 2D volume-integrated Ar CSD is shown by triangles (red) in Fig. 5(*a*). Comparing with the experimental result, it is seen that the ion yield of +5 is slightly overestimated, while those of +6 and +7, excluded in the cost function, are significantly underestimated. A possible reason for this discrepancy is the neglect of higher-order many-electron corrections in our theoretical model.
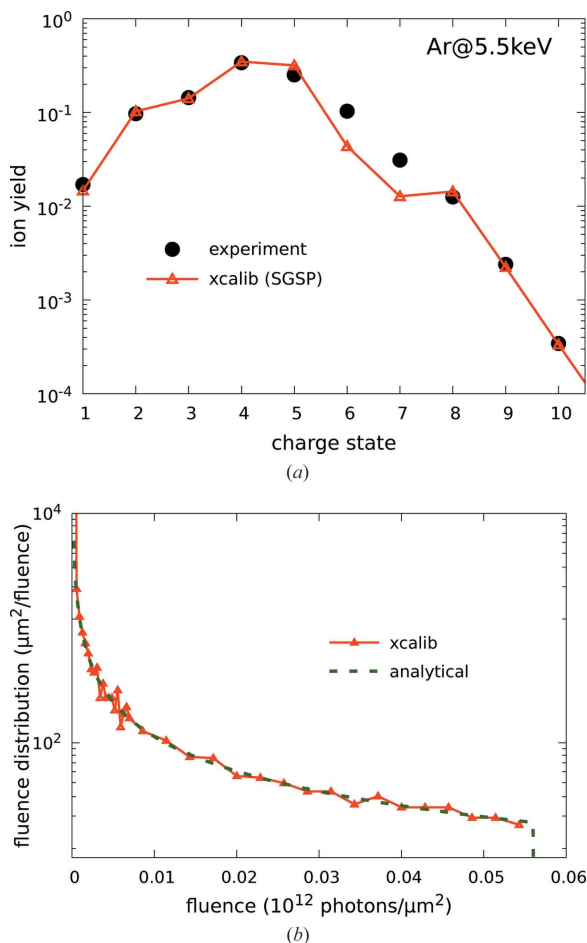
Figure 5
(a) Ar CSD at 5.5 keV by Fukuzawa et al. (2013) is shown. The black line with dots shows the experimental results. The red line with triangles shows the theoretical results obtained by *xcalib* with the SGSP, equation (22). (b) Numerical and analytical FDFs for the SGSP, equation (22). The analytical formula is given by equation (42).

We further examine *xcalib* by comparing numerical and analytical fluence distribution functions (FDFs). The FDF is defined by the area per unit fluence occupied by a certain fluence value in a given spatial profile. For instance, the FDF vanishes at the peak fluence and goes to infinity at zero fluence. A numerical FDF is obtained by making a histogram of fluence values multiplied by a constant. The constant is given by a unit volume made by spatial grid points divided by the fluence bin size. An analytical formula is obtained for the case of the SGSP, equation (22) (see Appendix E). In Fig. 5(b), the numerical and analytical FDFs are shown by solid (red) and dashed lines (green). It is seen that both of them agree very well. The FDF vanishes above the peak fluence $F_0 = 0.056 \times 10^{12}$ photons μm^{-2}. The noisy behavior in the numerical result for fluence values comes from low statistics.

3.2. Ar calibration with a double Gaussian spatial profile

Here we revisit the Ar calibration at 805 eV performed for the C_{60} experiment at LCLS (Murphy et al., 2014). The pulse energy was 1.15 mJ and the nominal focal area was $\Delta^2 = 1.38 \, \mu m \times 1.38 \, \mu m$. The nominal pulse duration was 30 fs

FWHM. The slit size was 1.6 mm (Osipov, 2013). The experimental Ar CSD is shown by black dots in Fig. 6(a). The charge states of +5 and +10 are not shown because of experimental uncertainty. Thus, when evaluating equations (3) and (4), the set of charge states consists of +1 to +14 without +5 and +10. The set of selected charge states to calculate the cost function, equation (20), is constructed without the charge states of +5, +6, +7 and +10.

The calibration was first attempted employing a SGSP, equation (7), taking the Rayleigh range into consideration. The photon energy is smaller than that used in Section 3.1, so the Rayleigh range is smaller here. The calculated value is $z_{R1} = 2.8$ mm, which is now comparable with the slit size. Therefore, we use the three-dimensional integration in equation (2) with the three-dimensional SGSP in equation (7), in order to obtain the CSDs. The optimized peak fluence value is $F_0 = 0.279 \times 10^{12}$ photons μm^{-2} and the calculated CSD is shown using triangles (red) in Fig. 6(a). The calibration clearly fails to reproduce the experimental result. The ion yields for low charge states are underestimated, whereas those for middle to high charge states are overestimated, as was also
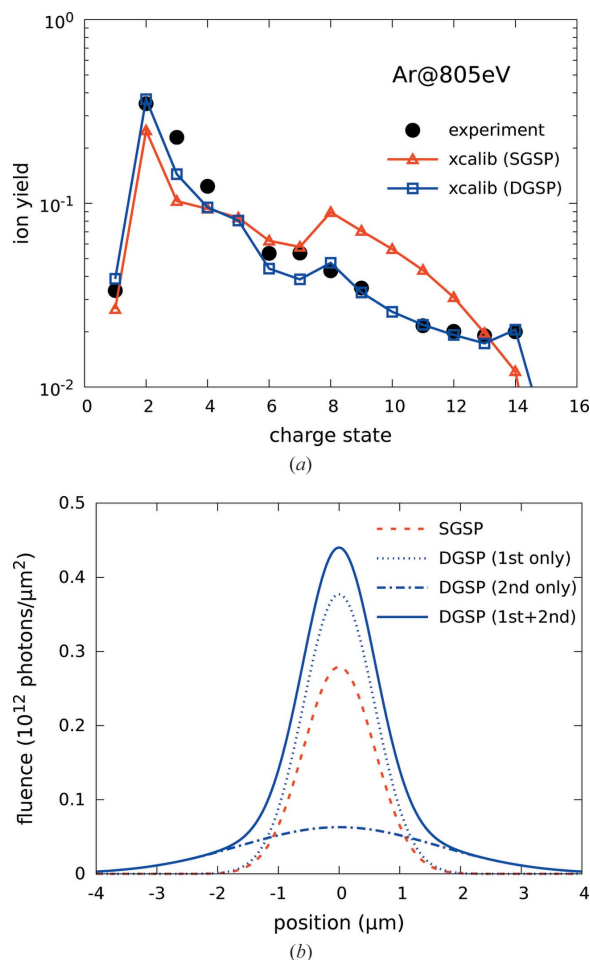

Figure 6
(a) The xperimental Ar CSD at 805 eV obtained by Murphy et al. (2014) is shown by black dots. Triangles indicate the volume-integrated results with calibrated SGSP in equation (7), and squares are those with calibrated DGSP in equation (12) in this work. (b) Calibrated SGSP, equation (7), and DGSP, equation (12), at $y = 0$.

found by Murphy *et al.* (2014). This failure implies that the beam profile is more complex than a SGSP, so a DGSP was introduced by Murphy *et al.* (2014) to overcome this problem.

Here we also employ a DGSP, equation (12). Twenty-seven initial guesses, uniformly distributed in the parameter range

$$F_G \ (10^{12} \ \text{photons} \ \mu\text{m}^{-2}) \in [0.357, 1.79],$$
$$w_r \in [1.1, 3.1], \qquad (23)$$
$$f_r \in [0.1, 0.5],$$

were employed to perform the optimization. The solution giving the lowest local cost function value for the set of selected charge states is given by $\mathbf{P} = (F_G, w_r, f_r) = (0.44 \times 10^{12} \ \text{photons} \ \mu\text{m}^{-2}, 2.77, 0.167)$. The result is given by squares (blue) in Fig. 6(*a*). The agreement between the result of *xcalib* and the experimental result is now much better. The low fluence wing supported by the second Gaussian profile in equation (12) enhances the ion yields of low charge states. The DGSP, equation (12), with the calibrated parameter set at $y = 0$ is depicted in Fig. 6(*b*) together with the SGSP, equation (7).

3.3. Ar calibration with attenuated beams

In this demonstration, we calibrate the three data sets of Ar CSDs at 6.5 keV taken for a recent experiment on Xe (Rudek *et al.*, 2018). The nominal pulse duration was 30 fs FWHM. The first data set consists of Ar CSDs measured without an attenuator. The pulse energy fluctuates from shot to shot, so the Ar CSD data are binned according to the pulse energies of 4.5, 4.3 and 4.1 mJ. A silicon attenuator was used to reduce the pulse energies to 58% of the full-power beam, providing the pulse energies of 2.61, 2.49 and 2.38 mJ (the second data set), and 20% corresponding to 0.90, 0.86 and 0.82 mJ (the third data set). We calibrate these three data sets (100%, 58% and 20%), as listed in Table 1, to study the effect of the attenuator on the spatial fluence profile. It has been believed that an attenuator would not change the spatial profile of XFEL pulses, but no comprehensive studies have been reported so far on this subject.

The nominal focal spot area in the experiment was estimated as $\Delta^2 = 0.35 \ \mu\text{m} \times 0.3 \ \mu\text{m}$ with an elliptic focal shape. In the numerical method described in Section 2, a circularly shaped focal area is assumed. We numerically confirmed that the volume-integrated Ar CSD and the FDF at 4.3 mJ do not change noticeably when an elliptic focal shape is explicitly used. Hence, we keep using a circularly shaped focal spot, *i.e.* $\Delta^2 = 0.324 \ \mu\text{m} \times 0.324 \ \mu\text{m}$, in the following. The focal size is much smaller than that used in Section 3.1, so the Rayleigh range is smaller. It is calculated as $z_{R1} = 1.25$ mm, which is quite comparable with the slit size used in experiment ($z_0 = 1.0$ mm) (Rudek *et al.*, 2018). Therefore, we start with three-dimensional (3D) volume integration, but two-dimensional (2D) volume integration will also be tested later on. Here we assume a DGSP, because a DGSP is more general than a SGSP. We explain the numerical procedure for the pulse energy of 4.3 mJ in the following. The same procedure is applied for the other pulse energies.

Table 1
Calibrated DGSPs, equation (12), for non-attenuated (100%) and attenuated (58% and 20%) beams at 6.5 keV. These parameters were extracted from the experimental Ar CSDs of Rudek *et al.* (2018).

	E (mJ)	F_G (photons μm^{-2})	w_r	f_r	$w_r^2 f_r$	T/Δ^2 (μm^{-2})
Non-attenuated (100%)	4.5	4.01×10^{12}	2.91	0.266	2.25	2.70
	4.3	3.93×10^{12}	3.10	0.281	2.70	3.12
	4.1	3.90×10^{12}	3.29	0.275	2.98	3.50
Attenuated to 58% intensity	2.61	2.95×10^{12}	3.69	0.160	2.18	3.65
	2.49	3.02×10^{12}	3.98	0.168	2.66	4.49
	2.38	2.95×10^{12}	4.16	0.171	2.96	4.95
Attenuated to 20% intensity	0.90	1.27×10^{12}	5.37	0.135	3.89	7.18
	0.86	1.44×10^{12}	6.02	0.172	6.23	12.19
	0.82	1.39×10^{12}	6.30	0.175	6.95	13.53

In principle, the set of all charge states for Ar ranges from +1 to +18. In the recent experiment with ultra-intense hard-X-ray pulses, the charge state of +18, a bare nucleus of Ar, was observed (Rudenko *et al.*, 2017). The set of selected charge states is obtained by removing +6 and +7, whose ion yields are inaccurate in theory. Twenty-seven initial guesses were uniformly distributed in the parameter range given by

$$F_G \ (10^{12} \ \text{photons} \ \mu\text{m}^{-2}) \in [0.357, 1.79],$$
$$w_r \in [1.1, 3.1], \qquad (24)$$
$$f_r \in [0.1, 0.5].$$

The solution corresponding to the lowest cost function value, equation (20), is given by

$$\mathbf{P}_{1st}^{3D} = (F_G, w_r, f_r)$$
$$= (3.85 \times 10^{12} \ \text{photons} \ \mu\text{m}^{-2}, 3.04, 0.234). \quad (25)$$

The beam geometry of the calibrated DGSP, equation (12), for the set of parameters $\mathbf{P}_{1st}^{3D}$ was shown in Fig. 1. In Fig. 3(*b*), the volume-integrated ion yields, equation (2), for several charge states for the set $\mathbf{P}_{1st}^{3D}$ were shown. The ion yields were calculated by changing F_G values with w_r and f_r values being fixed. As discussed earlier, because of the saturation effect, the ion yields of low charge states are less sensitive to the change of the global fluence value around $F_G = 3.85 \times 10^{12}$ photons μm^{-2} in optimization, which might give rise to an ambiguity in the determination of the global fluence value. Specifically, the ion yields of the charge states from +1 to +9 have a slope that is less than unity. Therefore, a second optimization was performed removing these charge states from the set of charge states selected for the optimization procedure. Then we obtained the optimized set of parameters given by

$$\mathbf{P}_{2nd}^{3D} = (F_G, w_r, f_r)$$
$$= (3.93 \times 10^{12} \ \text{photons} \ \mu\text{m}^{-2}, 3.10, 0.282). \quad (26)$$

These two different optimizations using different sets of charge states give rather similar parameters. Also both methods, $\mathbf{P}_{1st}^{3D}$ (red triangles) and $\mathbf{P}_{2nd}^{3D}$ (blue squares), provide almost the same Ar CSDs, as illustrated in Fig. 7. In the following applications, we keep the charge selection procedure

using the slope of volume-integrated ion yields. The transmission calculated using equation (17d) with $\mathbf{P}_{2nd}^{3D}$ is 32.8%.

Here we attempt to compare 2D and 3D volume integrations to perform the calibration. The result obtained via 2D volume integration after the charge selection is given by

$$\mathbf{P}_{2nd}^{2D} = \left(F_G, w_r, f_r\right)$$
$$= (3.93 \times 10^{12}\ \text{photons}\,\mu\text{m}^{-2}, 3.10, 0.281), \quad (27)$$

which practically coincides with the result of equation (26). The transmission is calculated using equation (17d) as 32.6%, and the volume-integrated CSD with this calibrated parameter set looks quite similar to those in Fig. 7. The almost identical results of $\mathbf{P}_{2nd}^{3D}$ and $\mathbf{P}_{2nd}^{2D}$ indicate that the integration for the z-direction does not affect the final Ar CSDs. We confirm that this is also true for other pulse energies. Therefore, we will show further results for $\mathbf{P}_{2nd}^{2D}$ only.

In Table 1, we list all numerical results for three data sets (nine different pulse energies). In each row, the global peak fluence F_G, the width ratio w_r, and the fluence ratio f_r are given. The energy ratio (or the ratio of the total numbers of photons) between the first and second Gaussians is calculated using equation (17b), $n_2/n_1 = w_r^2 f_r$. The ratios T/Δ^2 calculated using equation (18) are listed in the last column of the table, assuming that the input pulse energy E is known. When inspecting Table 1, we find interesting trends. Our calibration suggests that F_G decreases more slowly than E. Note that if the shape of the spatial fluence distribution were unchanged, F_G would be exactly proportional to E. Our calibrated parameters imply that the shape of the spatial fluence distribution of the X-ray beam appears to be affected by the attenuation process. This fact becomes even more evident when we consider the width ratio w_r, which increases monotonically with decreasing pulse energy. In other words, the attenuation process causes the second Gaussian to become wider relative to the first Gaussian in the DGSP (or the first Gaussian becomes narrower in comparison with the second Gaussian), with
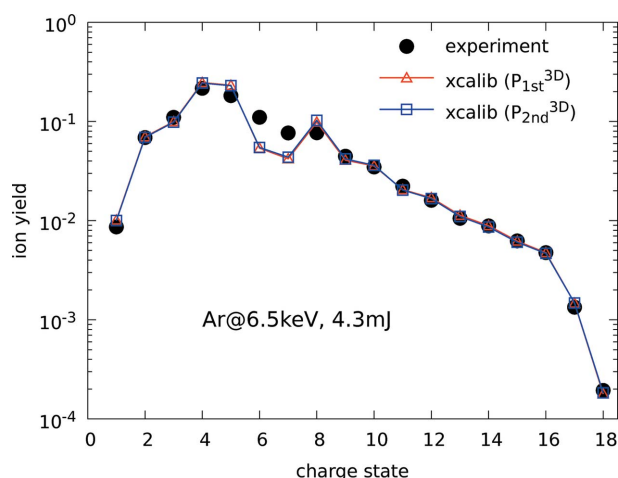
increasing attenuation. Similarly, we observe that there is a tendency for the fluence ratio f_r to decrease with stronger attenuation. The energy ratio $w_r^2 f_r$ shows that the 100% and 58% cases are similar, while the ratio becomes larger in the 20% case. It might indicate that the first Gaussian is more attenuated than the second Gaussian in the 20% case.

The most pronounced effect that we observe in Table 1 is the ratio of the transmission and the focal area of the first Gaussian, T/Δ^2: decreasing the pulse energy by a factor of five increases this ratio by a factor of five. Also note that the ratio changes by almost a factor of two between the two bins of the same data set in the 20% attenuation case. Again, it is supposed to be the same ratio in the ideal case where the attenuator would not change the spatial fluence distribution. Either the transmission through the X-ray focusing optics increases or the focal area of the first Gaussian decreases as the pulse energy decreases, or both change when attenuating the pulse energy. The reduction of the focal area of the first Gaussian would seem rather surprising because any wavefront distortion induced by the inserted foil would cause the focal spot to become larger, but not smaller. Thus, we speculate as to the possibility of creating a very irregular focal shape with the hot spot of the beam, which may not be captured by a simple Gaussian profile. The increase of the beamline transmission would likely mean that at higher fluences the focusing or transport mirrors are heated up and lose part of their reflectivity. Even though this effect was found to be small (Ryutov *et al.*, 2009), it may not be completely excluded in our case where high pulse energy (up to 4.5 mJ) was employed. Since T/Δ^2 is inversely proportional to the pulse energy in equations (11) and (18), this conclusion requires the assumption that the experimental attenuation coefficients are correct. Our calibration cannot determine T and Δ independently of each other, and both of them are not always well known for different experimental configurations, which currently prevents a detailed understanding of the attenuation behavior. Nevertheless, our results indicate that the attenuation process does not only reduce the pulse energy, but also it might influence the spatial fluence profile. In view of these observations, we consider it advisable to calibrate the spatial fluence profile for each and every experimental condition when using attenuated beams.

Figure 7
Ar CSDs at 6.5 keV and 4.3 mJ. Circles (black) represent experimental result. Triangles (red) and squares (blue) represent the solutions $\mathbf{P}_{1st}^{3D}$, equation (25), and $\mathbf{P}_{2nd}^{3D}$, equation (26), respectively, obtained by *xcalib*. These are obtained by including or excluding some charge states whose volume-integrated ion yields are saturated.

4. Conclusion

In this work, we have developed the *xcalib* toolkit to calibrate the spatial fluence distribution of an XFEL pulse at its focal spot using the CSDs of light atoms. The calibration of the spatial fluence profile is essential to calculate volume-integrated CSDs and to make a quantitative comparison between theoretical and experimental results. We formulated the calibration procedure based on reinforcement learning (Raschka, 2015) by using optimization modules in Python. The automated calibration procedure in *xcalib* is more efficient than the procedure of manually exploring the parameter space of the spatial fluence profile performed in previous studies

(Rudek *et al.*, 2012, 2018; Fukuzawa *et al.*, 2013; Murphy *et al.*, 2014).

Using the *xcalib* toolkit, we revisited the Ar calibrations performed for previous experiments (Fukuzawa *et al.*, 2013; Murphy *et al.*, 2014; Rudek *et al.*, 2018). In the first demonstration, we revisited the Ar calibration at 5.5 keV employing the SGSP, which was performed for a Xe experiment at SACLA (Fukuzawa *et al.*, 2013). Our result could reproduce the Ar CSD shown by Fukuzawa *et al.* (2013). In that reference, the peak fluence was calibrated manually by comparing the yields of ions created by two-photon absorption to the yields of ions created by one-photon absorption. In contrast to this manual procedure, we employed the cost function defined by the logarithmic difference between the volume-integrated theoretical result and the experimental data. Using *xcalib*, it is possible to perform calibrations without introducing a measure involving such physical considerations. We also confirmed that the numerical FDF of the SGSP agrees with the analytical formula. In the second demonstration, we revisited the Ar calibration at 805 eV performed for the C_{60} experiment by Murphy *et al.* (2014). We confirmed that the low fluence tail modeled by the second Gaussian profile in the DGSP is necessary to reproduce the experimental Ar CSD as demonstrated by Murphy *et al.* (2014). In the third demonstration, we performed the Ar calibration at 6.5 keV for a recent Xe experiment (Rudek *et al.*, 2018) to study the effect of an attenuator on the functional form of the fluence spatial profile. We found that the attenuation process appears to cause a significant modification of the spatial fluence profiles. Therefore, when using attenuated beams, a beam profile calibration is advisable for each and every experimental condition.

Our development is essential to automatize the optimization procedure with flexibility, combining different optimization algorithms, fluence profiles, charge states, and a wide range of parameter space, which is far beyond manual procedures employed in the literature (Rudek *et al.*, 2012, 2018; Fukuzawa *et al.*, 2013; Murphy *et al.*, 2014). Moreover, *xcalib* has the capability of handling massive amounts of experimental data through automatized optimization procedure. The calibrated Ar CSDs in this work agree well with the experimental data except for several charge states as previously found (Fukuzawa *et al.*, 2013; Murphy *et al.*, 2014; Rudek *et al.*, 2018). One way to improve the calibration is improving the level of electronic structure theory being employed by including shake-off process and/or double Auger decay when computing the ionization dynamics with *xatom*. In the future, *xcalib* will be employed to a more complex problem such as calibrating pump and probe pulses. *xcalib* offers us a tool to calibrate such cases with high efficiency powered by automation.

APPENDIX *A*
Dependence of the charge state distribution on the pulse duration

Fig. 8 shows Ar CSDs as a function of pulse duration, calculated using *xatom* (Son *et al.*, 2018; Toyota *et al.*, 2017). The
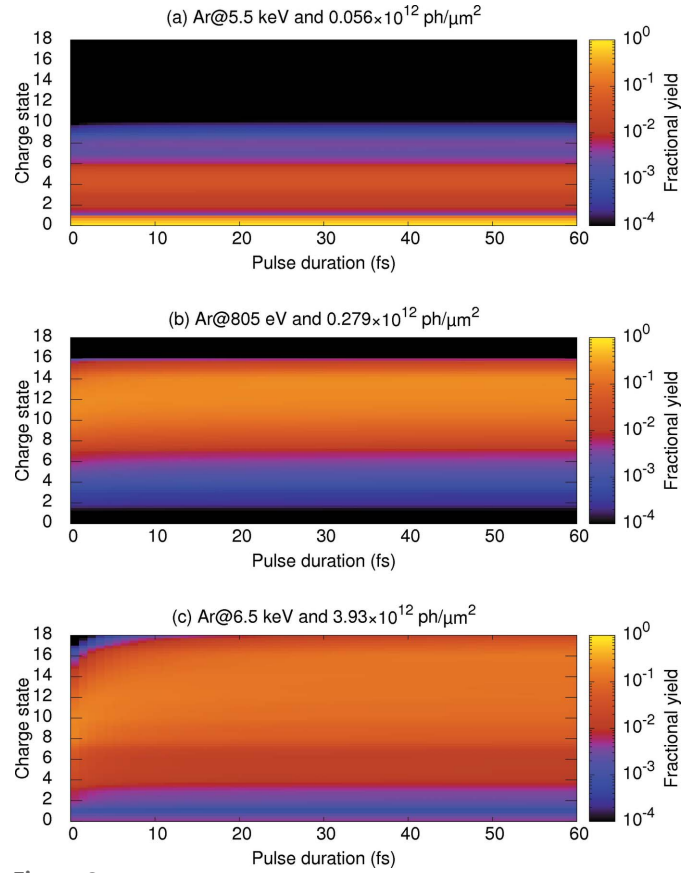
Figure 8
Calculated Ar CSDs as a function of pulse duration at (*a*) 5.5 keV, (*b*) 805 eV and (*c*) 6.5 keV.

color indicates the fractional yield of the charge state in a logarithmic scale. The photon energies (5.5 keV, 805 eV and 6.5 keV) were chosen from the three examples in the *Results* section, and the optimized peak fluence for each photon energy was employed. When the pulse duration is shorter than ∼10 fs, the calculated CSDs, particularly at 6.5 keV, are sensitive to the pulse duration. For a pulse duration longer than ∼10 fs, however, they show almost no dependence on the pulse duration for all cases considered.

APPENDIX *B*
Flipping the first and second Gaussian profiles

Let $\mathbf{P} = (F_0, f_r, w_r)$ be a set of parameters for the DGSP, equation (12), such that the second Gaussian profile is narrower and higher than the first Gaussian, namely, $f_r > 1$ and $w_r < 1$. In such a case, $\tilde{\mathbf{P}} = (f_r F_0, 1/f_r, 1/w_r)$ gives an equivalent solution having the same cost function value.

The inverses of f_r and w_r make the first Gaussian narrower and higher than the second Gaussian profile. The scaling of $F_0 \rightarrow f_r F_0$ keeps the global peak fluence F_G, equation (13), fixed. From equation (17*c*), the global peak fluence F_G for the parameter set $\tilde{\mathbf{P}}$ is given by

$$F_G(\tilde{\mathbf{P}}) = F_0 + f_r F_0 = \frac{a}{\Delta^2} \frac{1+f_r}{1+w_r^2 f_r} \frac{TE}{\omega} \times w_r^2. \qquad (28)$$

Hence, we also scale the transmission T as $T \to T/w_r^2$, so that $F_G(\tilde{\mathbf{P}})$ can satisfy the relation equation (17c) for the parameter set $\mathbf{P}$. For the solution $\tilde{\mathbf{P}}$, the absolute ion yield of a charge state q, equation (1), is given by

$$Y_{theo}^{(+q)}(\tilde{\mathbf{P}}) = \int \mathcal{Y}_{theo}^{(+q)}[F(\mathbf{r}; \tilde{\mathbf{P}})]\, d^3 r$$
$$= \int \mathcal{Y}_{theo}^{(+q)}[F(\mathbf{r}'; \tilde{\mathbf{P}})[\, d^3 r'$$
$$= \bar{Y}_{theo}^{(+q)}(\mathbf{P}). \qquad (29)$$

On the second line, the integration variables are changed to $\mathbf{r}' = (x', y', z') = (w_r x, w_r y, z/w_r^2)$. After changing the integration variables, the functional form of the fluence spatial profile coincides with that for the parameter set $\mathbf{P}$.

APPENDIX C
Grid scheme for the double Gaussian fluence distribution

During optimizations for the DGSP, the width of one Gaussian could become narrower or wider than the other, in contrast to the initial guess. Sufficiently fine grid points must be used to describe the narrower Gaussian to accurately calculate volume integrations. To handle such unpredictable situations, the grid points have to be dynamically changed at each optimization steps. The first step to define our grid points is to find out which Gaussian width is narrower or wider, namely

$$\Delta_< = \min(\Delta_1, \Delta_2), \qquad (30a)$$
$$\Delta_> = \max(\Delta_1, \Delta_2). \qquad (30b)$$

Using the narrower width, we define the grid spacing for the narrower Gaussian in the interval of $[-L_</2, L_</2]$,

$$l_< = \frac{L_<}{n_{grid}}, \qquad (30c)$$

where n_{grid} is the default number of grid points. The length of the interval $L_<$ is determined as small as possible such that the amplitudes of the narrower Gaussian at the borders can satisfy the condition

$$\exp\left(-\frac{L_<^2}{2\Delta_<^2}\right) \ll 1. \qquad (30d)$$

Using the quantity $l_<$, we define the grid points for both of two Gaussian profiles,

$$n = \left[\frac{L_>}{l_<}\right], \qquad (30e)$$

where the bracket is the operator which returns the nearest integer of the argument. In this work, we use $L_{<,>} = 7\Delta_{<,>}$, then the left-hand side of equation (30d) is $\sim 2.2 \times 10^{-3}$.

APPENDIX D
Slope value in the low fluence limit

Using $\ln(1 + x) \simeq x$ for $|x| \ll 1$, the slope of the logarithm of the absolute ion yield at fluence value F_0 may be computed via

$$\ln Y_{theo}^{(+q)}(F_0 + \Delta F_0) - \ln Y_{theo}^{(+q)}(F_0) \simeq \frac{1}{Y_{theo}^{(+q)}(F_0)} \frac{\partial Y_{theo}^{(+q)}(F_0)}{\partial F_0} \Delta F_0. \qquad (31)$$

Here the dependency of the absolute ion yield on other parameters is omitted for simplicity. In the same way,

$$\ln(F_0 + \Delta F_0) - \ln F_0 \simeq \frac{\Delta F_0}{F_0}. \qquad (32)$$

Using equations (31) and (32), the slope, equation (19), is given by

$$s(F_0) = \lim_{\Delta F_0 \to 0} \frac{\ln Y_{theo}^{(+q)}(F_0 + \Delta F_0) - \ln Y_{theo}^{(+q)}(F_0)}{\ln(F_0 + \Delta F_0) - \ln F_0}$$
$$= F_0 \frac{\partial}{\partial F_0} \ln Y_{theo}^{(+q)}(F_0). \qquad (33)$$

Substituting the definition of the absolute yield, equation (1),

$$s(F_0) = F_0 \left(\int \frac{\partial \mathcal{Y}_{theo}^{(+q)}(F_0)}{\partial F} \frac{\partial F}{\partial F_0}\, d^3 r\right) \bigg/ \int \mathcal{Y}_{theo}^{(+q)}(F_0)\, d^3 r. \qquad (34)$$

We assume that the spatial fluence profile may be written as

$$F(\mathbf{r}; \mathbf{P}) = F_0 f(\mathbf{r}; \mathbf{p}), \qquad (35)$$

where the symbol $\mathbf{p}$ represents other parameters, namely

$$\mathbf{P} = (F_0, \mathbf{p}). \qquad (36)$$

In the low fluence limit, substituting

$$\mathcal{Y}_{theo}^{(+q)}[F(\mathbf{r}; \mathbf{P})] \propto [F(\mathbf{r}; \mathbf{P})]^{n_q} \qquad (37)$$

and equation (35) into equation (2), it may be shown that

$$\lim_{F_0 \to 0} s(F_0) = n_q. \qquad (38)$$

APPENDIX E
Fluence distribution function

The definition of fluence distribution function is given by

$$V(f) = \int \delta[F(\mathbf{r}) - f]\, d^3 r. \qquad (39)$$

Switching to polar coordinates, and assuming $F(\mathbf{r}) = F(r)$,

$$V(f) = 2\pi \int_0^\infty \frac{\delta(r - r_f)}{|F'(r_f)|} r\, dr, \qquad (40)$$

where r_f satisfies $F(r_f) = f$. For the SGSP,

$$F'(r_f) = -\frac{2\pi a r_f}{\Delta^2} f. \qquad (41)$$

We thus obtain

$$V(f) = \begin{cases} (\Delta^2/a)(1/f) & f \leq F_0, \\ 0 & f > F_0. \end{cases} \qquad (42)$$

Note that equation (42) depends on the fit parameter F_0 only through the cutoff condition at F_0.

In the case of the DGSP, there is no analytical formula for the fluence distribution function. However, it is still possible to understand its overall behavior considering the model

$$f_\mathrm{r} \ll 1 \quad \text{and} \quad w_\mathrm{r} \gg 1. \tag{43}$$

This model was employed before (Murphy *et al.*, 2014) to describe the halo in the spatial profile of XFELs utilizing the second Gaussian profile. First we consider the solution r_f for a fluence value f much smaller than the peak fluence of the second Gaussian profile $f_\mathrm{r} f$, such that $f \ll f_\mathrm{r} F_0$. In such a situation the first Gaussian can be ignored. Then the solution r_f roughly satisfies

$$F(r_f) \simeq f_\mathrm{r} F_0 \exp\left[-\pi a \frac{r_f^2}{(w_\mathrm{r}\Delta)^2}\right], \tag{44}$$

$$F'(r_f) \simeq -\frac{2\pi a}{(w_\mathrm{r}\Delta)^2} r_f f. \tag{45}$$

Therefore, following the derivation for the SGSP, we obtain

$$V(f) \simeq \frac{(w_\mathrm{r}\Delta)^2}{a} \frac{1}{f} \qquad (f \ll f_\mathrm{r} F_0). \tag{46}$$

It is important to realize that the result of equation (46) depends on the additional fit parameter w_r. Therefore comparing with the single Gaussian case, in that sense, the behavior of the fluence distribution function for small fluence values is not universal. The fluence distribution function for a fluence value $f \gg f_\mathrm{r} F_0$ can be also derived ignoring the second Gaussian profile. The result coincides with that of equation (42) in the single Gaussian case.

Funding information

References

Barty, A., Soufli, R., McCarville, T., Baker, S. L., Pivovaroff, M. J., Stefan, P. & Bionta, R. (2009). *Opt. Express*, **17**, 15508–15519.
Chalupský, J., Bohacek, P., Burian, T., Hajkova, V., Hau-Riege, S., Heimann, P., Juha, L., Messerschmidt, M., Moeller, S., Nagler, B., Rowen, M., Schlotter, W., Swiggers, M., Turner, J. & Krzywinski, J. (2015). *Phys. Rev. Appl.* **4**, 014004.
Chalupský, J., Bohacek, P., Hajkova, V., Hau-Riege, S., Heimann, P., Juha, L., Krzywinski, J., Messerschmidt, M., Moeller, S., Nagler, B., Rowen, M., Schlotter, W. F., Swiggers, M. L. & Turner, J. J. (2011). *Nucl. Instrum. Methods Phys. Res. A*, **631**, 130–133.
Chalupský, J., Juha, L., Kuba, J., Cihelka, J., Hájková, V., Koptyaev, S., Krása, J., Velyhan, A., Bergh, M., Caleman, C., Hajdu, J., Bionta, R. M., Chapman, H., Hau-Riege, S. P., London, R. A., Jurek, M., Krzywinski, J., Nietubyc, R., Pelka, J. B., Sobierajski, R., Meyer-Ter-Vehn, J., Tronnier, A., Sokolowski-Tinten, K., Stojanovic, N., Tiedtke, K., Toleikis, S., Tschentscher, T., Wabnitz, H. & Zastrau, U. (2007). *Opt. Express*, **15**, 6036–6043.
Chalupský, J., Krzywinski, J., Juha, L., Hájková, V., Cihelka, J., Burian, T., Vyšín, L., Gaudin, J., Gleeson, A., Jurek, M., Khorsand, A. R., Klinger, D., Wabnitz, H., Sobierajski, R., Störmer, M., Tiedtke, K. & Toleikis, S. (2010). *Opt. Express*, **18**, 27836–27845.
Chung, H.-K., Chen, M. H. & Lee, R. W. (2007). *High Energy Density Phys.* **3**, 57–64.
Ciricosta, O., Chung, H.-K., Lee, R. W. & Wark, J. S. (2011). *High Energy Density Phys.* **7**, 111–116.
Doumy, G., Roedig, C., Son, S.-K., Blaga, C. I., DiChiara, A. D., Santra, R., Berrah, N., Bostedt, C., Bozek, J. D., Bucksbaum, P. H., Cryan, J. P., Fang, L., Ghimire, S., Glownia, J. M., Hoener, M., Kanter, E. P., Krässig, B., Kuebel, M., Messerschmidt, M., Paulus, G. G., Reis, D. A., Rohringer, N., Young, L., Agostini, P. & DiMauro, L. F. (2011). *Phys. Rev. Lett.* **106**, 083002.
Flöter, B., Juranić, P., Kapitzki, S., Keitel, B., Mann, K., Plönjes, E., Schäfer, B. & Tiedtke, K. (2010). *New J. Phys.* **12**, 083015.
Fukuzawa, H., Son, S.-K., Motomura, K., Mondal, S., Nagaya, K., Wada, S., Liu, X.-J., Feifel, R., Tachibana, T., Ito, Y., Kimura, M., Sakai, T., Matsunami, K., Hayashita, H., Kajikawa, J., Johnsson, P., Siano, M., Kukk, E., Rudek, B., Erk, B., Foucar, L., Robert, E., Miron, C., Tono, K., Inubushi, Y., Hatsui, T., Yabashi, M., Yao, M., Santra, R. & Ueda, K. (2013). *Phys. Rev. Lett.* **110**, 173005.
Gao, C., Zeng, J., Li, Y., Jin, F. & Yuan, J. (2013). *High Energy Density Phys.* **9**, 583–593.
Gao, C., Zeng, J. & Yuan, J. (2015). *Contrib. Plasma Phys.* **55**, 123–127.
Ho, P. J., Bostedt, C., Schorb, S. & Young, L. (2014). *Phys. Rev. Lett.* **113**, 253001.
Ho, P. J., Kanter, E. P. & Young, L. (2015). *Phys. Rev. A*, **92**, 063430.
Hoener, M., Fang, L., Kornilov, O., Gessner, O., Pratt, S. T., Gühr, M., Kanter, E. P., Blaga, C., Bostedt, C., Bozek, J. D., Bucksbaum, P. H., Buth, C., Chen, M., Coffee, R., Cryan, J., Dimauro, L., Glownia, M., Hosler, E., Kukk, E., Leone, S. R., McFarland, B., Messerschmidt, M., Murphy, B., Petrovic, V., Rolles, D. & Berrah, N. (2010). *Phys. Rev. Lett.* **104**, 253002.
Jurek, Z., Santra, R., Son, S.-K. & Ziaja, B. (2016a). *Xraypac – a software package for modeling x-ray-induced dynamics of Matter*, Version 1.0, https://www.desy.de/~xraypac/.
Jurek, Z., Son, S.-K., Ziaja, B. & Santra, R. (2016b). *J. Appl. Cryst.* **49**, 1048–1056.
Kayser, Y., Rutishauser, S., Katayama, T., Ohashi, H., Kameshima, T., Flechsig, U., Yabashi, M. & David, C. (2014). *Opt. Express*, **22**, 9004–9015.
Keitel, B., Plönjes, E., Kreis, S., Kuhlmann, M., Tiedtke, K., Mey, T., Schäfer, B. & Mann, K. (2016). *J. Synchrotron Rad.* **23**, 43–49.
Kobayashi, A., Sekiguchi, Y., Oroguchi, T., Yamamoto, M. & Nakasako, M. (2018). *Sci. Rep.* **8**, 831.
Kumagai, Y., Jurek, Z., Xu, W., Fukuzawa, H., Motomura, K., Iablonskyi, D., Nagaya, K., Wada, S. I., Mondal, S., Tachibana, T., Ito, Y., Sakai, T., Matsunami, K., Nishiyama, T., Umemoto, T., Nicolas, C., Miron, C., Togashi, T., Ogawa, K., Owada, S., Tono, K., Yabashi, M., Son, S.-K., Ziaja, B., Santra, R. & Ueda, K. (2018). *Phys. Rev. Lett.* **120**, 223201.
Le Pape, S., Zeitoun, P., Idir, M., Dhez, P., Rocca, J. J. & François, M. (2002). *Phys. Rev. Lett.* **88**, 183901.
Loh, N. D., Starodub, D., Lomb, L., Hampton, C. Y., Martin, A. V., Sierra, R. G., Barty, A., Aquila, A., Schulz, J., Steinbrener, J., Shoeman, R. L., Kassemeyer, S., Bostedt, C., Bozek, J., Epp, S. W.,

Erk, B., Hartmann, R., Rolles, D., Rudenko, A., Rudek, B., Foucar, L., Kimmel, N., Weidenspointner, G., Hauser, G., Holl, P., Pedersoli, E., Liang, M., Hunter, M. S., Gumprecht, L., Coppola, N., Wunderer, C., Graafsma, H., Maia, F. R., Ekeberg, T., Hantke, M., Fleckenstein, H., Hirsemann, H., Nass, K., White, T. A., Tobias, H. J., Farquar, G. R., Benner, W. H., Hau-Riege, S., Reich, C., Hartmann, A., Soltau, H., Marchesini, S., Bajt, S., Barthelmess, M., Strueder, L., Ullrich, J., Bucksbaum, P., Frank, M., Schlichting, I., Chapman, H. N. & Bogan, M. J. (2013). *Opt. Express*, **21**, 12385–12394.

Mehrjoo, M., Giewekemeyer, K., Vagovič, P., Stern, S., Bean, R., Messerschmidt, M., Keitel, B., Plönjes, E., Kuhlmann, M., Mey, T., Schneidmiller, E. A., Yurkov, M. V., Limberg, T. & Mancuso, A. P. (2017). *Opt. Express*, **25**, 17892–17903.

Motomura, K., Fukuzawa, H., Son, S.-K., Mondal, S., Tachibana, T., Ito, Y., Kimura, M., Nagaya, K., Sakai, T., Matsunami, K., Wada, S., Hayashita, H., Kajikawa, J., Liu, X.-J., Feifel, R., Johnsson, P., Siano, M., Kukk, E., Rudek, B., Erk, B., Foucar, L., Robert, E., Miron, C., Tono, K., Inubushi, Y., Hatsui, T., Yabashi, M., Yao, M., Santra, R. & Ueda, K. (2013). *J. Phys. B At. Mol. Opt. Phys.* **46**, 164024.

Murphy, B., Osipov, T., Jurek, Z., Fang, L., Son, S.-K., Mucke, M., Eland, J. H. D., Zhaunerchyk, V., Feifel, R., Avaldi, L., Bolognesi, P., Bostedt, C., Bozek, J. D., Grilj, J., Guehr, M., Frasinski, L. J., Glownia, J., Ha, D. T., Hoffmann, K., Kukk, E., McFarland, B. K., Miron, C., Sistrunk, E., Squibb, R. J., Ueda, K., Santra, R. & Berrah, N. (2014). *Nat. Commun.* **5**, 4281.

Nagler, B., Aquila, A., Boutet, S., Galtier, E. C., Hashim, A. S., Hunter, M., Liang, M., Sakdinawat, A. E., Schroer, C. G., Schropp, A., Seaberg, M. H., Seiboth, F., van Driel, T., Xing, Z., Liu, Y. & Lee, H. J. (2017). *Sci. Rep.* **7**, 13698.

Nilsson, D., Uhlén, F., Holmberg, A., Hertz, H. M., Schropp, A., Patommel, J., Hoppe, R., Seiboth, F., Meier, V., Schroer, C. G., Galtier, E., Nagler, B., Lee, H. J. & Vogt, U. (2012). *Opt. Lett.* **37**, 5046–5048.

Osipov, T. (2013). Personal communication.

Peyrusse, O. (2000). *J. Phys. B At. Mol. Opt. Phys.* **33**, 4303–4321.

Peyrusse, O., Deschaud, B. & Rolles, D. (2014). *J. Phys. B At. Mol. Opt. Phys.* **47**, 011001.

Raschka, S. (2015). *Python Machine Learning.* Packt Publishing Ltd.

Rohringer, N. & Santra, R. (2007). *Phys. Rev. A*, **76**, 033416.

Rösner, B., Döring, F., Ribič, P. R., Gauthier, D., Principi, E., Masciovecchio, C., Zangrando, M., Vila-Comamala, J., De Ninno, G. & David, C. (2017). *Opt. Express*, **25**, 30686–30695.

Rudek, B., Rolles, D., Son, S., Foucar, L., Erk, B., Epp, S., Boll, R., Anielski, D., Bostedt, C., Schorb, S., Coffee, R., Bozek, J., Trippel, S., Marchenko, T., Simon, M., Christensen, L., De, S., Wada, S., Ueda, K., Schlichting, I., Santra, R., Ullrich, J. & Rudenko, A. (2013). *Phys. Rev. A*, **87**, 023413.

Rudek, B., Son, S.-K., Foucar, L., Epp, S. W., Erk, B., Hartmann, R., Adolph, M., Andritschke, R., Aquila, A., Berrah, N., Bostedt, C., Bozek, J., Coppola, N., Filsinger, F., Gorke, H., Gorkhover, T., Graafsma, H., Gumprecht, L., Hartmann, A., Hauser, G., Herrmann, S., Hirsemann, H., Holl, P., Hömke, A., Journel, L., Kaiser, C., Kimmel, N., Krasniqi, F., Kühnel, K., Matysek, M., Messerschmidt, M., Miesner, D., Möller, T., Moshammer, R., Nagaya, K., Nilsson, B., Potdevin, G., Pietschner, D., Reich, C., Rupp, D., Schaller, G., Schlichting, I., Schmidt, C., Schopper, F., Schorb, S., Schröter, C., Schulz, J., Simon, M., Soltau, H., Strüder,

L., Ueda, K., Weidenspointner, G., Santra, R., Ullrich, J., Rudenko, A. & Rolles, D. (2012). *Nat. Photon.* **6**, 858–865.

Rudek, B., Toyota, K., Foucar, L., Erk, B., Boll, R., Bomme, C., Correa, J., Carron, S., Boutet, S., Williams, G. J., Ferguson, K. R., Alonso-Mori, R., Koglin, J. E., Gorkhover, T., Bucher, M., Lehmann, C. S., Krässig, B., Southworth, S. H., Young, L., Bostedt, C., Ueda, K., Marchenko, T., Simon, M., Jurek, Z., Santra, R., Rudenko, A., Son, S. K. & Rolles, D. (2018). *Nat. Commun.* **9**, 4200.

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

Rutishauser, S., Samoylova, L., Krzywinski, J., Bunk, O., Grünert, J., Sinn, H., Cammarata, M., Fritz, D. M. & David, C. (2012). *Nat. Commun.* **3**, 947.

Ryutov, D. D., Bionta, R. M., Hau-Riege, S. P., Kishiyama, K. I., McMahon, D., Roeben, M. D., Shen, S. & Stefan, P. M. (2009). *The Physics of the Gas Attenuator for the Linac Coherent Light Source (LCLS).* Technical report. Lawrence Livermore National Laboratory (LLNL), Livermore, CA, USA.

Sanchez-Gonzalez, A., Micaelli, P., Olivier, C., Barillot, T. R., Ilchen, M., Lutman, A. A., Marinelli, A., Maxwell, T., Achner, A., Agåker, M., Berrah, N., Bostedt, C., Bozek, J. D., Buck, J., Bucksbaum, P. H., Montero, S. C., Cooper, B., Cryan, J. P., Dong, M., Feifel, R., Frasinski, L. J., Fukuzawa, H., Galler, A., Hartmann, G., Hartmann, N., Helml, W., Johnson, A. S., Knie, A., Lindahl, A. O., Liu, J., Motomura, K., Mucke, M., O'Grady, C., Rubensson, J.-E., Simpson, E. R., Squibb, R. J., Såthe, C., Ueda, K., Vacher, M., Walke, D. J., Zhaunerchyk, V., Coffee, R. N. & Marangos, J. P. (2017). *Nat. Commun.* **8**, 15461.

Schäfer, B., Lübbecke, M. & Mann, K. (2006). *Rev. Sci. Instrum.* **77**, 053103.

Schneider, J. R. (2010). *Rev. Accl. Sci. Tech.* **03**, 13–37.

Schneider, M., Günther, C. M., Pfau, B., Capotondi, F., Manfredda, M., Zangrando, M., Mahne, N., Raimondi, L., Pedersoli, E., Naumenko, D. & Eisebitt, S. (2018). *Nat. Commun.* **9**, 214.

Schneider, M., Günther, C. M., von Korff Schmising, C., Pfau, B. & Eisebitt, S. (2016). *Opt. Express*, **24**, 13091–13100.

Schorb, S. (2012). PhD thesis, TU Berlin, Germany.

Schropp, A., Hoppe, R., Meier, V., Patommel, J., Seiboth, F., Lee, H. J., Nagler, B., Galtier, E. C., Arnold, B., Zastrau, U., Hastings, J. B., Nilsson, D., Uhlén, F., Vogt, U., Hertz, H. M. & Schroer, C. G. (2013). *Sci. Rep.* **3**, 1633.

Son, S.-K., Bekx, J. J., Toyota, K., Geffert, O., Slowik, J. M. & Santra, R. (2018). *Xatom – an integrated toolkit for x-ray and atomic physics.* Revision 2964. https://www.desy.de/~xraypac/xatom.html.

Son, S.-K., Young, L. & Santra, R. (2011). *Phys. Rev. A*, **83**, 033402.

Toyota, K., Son, S.-K. & Santra, R. (2017). *Phys. Rev. A*, **95**, 043412.

Xiang, W., Gao, C., Fu, Y., Zeng, J. & Yuan, J. (2012). *Phys. Rev. A*, **86**, 061401.

Young, L., Kanter, E. P., Krässig, B., Li, Y., March, A. M., Pratt, S. T., Santra, R., Southworth, S. H., Rohringer, N., Dimauro, L. F., Doumy, G., Roedig, C. A., Berrah, N., Fang, L., Hoener, M., Bucksbaum, P. H., Cryan, J. P., Ghimire, S., Glownia, J. M., Reis, D. A., Bozek, J. D., Bostedt, C. & Messerschmidt, M. (2010). *Nature*, **466**, 56–61.