PHYSICAL REVIEW E **106**, 015205 (2022)

# Ionization states for the multipetawatt laser-QED regime

I. Ouatu,[1,*] B. T. Spiers,[1,2] R. Aboushelbaya,[1] Q. Feng,[1] M. W. von der Leyen,[1] R. W. Paddock,[1]
R. Timmis,[1] C. Ticos,[3] K. M. Krushelnick,[4] and P. A. Norreys[1,2,5]
[1]*Department of Physics, Atomic and Laser Physics sub-Department, Clarendon Laboratory,*
*University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom*
[2]*Central Laser Facility, UKRI-STFC Rutherford Appleton Laboratory, Didcot, Oxon OX11 0QX, United Kingdom*
[3]*Extreme Light Infrastructure–Nuclear Physics (ELI-NP), Horia Hulubei National*
*Institute for Physics and Nuclear Engineering, Măgurele 077125, Romania*
[4]*Center for Ultra-Fast Optics, University of Michigan, Ann Arbor, Michigan, USA*
[5]*John Adams Institute, Denys Wilkinson Building, Oxford OX1 3RH, United Kingdom*

(Received 29 September 2021; revised 9 April 2022; accepted 30 June 2022; published 27 July 2022)

A paradigm shift in the physics of laser-plasma interactions is approaching with the commissioning of multipetawatt laser facilities worldwide. Radiation reaction processes will result in the onset of electron-positron pair cascades and, with that, the absorption and partitioning of the incident laser energy, as well as the energy transport throughout the irradiated targets. To accurately quantify these effects, one must know the focused intensity on target *in situ*. In this work, a way of measuring the focused intensity on target is proposed based upon the ionization of xenon gas at low ambient pressure. The field ionization rates from two works [Phys. Rev. A **59**, 569 (1999) and Phys. Rev. A **98**, 043407 (2018)], where the latter rate has been derived using quantum mechanics, have been implemented in the particle-in-cell code SMILEI [Comput. Phys. Commun. **222**, 351 (2018)]. A series of one- and two-dimensional simulations are compared and shown to reproduce the charge states without presenting visible differences when increasing the simulation dimensionality. They provide a way to accurately verify the intensity on target using *in situ* measurements.

DOI: 10.1103/PhysRevE.106.015205

## I. INTRODUCTION

There has been tremendous progress towards the construction of multipetawatt (PW) class laser facilities [1,2], providing extreme powers to target [3,4]. Equally intense lasers will be built in the United Kingdom [5], France [6], the United States [7,8], South Korea [9], China [10], and the Russian Federation [11]. Focused laser intensities greater than $10^{23}$ W/cm$^2$ will determine which theoretical model for radiation reaction (RR) applies to laser-plasma interactions such as relativistic tunneling, quantum electrodynamics cascades [12], and prolific pair production [13,14], naming only a few [15].

Historically, the focused intensity has been diagnosed by separately measuring the pulse's spatial and temporal parameters at low power and then extrapolating to the required, higher, peak power [16]. However, realistic wavefront, spectral phase, and temporal distortions make the peak intensities achieved in experiments to be drastically different from their theoretically expected values [17]. This is due to the spatiotemporal couplings (STCs) from a beam pulse of a $O(10)$ PW laser facility: they become increasingly significant as the beam diameter increases and the pulse's duration decreases. The ELI-NP (Apollon) beam diameter is 550 mm [18] (400 mm [17]). To illustrate this effect [19], Fig. 1 shows the degradation of a beam-pulse intensity as a function of pulse front tilt (PFT). For a beam with a PFT of 0.2 fs/mm (where 0.1 fs/mm corresponds to an angular deviance in the parallelism of the compressor gratings of around 20 $\mu$rad [17]) this corresponds to as much as one order of magnitude reduction in the actual peak intensity to target. This will prevent researchers from distinguishing between different theoretical models in the laser-plasma interaction with confidence.

There is no easy way to account for chromatic aberrations and frequency chirps, constraining conventional (direct) methods for ultra-high-power laser pulse characterization; indirect methods have to be used. Work done towards such approaches can be categorized into three main groups: Thomson scattering [20–25], ponderomotive scattering of charged particles from the focus of the laser [26–30], and laser-field ionization of high-$Z$ low-density targets [31–34].

This article describes a method to diagnose the peak intensity, based on laser-field ionization of a rarefied high-$Z$ noble gas, self-consistently modeled using a fully relativistic particle-in-cell (PIC) simulation code. Evidence is provided that computationally efficient one-dimensional (1D) PIC simulations are accurate enough to infer an ultra-high-power laser's peak intensity. In contrast to [32–34], the effect of radiation reaction has been included. The validity of the approach was confirmed by comparison with two-dimensional (2D) simulations: increasing the dimensionality of the PIC simulations enables new effects to appear. Considerations about the practical realization of the diagnostic and how this

---

[*] iustin.ouatu@physics.ox.ac.uk

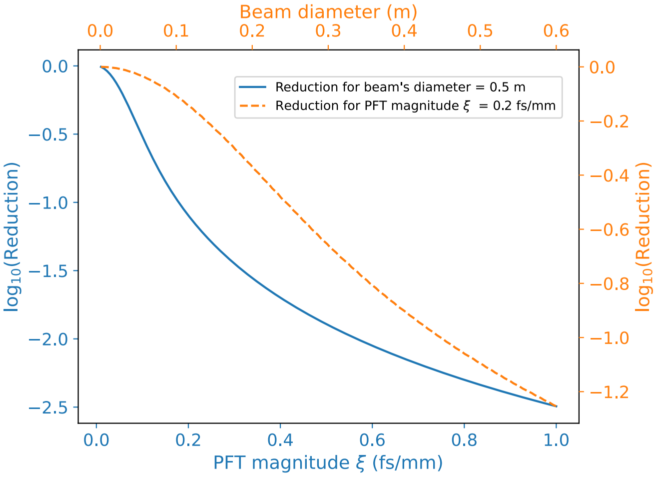

FIG. 1. Ratio of the focused peak intensity of a beam pulse of central wavelength 800 nm and bandwidth of 50 nm having PFT degradation to the focused peak intensity of the STC-free beam pulse (thus called reduction), as a function of the beam's diameter for fixed degradation or as a function of the degradation's magnitude for fixed diameter.

can be used in high-repetition-rate systems are included. Particular attention is paid to the transition between the tunnel ionization regime (TI) and the barrier suppression ionization regime (BSI) (having two subregimes, each described by a different ionization rate). The simulations have been implemented in the SMILEI open-source PIC code [35], along with the piecewise defined rate describing the three regimes based upon state-of-the art research on field ionization [12,36–38].

## II. LASER-FIELD IONIZATION RATES

Experiments confirming TI rates have already been performed, with the assumption that the laser intensity is known so that fits of the collected results were made based upon the theoretical rates [39–43].

The well-known Keldysh parameter [44] is

$$\gamma_K = \frac{E_K}{E_l} = \omega_L \frac{\sqrt{2m_e\,\mathrm{IP}_i}}{e\,E_l}, \tag{1}$$

where $m_e$ is the electron's rest mass, $\mathrm{IP}_i$ the ionization potential (IP) of the atom, $\omega_L$ the frequency of the light, $e$ the fundamental unit of charge, and $E_l$ the peak electric field of the laser pulse. $\gamma_K$ dictates in which regime the ionization takes place: multiphoton (MPI) if $\gamma_K > 1$ or tunnel (TI) if $\gamma_K < 1$. For current state-of-the-art laser-matter interaction experiments, where $\gamma_K \ll 1$, the Perelomov, Popov, and Terent'Ev (PPT) theory for field ionization rates [45,46] is often used. Following [36], the PPT static electric field ionization rate is

$$\begin{aligned} w_{lm} &= \omega_a\, \kappa^2\, C^2_{\kappa l^*}(2l+1)\left(\frac{2}{F}\right)^{2n^*-|m|-1} \\ &\quad\times \frac{(l+|m|)!}{2^{|m|}(|m|)!(l-|m|)!}\ \exp\left(\frac{-2}{3F}\right), \\ C^2_{\kappa l^*} &= \frac{2^{2n^*}}{n^*\Gamma(n^*+l^*+1)\Gamma(n^*-l^*)}, \end{aligned} \tag{2}$$

where $l$ and $m$ are the orbital and magnetic quantum numbers of the electron, respectively, $\omega_a$ the atomic unit of frequency, $\kappa^2 = \mathrm{IP}_i/\mathrm{IP}_H$ with $\mathrm{IP}_H = m_e e^4/(2\hbar^2) \approx 13.6\,\mathrm{eV}$, $F = E/(\kappa^3 E_a)$ with $E$ the local value of the electric field and $E_a$ the atomic unit of electric field, $n^* = Z/\kappa$, $l^* = n^* - 1$, with $Z$ the ion charge number, and $\Gamma(x)$ the Gamma function. This PPT rate was already implemented in SMILEI, and it approaches the Ammosov-Delone-Krainov (ADK) rate [47] when $n^* \gg 1$.

SMILEI resolves the fast temporal oscillations of the laser electric field, thus Eq. (2) does not need to be averaged over one laser period.

However, quantum tunneling occurs only when the potential barrier caused by the combined atomic and laser fields is higher than the initial unperturbed energy level of the electron. If this is not the case, the electron classically escapes above the potential barrier [48]: BSI happens for $E \gtrsim E_{cr} = E_a\,\kappa^4/(16Z)$. For $E > E_{cr}$ the tunnel rates strongly deviate from numerically calculated rates [49], as expected. The transition region between TI and BSI is correctly described by the Bauer-Mulser (BM) rate $w_{\mathrm{BM}} = 2.4\,\omega_a(E/E_a)^2(IP_H/IP_i)^2$, proven to work for $E \sim E_{cr}$ [38]. Furthermore, $w_{\mathrm{BM}}$ deviates from numerically calculated rates as $E \gg E_{cr}$ [36,38]. Derived both classically and quantum mechanically (in the motionless and free electron approximation), the rate dependence on $E$ for $E \gg E_{cr}$ is linear [36]: $w_{\mathrm{BSI}} = p\,\omega_a(E/E_a)\sqrt{(\mathrm{IP}_H/\mathrm{IP}_i)}$. The coefficient $p$ ranges from 0.62 to 0.87, depending on the potential used in the derivation: for the three-dimensional (3D) Coulomb potential $p \approx 0.8$. Numerically integrating the time-dependent Schrödinger equation shows that both approximations used in the quantum derivation are indeed applicable [36]. $w_{lm}$, $w_{\mathrm{BM}}$, and $w_{\mathrm{BSI}}$ are grouped into a piecewise-defined rate [37] which describes the whole range of possible $E$ values at the location of the atom:

$$w(E) \approx \begin{cases} w_{lm}\,, & E \leqslant E_1 \\ w_{\mathrm{BM}}\,, & E_1 < E \leqslant E_2 \\ w_{\mathrm{BSI}}\,, & E > E_2 \end{cases}, \tag{3}$$

where $E_1$ and $E_2$ are found by imposing continuity on $w$ at the transition points. This piecewise-defined rate has been implemented as a functionality in SMILEI.

## III. NUMERICAL SIMULATIONS

### A. One-dimensional particle-in-cell simulations

For a fixed value of the laser's normalized vector potential $a_0 = eE_l/m_e\omega_L c$, a 1D PIC simulation shows how the proportion of the different charge states evolves during the passage of the laser pulse. Running 1D simulations to cover the range $a_0 \in [6.8, \ldots, 2162.8]$ ($\sim I \in [10^{20}, \ldots, 10^{25}]$ PW for a plane wave), the ionization behavior as $I$ is varied can be obtained. Fifteen thousand 1D simulations were run to finely sample this intensity range. Xenon (Xe), atomic number 54, was used, and the initially neutral gas density was chosen such that the electron number density for the fully ionized gas, $n_e$, is low enough to make both collective effects and relativistic self-focusing negligible. The first condition led to us setting the plasma frequency $\omega_p = \sqrt{(n_e\,e^2/m_e\,\epsilon_0)}$ to be such that the plasma oscillation period is significantly longer than the laser

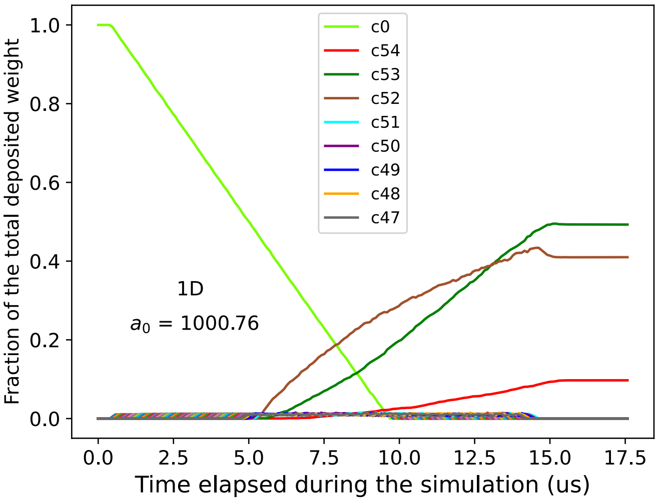

FIG. 2. Xe charge states in a 1D PIC simulation. The gas density was low enough to allow neglecting collective plasma effects and relativistic self-focusing.

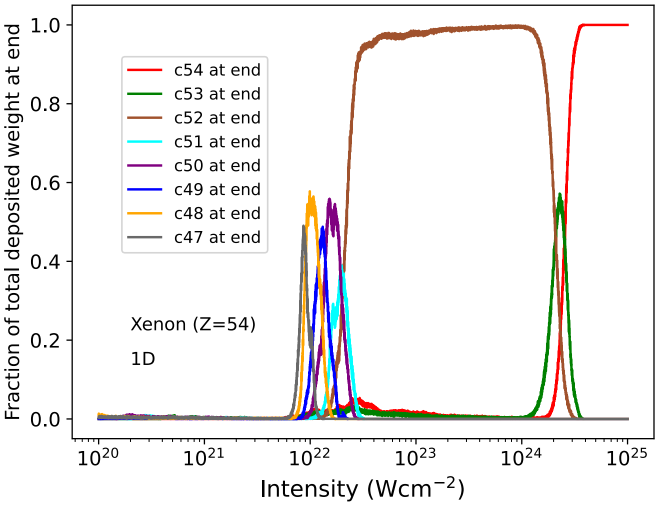

FIG. 3. Xe representative charge states' populations at the end of 15 000 1D PIC simulations. These $c_n(I)$ curves can be used to calculate the number of ions produced in the focus.

pulse [of $O(10)$ fs at ultra-high-power laser facilities]. The second condition led to us setting $n_e$ to be such that $P_L = 10\,\mathrm{PW} < P_{cr} = 17(\omega_L/\omega_p)^2\mathrm{GW}$. To meet both constraints, the neutral Xe number density $n_0 = 1.98 \times 10^{12}\,\mathrm{cm}^{-3}$ was chosen.

Typical results of a 1D simulation are shown in Fig. 2. The charge state dynamics are displayed as a function of time during the simulation. The $y$ axis represents the deposited weights from a particle binning diagnostic for different ion charge states, as a fraction of the total Xe species population's deposited weight. Loosely speaking, one value of any of the $c_n$ curves represents what fraction from the total number of the ions that particular charge state represents, at that particular time during the simulation. The electrons resulting from ionization were injected into the simulation as a separate species which had its RR module turned on. The laser had a wavelength $\lambda_L = 800\,\mathrm{nm}$ and a Gaussian temporal profile with a full-width at half-maximum (FWHM) of six optical cycles (6 $T_L$). The extent of the laser's temporal profile was limited to 10 $T_L$. The simulation box was 12 $\lambda_L$ long, and the spatial resolution was 8. On both sides, Silver-Müller boundary conditions were chosen for the fields, and periodic boundary conditions were chosen for the particles. The simulations lasted for 22 $T_L$, and the number of particles per cell (ppc) for Xe was 32. The same simulations, but with 128 ppc, gave the same results as for those with 32 ppc.

The results for the whole range of intensities appear in Fig. 3. Each data point from one $c_n$ curve corresponds to the fraction of the deposited weight of the charge state Xe$^{n+}$ relative to the total deposited weights of all states at the end of one simulation at a particular input intensity. The charge states values at the end of each simulation are plotted against the corresponding intensities.

### B. Focal averaging and two-dimensional particle-in-cell simulations

To diagnose $I$, the numerically-estimated functions $c_n(I)$'s shown in Fig. 3 are used to infer the number of Xe ions expected to originate from the focus of the laser pulse. A suitable model for $I(\vec{r}\,)$ is chosen, and the number of Xe ions of charge state $n+$ is obtained from an integral over the focal volume:

$$N(\mathrm{Xe}^{n+}) = n_0 \iiint c_n[I(\vec{r}\,)]d^3r. \tag{4}$$

where

$$I(\vec{r}) = \frac{I_m}{1+\frac{z^2}{z_R^2}} \exp\left(-\frac{2\rho^2}{w_z^2}\right) \tag{5}$$

and $\rho = \sqrt{x^2+y^2}$, $I_m$ is the peak intensity, $z_R = \pi w_0^2/\lambda_L$ is the Rayleigh range, $w_0 = 3\,\mu\mathrm{m}$ is the beam waist, and $w_z = w_0\sqrt{1+(z^2/z_R^2)}$.

To numerically calculate the integral from Eq. (4), linear 1D interpolators are used to obtain the values of $c_n$ at $I$ values not simulated using the PIC code but required by the integration routine. For $I$ values outside the PIC-simulated range of $[10^{20}, \ldots, 10^{25}]\,\mathrm{W/cm^2}$, manual extrapolation for the highly charged states (i.e., $c_{47}$ to $c_{54}$ ones) is done. From physical considerations (confirmed by Fig. 3), any highly charged state is not excited for $I < 10^{20}$ W/cm$^2$. Additionally, only the fully ionized state remains for $I > 10^{25}$ W/cm$^2$.

A midpoint rule in two dimensions ($z$ and $\rho$) is employed (details are given in Appendix A).

To verify that the reduced dimensionality of the simulations was not impacting the results, the relationship between $I$ and charge state distribution in two-dimensional (2D) simulations was investigated. For a beam with a Gaussian transverse mode, a 3 $\mu$m radius at focus, and the same $\lambda_L$ and temporal profile as above, the charge states at the end of the simulation for 1000 different $a_0$ values are presented in the left panel of Fig. 4. The middle panel of Fig. 4 shows the spatial distribution of a representative beam's squared electric field from one such 2D simulation. The simulation box was 12 $\lambda_L$ along $x$ and 20 $\lambda_L$ along $y$, with Silver-Müller boundary conditions along $x$ and periodic boundary conditions along $y$.

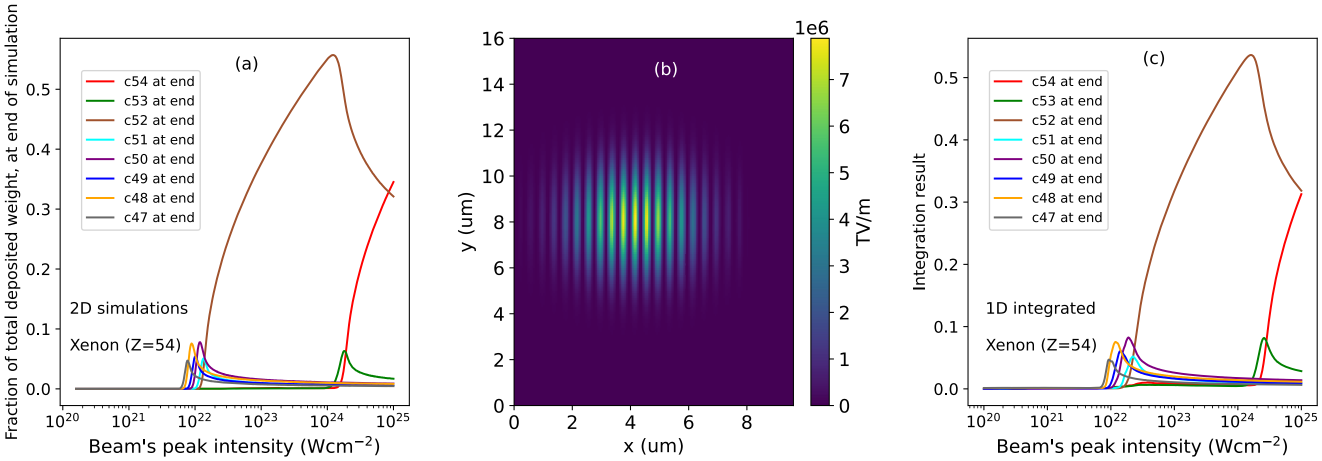

FIG. 4. (a) Xenon representative charge states' populations at the end of 1000 2D Gaussian beam PIC simulations. (b) Squared electric field at one time step, across the 2D PIC simulation box, for a simulation with $a_0 = 707.19$. (c) Results of the integration of the 1D charge states across the 2D PIC simulation box.

To check the consistency of the results between the 1D and 2D simulations, a 2D version of the 3D integral from Eq. (4) of the $c_n$ was performed, while keeping the grid size the same as the 2D simulation box. The right panel of Fig. 4 confirms that the charge states at the end of a 2D simulation reproduce the 1D results.

One needs to estimate the numerical errors of the midpoint rule to ensure the veracity of the results. When done as in Appendix A, the error bars are too small to be visible on the plot of $N(\text{Xe}^{n+})$'s from Fig. 5.

The ionization-induced defocusing effect [50,51] is taken into account by default by the self-consistent modeling of the interaction performed by the PIC code. For the gas densities which are proposed to be used, this effect does not invalidate our method.

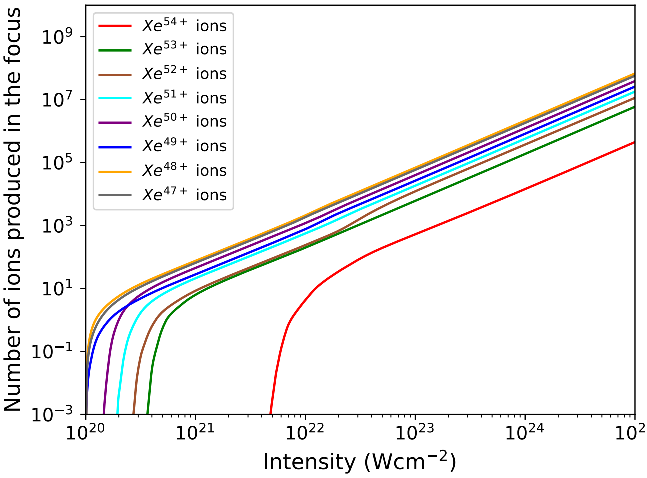

FIG. 5. Number of xenon ions produced in the focus of the Gaussian beam from Eq. (5). One data point on each curve corresponds to the number of ions arising from one, fixed, peak intensity $I_m$ value in Eq. (5). Results on the $y$ axis are numerical calculations for $N(\text{Xe}^{n+})$ from Eq. (4).

## IV. DETECTION OF THE IONS ORIGINATING FROM THE FOCUS OF THE BEAM

The curve $N(\text{Xe}^{n+})$ as function of $I$ was fitted using a relevance vector machine (RVM) [52], as explained in Appendix B. The method can be applied to any curve from Fig. 5. For $N(\text{Xe}^{54+})$, only eight scalar quantities are needed for the posterior predictive distribution, and the mean-squared error obtained is $\sim 9 \times 10^{-3}$ when $\log_{10}$ was taken for both $N(\text{Xe}^{54+})$ and $I$ and then standardized for higher robustness of the algorithm to numerical instabilities.

To count the ions, filtering based on their charge-to-mass ratio is needed, and this is normally done using a Thomson spectrometer. A reasonable estimate for the minimum number of ions of a given charge state to be collected for a clear differentiation of that charge state is $N_{\min} = 50$. One therefore needs to calculate the minimum initial gas density to allow $N_{\min}$ ions to be collected, taking into account realistic distances from the spectrometer pinhole to the interaction point.

The ions are assumed to expand isotropically from the focus (confirmed by the 2D PIC simulations). A circular aperture of diameter $D = 500\,\mu\text{m}$ gives a solid angle of $\Omega = (\pi D^2/4)/R^2$ (the planar area of the aperture is used because its dimensions need to be at least one order of magnitude smaller than the source-aperture distance $R$ based on experimental considerations). The total number of Xe ions produced in the focus is read from Fig. 5 for any charge state and is denoted as $N^{n+}_{\text{total}}$ for charge state $n+$. $N^{n+}_{\text{total}}$ corresponds to the full $4\pi$ sr for an isotropic ejection of the ions.

Only $N_{\text{Th}} = N^{n+}_{\text{total}}\ \Omega/(4\pi)$ ions will be actually collected by the pinhole, for any $n+$. For 50 ions of each charge state to pass through the pinhole, the aperture has to be placed at around or less than $R = 0.1$ cm from the beam's focus if the focused intensity is around $10^{23}\text{W/cm}^2$. This is not realizable in practice. One needs to increase the interaction point-pinhole distance. Doing so, multiple shots are needed to collect enough statistics for a good differentiation of the parabolas. Increasing $D$ decreases the spectrometer's capability to discriminate between different charge states. While providing a higher ion flux at the detector, a very large aperture is not suitable for the envisioned experiments.

The solution for a single-shot experiment is to increase the neutral gas number density $n_0$. However, one needs to avoid relativistic self-focusing: it would provide an artificially high estimate of the peak intensity. The results presented so far are for simulations with $n_0 = 1.98 \times 10^{12}\ \mathrm{cm}^{-3}$ for Xe. These were repeated for $n_0$'s up to three orders of magnitude higher. No self-focusing was observed in the 2D simulations. Increasing $n_0$ leads to a higher number of ions produced in the focus: a single-shot measurement of the peak intensity using Xe is more easily achieved. Depending on the experimental constraints, one can run additional simulations to check when self-focusing appears and choose the most suitable $n_0$ for the optimum source-pinhole distance.

The prepulse of the laser does not influence the results obtained via our diagnostic. For a contrast ratio of $10^8$:1, shooting a pulse at $10^{23}\ \mathrm{W/cm^2}$, the pedestal is of $10^{15}\ \mathrm{W/cm^2}$ and laser-field ionization only starts to develop: only charge states from the lower end will be populated. Reference [32] demonstrated that any initial state of the target, as long as it contains nonzero populations only in the lower end of the charge states, produces the same charge state distribution after the passage of the laser pulse.

Simulations for argon were also performed, and they show that it is completely ionized at $\sim I_{\max}^{\mathrm{Ar}} = 3 \times 10^{21}\ \mathrm{W/cm^2}$. For facilities delivering ultra-high power pulses at high repetition rates, assuming the beam leaks through a mirror after the compressor, this method is suitable to be an *in situ* diagnostic for the peak intensity. Alternatively, a reflection from a blast-shield glass plate or transmission through a turning mirror are equally well suited to transport the beam into a secondary chamber to perform ionization measurements there. This allows measuring the peak intensity regardless of the experiment conducted in the main interaction chamber.

## V. CONCLUSION

To conclude, a method to diagnose the peak intensity of ultra-intense laser pulses has been presented. It can be easily implemented and requires modest computational resources and access to the open-source modified SMILEI PIC code [53]. The numerical experiments run on an Intel i7-7500U in one day of wall clock time. This work allows transitioning from currently performed commissioning experiments on $O(10)$ PW laser facilities to the community's first experiments by providing an accurate assessment of the true intensity on the target.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of the staff of the Central Laser Facility, Rutherford Appleton Laboratory. Useful discussions on electrons' trajectories tracking with Robert Bajbak and Jan Psikal are acknowledged. Part of the simulations presented herein were run on the ARCHER2 UK National Supercomputing Service [54] under UKRI-EPSRC Grant No. EP/R029148/1. The work was carried out under the John Adams Institute Grant No. UKRI-STFCST/V001655/1. I.O. acknowledges support from an UKRI-EPSRC DTP grant to the University of Oxford. C.T. acknowledges support from the PN19060105-2021 grant from MCID. R.A. and P.A.N. acknowledge support from the Oxford-ShanghaiTech collaboration agreement, and K.M.K. and P.A.N. acknowledge support from the Leverhulme Trust.

## APPENDIX A: NUMERICAL INTEGRATION DETAILS

A 2D integration scheme is used because Gaussian beams are symmetric with respect to the azimuthal dimension. A 2D cylindrical integral is thus performed, being then multiplied by $2\pi$ to calculate the full 3D integral.

The integration ranges are obtained by considering that below a threshold intensity the $c_{47}$ to $c_{54}$ charge states are not excited. From Fig. 3, $I_{\mathrm{thr}} = 10^{20}\ \mathrm{W/cm^2}$.

The integral in $z$ is performed across $[-z_{\max}, z_{\max}]$, with $z_{\max} = z_R (I_m/I_{\mathrm{thr}} - 1)^{0.5}$.

At each step along $z$, the integration along $\rho$ is performed across $[0, \rho_{\max}]$, where $\rho_{\max}$ is obtained by imposing that the intensity at that particular $z$ and $\rho_{\max}$ equals $I_{\mathrm{thr}}$, thus

$$\rho_{\max} = \sqrt{\frac{-w_z^2}{2} \ln\left[\left(1 + \frac{z^2}{z_R^2}\right)\frac{I_{\mathrm{thr}}}{I_m}\right]}. \tag{A1}$$

The number of steps in $z$ was $N_z = 2^{12}$ and the number of steps in $\rho$, for each $z$ value, was $N_\rho = 2^{12}$.

One needs to estimate the numerical errors of the computational method to ensure the veracity of the results. The $N(\mathrm{Xe}^{n+})$'s from Eq. (4), obtained using the 2D midpoint rule, are presented in Fig. 5. There integral estimates in the theory of converging sequences of numerical approximations [55] are shown (after being multiplied by $n_0$). The result estimate for one integral is denoted as $I_M$, where $M$ is the highest number of points the laptop can accommodate, $M = 2^{12}$ for both $z$ and $\rho$ axes. The same integral, with half the number of points along both $z$ and $\rho$, was performed and stored as $I_{M/2}$. The numerical error on a result $I_M$ was determined to be $|I_M - I_{M/2}|$. After being multiplied by $n_0$, these approximations to the numerical errors are too small to be visible on the plot.

## APPENDIX B: FIT FOR THE NUMBER OF IONS ORIGINATING FROM THE FOCUS OF THE LASER BEAM

It is possible to train a regressor on the curves from our Fig. 4 and then query it on unseen examples. Curve fitting when the underlying data-generating distribution is unknown is more and more prevalent nowadays in the form of machine learning.

Bayesian thinking was adopted: this framework naturally penalizes heavily complex models which overfit drastically when tested on the hold-out set. The resulting fitting equation is envisioned to be practical, thus Bayesian linear regression with the basis functions being the simple polynomials was performed. Our regressor acquires a mean-squared error on the test set of around $2 \times 10^{-3}$ when the targets have been standardized (mean made to be equal to 0 and variance made to be equal to 1). This is comparable to results which have been obtained using state-of-the-art ML algorithms, such as the XGBoost [56] version of the ensemble of weak learners (i.e., boosted trees). The optimization algorithms implemented in Optuna [57] were used for setting the hyperparameters of the XGboost (in particular an algorithm based on the evolutionary optimization strategy CMA-ES and an algorithm based on trees called TPE). These algorithms were let run to try to find suitable values for the hyperparameters for 15 000 trials, and the objective of their minimization was the mean-squared-error (MSE) on the validation set.

The data set is the hardest-to-model curve from our Fig. 4, i.e., the one containing the fewest data points: the fully ionized ions. It is denoted as $\mathcal{D} = \{(x_i, t_i)\colon i = 1, 2, \ldots, N\}$, where $N = 681$, $x_i$ represents the intensity, and $t_i$ represents the number of ions from the focus. The usual over-top arrow for vectors is used, and a vector will be expressed through a lower-case letter, while a matrix will be expressed as an upper-case letter. Anything written in normal font and lower-case letters is a real-valued scalar. The noise is modeled as Gaussian, i.e., it comes from a normal distribution with mean 0 and variance $1/\beta^2$. The polynomial features are expressed, for one input example $x_n$, as $\vec{\phi}(x_n) = [1, x_n, x_n^2, \ldots, x_n^M]$. The 2-dimensional design matrix $\Phi$ contains across its rows the $\vec{\phi}(x_n)$'s, for $n \in \{1, 2, \ldots, N\}$. Although non-linear in the input $x$, the model is linear in the fit parameters:

$$p(t|x, \vec{\theta}) = \mathcal{N}\left(t \mid \vec{w}^T \vec{\phi}(x), \frac{1}{\beta^2}\right), \tag{B1}$$

where $\vec{\theta}$ is a generic notation for the parameters of the model and $\vec{w}$ is a column vector with $M$ elements. The likelihood of the data $\vec{t} = [t_1, t_2, \ldots, t_N]^T$ and $X_i = [x_i]$ for $i \in \{1, 2, \ldots, N\}$ (thus $X$ is shape $N \times 1$), i.e., the joint likelihood, given the assumption that the data points are drawn in an i.i.d. fashion (identically and independently distributed), is

$$p(\vec{t} \mid X, \vec{w}, \beta) = \prod_{i=1}^{N} p(t_n|x_n, \vec{w}, \beta), \tag{B2}$$

thus a Gaussian probability distribution function. A prior on the weights $\vec{w}$ is put as

$$p(\vec{w}|\vec{\alpha}) = \prod_{i=1}^{M} \mathcal{N}\left(w_i \mid 0, \frac{1}{\alpha_i}\right), \tag{B3}$$

where $\vec{\alpha} = [\alpha_0, \alpha_1, \alpha_2 \ldots, \alpha_M]^T$. The posterior on the weights, from the Bayes' rule, is proportional to the product of the joint likelihood and the prior on the weights. As both of the terms are Gaussian, the resulting product is again Gaussian:

$$p(\vec{w}|\vec{t}, X, \vec{\alpha}, \beta) = \mathcal{N}(\vec{w} \mid \vec{m}, \Sigma), \tag{B4}$$

where

$$\vec{m} = \beta \Sigma \Phi^T \vec{t}, \quad \Sigma = (A + \beta \Phi^T \Phi)^{-1}, \tag{B5}$$

where $A = \operatorname{diag}(\vec{\alpha})$ (thus shape $M \times M$). For an infinitely broad prior, $\alpha_i \to 0$ for $i \in \{1, 2, \ldots, M\}$, the mode of the posterior $\vec{w}_{\text{MAP}} = \arg\max_{\vec{w}} p(\vec{w}|\vec{t}, X, \vec{\alpha}, \beta)$ tends to the maximum likelihood estimate of type I for the weights, $\vec{w}_{\text{MLE}}$, commonly used in simple fits of data (which almost always tend to overfit, especially in the small-data-set regime).

The priors on $\vec{\alpha}$ and $\beta$ are obtained through Empirical Bayes (maximum likelihood type II). They are those values which maximize the marginal (log) likelihood, i.e., the integral of the likelihood against the prior on the weights across all the parameters space $\vec{w}$:

$$p(\vec{t}|X, \vec{\alpha}, \beta) = \int_{\vec{w}} p(\vec{t}|X, \vec{w}, \beta) p(\vec{w}|\vec{\alpha})\, d\vec{w}. \tag{B6}$$

This integral is the correlation of two Gaussians and using the "Bayesian-Gaussian identity": $\int_{\vec{z}} \mathcal{N}(\vec{y} \mid W\vec{z} + \vec{b},\ \Sigma_y)\, \mathcal{N}(\vec{z} \mid \vec{\mu}_z,\ \Sigma_z)\, d\vec{z} = \mathcal{N}(\vec{y} \mid W\vec{\mu}_z + \vec{b},\ \Sigma_y + W\Sigma_y W^T)$, we obtain that

$$p(\vec{t} \mid X, \vec{\alpha}, \beta) = \mathcal{N}(\vec{t} \mid \vec{0}, C), \tag{B7}$$

where

$$C = \beta^{-1}\mathbb{1} + \Phi A^{-1} \Phi^T, \tag{B8}$$

with $\mathbb{1}$ being the identity matrix. This marginal likelihood has to be maximized w.r.t. $\vec{\alpha}$ and $\beta$, and this is a nonconvex optimization problem. One could perform the expectation-maximization (EM) algorithm [58] on it, but Refs. [52,59] have been followed here: $\vec{\alpha}$ and $\beta$ have been iteratively updated as below, up until a convergence criterion was achieved:

$$\alpha_i^{\text{new}} = \frac{\gamma_i}{m_i^2}, \quad (\beta^{\text{new}})^{-1} = \frac{||\vec{t} - \Phi\vec{m}||^2}{N - \sum_i \gamma_i}, \tag{B9}$$

where $\gamma_i = 1 - \Sigma_{ii}$, $m_i$ is the $i$th component of $\vec{m}$ and $||\vec{a}||^2 = \sum_i a_i^2$ for a real-valued $\vec{a}$ and where the sum runs over all elements of $\vec{a}$. In practice, the EM algorithm tends to provide a slower convergence than what we used [52]. The hierarchical Bayesian inference model has been stopped here: the optimization algorithm was simply started from very small values of $\vec{\alpha}$ and $\beta$. We mention that one could put a hyperprior on $\vec{\alpha}$ and $\beta$ as well in order to be even more Bayesian (Gamma distributions, which in turn have parameters, but set by hand to fixed values), an idea known in the literature [52].

After the optimization routine has run, it is found that some of the elements of $\vec{\alpha}$ were increased dramatically, up to infinity. This translates into the fact the the posterior on those weights $w_i$ for which the $\alpha_i \to \infty$ is centered at 0 and has a virtually 0 variance [52,59]. The explanation for this is that the model prunes the basis functions corresponding to these $w_i$'s: they have no predictive power for the problem we are trying to solve, given our data set. In this way the model naturally balances complexity with the avoidance of overfitting, beautifully incorporating the Occam's razor idea. Instead of using regularizers and expensive k-fold-cross-validation (CV) runs, this Bayesian approach to learning achieves (loosely speaking) what l1 and l2 regularizations, for example, achieve. Details can be found in [59].

During the run of the optimization routine we were confronted with numerical instabilities because of those $\alpha_i$ which shoot up to $\infty$, especially when the maximum order of the polynomial was really high (>30, for example). The original procedure [52] has been followed, and as explained in Appendix B 1 from that paper, the columns of $\Phi$ corresponding to those values of $\alpha_i$ which became large during the optimization routine were removed: in particular, when $\gamma_i < \epsilon$, where $\epsilon$ is the machine precision, that element $i$ from the vector $\vec{\alpha}$ is removed, similarly with column $i$ from $\Phi$. These removals cause no harm for the remaining steps of the optimization routine running to maximize the marginal likelihood (details are found in the same Appendix).

Having found the optimal $\vec{\alpha}$ and $\beta$ values, denoted as $\vec{\alpha}^*$ and $\beta^*$, the predictive distribution of the model is again a Gaussian and reads

$$p(t|x, X, \vec{t}, \vec{\alpha}^*, \beta^*) = \int_{\vec{w}} p(t|x, \vec{w}, \beta^*)p(\vec{w}|X, \vec{t}, \vec{\alpha}^*, \beta^*)\, d\vec{w} = \mathcal{N}[t|\vec{m}^T\vec{\phi}(x), \sigma^2(x)], \tag{B10}$$

where

$$\sigma^2(x) = (\beta^*)^{-1} + \vec{\phi}(x)^T \Sigma\, \vec{\phi}(x), \tag{B11}$$

with $\Sigma$ from Eq. (B5) with $\beta$ set to $\beta^*$ and $A$ arising from $\vec{\alpha}^*$.

Because it is envisioned that the fitted equation shall be practical, the maximum order of the polynomial which is proposed cannot be high. In order to choose which model to use, the Bayesian view was again adopted and the model evidence has been used to rank the models. That is, models with $M \in \{1, 2, 3, \dots\}$ were tried: Empirical Bayes was applied to each of them, and then the model evidence was calculated for each of them. The model evidence is the normalizing constant (the denominator) from the Bayes' rule for the parameters $\vec{w}$:

$$p(\vec{w}|D, \mathcal{H}) = \frac{p(D|\vec{w}, \mathcal{H})p(\vec{w}|\mathcal{H})}{p(D|\mathcal{H})}. \tag{B12}$$

For no *a priori* preference for a particular model, given a set of models $\{\mathcal{H}_j{:}j = 1, 2, \dots, J\}$, $p(\mathcal{H}_j) = 1/J$. The Bayes' rule reads

$$p(\mathcal{H}_j|\mathcal{D}) \propto p(\mathcal{D}|\mathcal{H}_j)p(\mathcal{H}_j), \tag{B13}$$

and the ranking of two competing models is done based on the Bayes' factor, which is equal to the ratio of posterior probabilities of the models in light of the data $\mathcal{D}$, in turn equal to the models' evidence for no *a priori* preferences over them:

$$B_{ab} = \frac{p(\mathcal{H}_a|\mathcal{D})}{p(\mathcal{H}_b|\mathcal{D})} = \frac{p(\mathcal{D}|\mathcal{H}_a)}{p(\mathcal{D}|\mathcal{H}_b)}. \tag{B14}$$

Calculating a generic model evidence is a complicated task and is well known in literature. In the model presented here, because everything is analytical, the model evidence is in closed form: usage of Markov-chain Monte Carlo (MCMC) or simplifying approximations, such as the Laplace one, or any of the information criteria appearing in the literature (Akaike, Bayesian, Schwarz, etc.) which are used for model comparison, has not been made. The log of the marginal likelihood, that is, the log of the model evidence, from Eq. (B7), reads, for one model,

$$\ln p(\vec{t}\,|\,X, \vec{\alpha}^*, \beta^*) = -\tfrac{1}{2}[\,N \ln(2\pi) + \ln|C| + \vec{t}^T\, C^{-1}\, \vec{t}\,], \tag{B15}$$

where $C$ is evaluated from Eq. (B8) with $\beta \leftarrow \beta^*$ and $\vec{\alpha} \leftarrow \vec{\alpha}^*$ and $|C|$ represents the determinant of the matrix $C$.

The obtained log-model evidence [from Eq. (B15)] as a function of the polynomial order of the model is shown in Fig. 6.

Because of how the RVM algorithm is designed, it removes and does not count on the basis functions which make the model infer a complexity penalty greater than the benefit obtained from a better fit to the training data. This is what has been explained above when $\alpha_i \to \infty$.

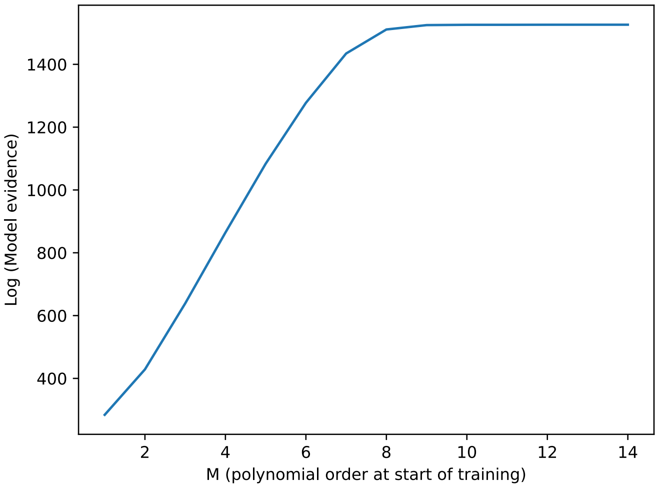

FIG. 6. Model evidence as a function of the polynomial order at the start of the training procedure.

Because the fitting equation has to be practical, the maximum polynomial order, that is, $M$, which was chosen is one for which the number of remaining (nonpruned) weights is less than a reasonable number, which translates in the ability to list all the nonzero elements from the resulting $\vec{\alpha}^*$. That reasonable number was chosen to be 7. For this, $M = 14$ in our case. There might be the case that as $M$ increases significantly more than what was chosen, there will be so many pruned basis functions in those high-$M$ models that the number of the remaining ones would be less than our reasonable number. This issue is raised because for those large $M$'s, the model evidence might be higher than the model evidence for what was chosen ($M = 14$). This was studied, and in the current work, the remaining number of basis functions greatly exceeds our reasonable number, making it clear that any of those higher-$M$ models cannot be reported.

The choice of the polynomial order can be analyzed also from the point of view of the MSE on the test-set diagnostic. The test set contained 103 examples. The MSE values as a function of the polynomial order are presented in Fig. 7. Although a stochastic quantity suffering from noise (especially in the low amounts of data regime), this MSE on the test set provides an insight into the performance of the model and is widely used in the frequentist view of model selection (albeit with expensive cross-validation runs).

We emphasize that the model evidence and the MSE on the test set cannot be simply thought to be perfectly anticorrelated: the maximum of the model evidence does not necessarily mean the minimum MSE on the test set, for any given data set $\mathcal{D}$. It can be shown that if the data were indeed generated from one of the models analyzed, on average the Bayes' factor between the correct model and any of the other models favors the correct model (giving it maximum evidence among them all). This is because the average Bayes' factor can be written as a KL divergence, which is always $>0$ unless one of the distributions is identically 0. Still, there might be some data sets on which this does not happen: the "correct" model might not be favored. This is why Bayesian

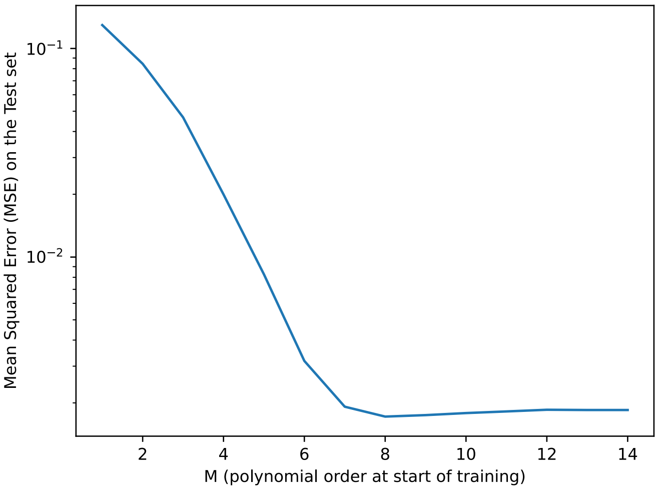

FIG. 7. Test-set mean-squared error as a function of the polynomial order at the start of the training procedure.

model averaging should be done (i.e., predictions shall be done using all the ranked models, each of the contributions being weighted by the corresponding evidence), but in this work a pointwise approximation to this has been used and the model with the highest evidence was chosen, understanding the limitations this approach poses, but also the benefits it offers. The relationship between Bayesian model selection and frequentist approaches is beyond the scope of the work from this Appendix [60].

The predictive distribution [Eq. (B10)] for one unseen example $x$ is again written in order to emphasize that one only needs $\vec{\alpha}^*$, $\beta^*$ and the train dataset in order to predict new values:

$$\begin{aligned} p(t|x, X, \vec{t}, \vec{\alpha}^*, \beta^*) &= \int_{\vec{w}} p(t|x, \vec{w}, \beta^*) p(\vec{w}|X, \vec{t}, \vec{\alpha}^*, \beta^*)\, d\vec{w} \\ &= \mathcal{N}[t|\vec{m}^T\vec{\phi}(x), \sigma^2(x)], \end{aligned} \tag{B16}$$

where

$$\vec{m} = \beta^*\Sigma\Phi^T\vec{t}, \tag{B17}$$

with

$$\Sigma = (A + \beta^*\Phi^T\Phi)^{-1} \tag{B18}$$

and

$$A = \mathrm{diag}(\vec{\alpha}^*), \tag{B19}$$

$$\sigma^2(x) = (\beta^*)^{-1} + \vec{\phi}(x)^T\Sigma\vec{\phi}(x). \tag{B20}$$

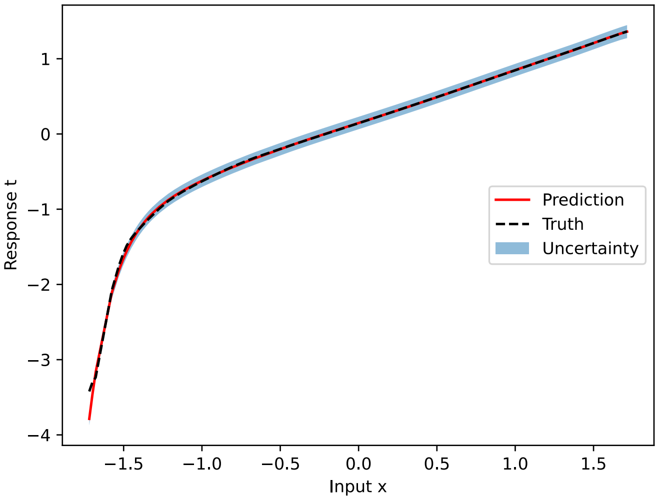

FIG. 8. Predicted response as a function of the input for the model for which we list the values of $\vec{\alpha}^*$ and $\beta^*$. The two-sigma deviation is highlighted in blue and represents the 95% uncertainty in the predictions, as from Eq. (B10).

As mentioned earlier in this Appendix, the data set has been standardized to have mean 0 and variance 1. This is a common technique to make the algorithms which will be deployed on the data set more robust to numerical instabilities. Before subtracting the mean and dividing by the standard deviation, the logarithm in base 10 of both the response but also of the predictors was taken, thus again of the whole data set. This was because the data set spans many orders of magnitude (intensity between $10^{20}$ W/cm$^2$ and $10^{25}$ W/cm$^2$ and the number of ions between $10^{-3}$ and $10^6$, or even $10^8$, depending on the considered charge state).

Thus when predicting the response at unseen examples, the input to the regressor shall be scaled accordingly, similarly for the obtained response.

The obtained nonzero elements of $\vec{\alpha}^*$ and the scalar $\beta^*$ are now listed to two significant digits:

$$\begin{aligned} &\alpha_0^* = 4.7 \times 10^1, \quad \alpha_1^* = 2.2, \quad \alpha_3^* = 1.6 \times 10^2, \\ &\alpha_4^* = 1.1 \times 10^3, \quad \alpha_5^* = 3.3 \times 10^3, \quad \alpha_{11}^* = 9.4 \times 10^4, \\ &\alpha_{12}^* = 3.7 \times 10^5, \quad \beta^* = 5.9 \times 10^3. \end{aligned} \tag{B21}$$

The plot of the predictions on the test set is presented in Fig. 8, where it is highlighted in blue the two-sigma uncertainty of the predicted quantity [the posterior predictive is a Gaussian; see Eq. (B10)].

[1] C. N. Danson, C. Haefner, J. Bromage, T. Butcher, J.-C. F. Chanteloup, E. A. Chowdhury, A. Galvanauskas, L. A. Gizzi, J. Hein, D. I. Hillier, N. W. Hopps, Y. Kato, E. A. Khazanov, R. Kodama, G. Korn, R. Li, Y. Li, J. Limpert, J. Ma, C. H. Nam *et al.*, High Power Laser Sci. Eng. **7**, E54 (2019).

[2] D. Strickland and G. Mourou, Opt. Commun. **55**, 447 (1985).

[3] J.-P. Chambaret, O. Chekhlov, G. Cheriaux, J. Collier, R. Dabu, P. Dombi, A. M. Dunne, K. Ertel, P. Georges, J. Hebling, J. Hein, C. Hernandez-Gomez, C. Hooker, S. Karsch, G. Korn, F. Krausz, C. Le Blanc, Zs. Major, F. Mathieu, T. Metzger *et al.*, Proc. SPIE **7721**, 77211D (2010).

[4] S. Weber, S. Bechet, S. Borneis, L. Brabec, M. Bučka, E. Chacon-Golcher, M. Ciappina, M. DeMarco, A. Fajstavr, K. Falk, E.-R. Garcia, J. Grosz, Y.-J. Gu, J.-C. Hernandez, M. Holec, P. Janečka, M. Jantač, M. Jirka, H. Kadlecova, D. Khikhlukha *et al.*, Matter Radiat. Extremes **2**, 149 (2017).

[5] CLF Vulcan 2020 Upgrade, https://www.clf.stfc.ac.uk/Pages/Vulcan-2020.aspx.

[6] B. Le Garrec and the Apollon team (Optical Society of America, San Jose, California United States, 2017), p. SF1K.3.

[7] J. Nees, A. Maksimchuk, G. Kalinchenko, B. Hou, Y. Ma, P. Campbell, A. McKelvey, L. Willingale, I. Jovanovic, C. Kuranz, A. Thomas, and K. Krushelnick (Optical Society of America, 2020), p. JW2B.9.

[8] J. Bromage, S.-W. Bahk, I. A. Begishev, C. Dorrer, M. J. Guardalben, B. N. Hoffman, J. B. Oliver, R. G. Roides, E. M. Schiesser, M. J. S. Iii, M. Spilatro, B. Webb, D. Weiner, and J. D. Zuegel, High Power Laser Sci. Eng. **7**, E4 (2019).

[9] J. W. Yoon, Y. G. Kim, I. W. Choi, J. H. Sung, H. W. Lee, S. K. Lee, and C. H. Nam, Optica **8**, 630 (2021).

[10] B. Shen, Z. Bu, J. Xu, T. Xu, L. Ji, R. Li, and Z. Xu, Plasma Phys. Controlled Fusion **60**, 044002 (2018).

[11] A. Bashinov, A. Gonoskov, A. Kim, G. Mourou, and A. Sergeev, Eur. Phys. J.: Spec. Top. **223**, 1105 (2014).

[12] I. I. Artemenko and I. Y. Kostyukov, Phys. Rev. A **96**, 032106 (2017).

[13] A. R. Bell and J. G. Kirk, Phys. Rev. Lett. **101**, 200403 (2008).

[14] C. P. Ridgers, C. S. Brady, R. Duclous, J. G. Kirk, K. Bennett, T. D. Arber, A. P. L. Robinson, and A. R. Bell, Phys. Rev. Lett. **108**, 165006 (2012).

[15] I. C. E. Turcu, F. Negoita, D. A. Jaroszynski, P. Mckenna, S. Balascuta, D. Ursescu, I. Dancus, M. O. Cernaianu, M. V. Tataru, P. Ghenuche, D. Stutman, A. Boianu, M. Risca, M. Toma, C. Petcu, G. Acbas, S. R. Yoffe, A. Noble, B. Ersfeld, E. Brunetti *et al.*, Romanian Reports in Physics **68**, S145 (2016).

[16] V. Yanovsky, V. Chvykov, G. Kalinchenko, P. Rousseau, T. Planchon, T. Matsuoka, A. Maksimchuk, J. Nees, G. Cheriaux, G. Mourou, and K. Krushelnick, Opt. Express **16**, 2109 (2008).

[17] G. Pariente, V. Gallet, A. Borot, O. Gobert, and F. Quéré, Nat. Photonics **10**, 547 (2016).

[18] S. Gales, K. A. Tanaka, D. L. Balabanski, F. Negoita, D. Stutman, O. Tesileanu, C. A. Ur, D. Ursescu, I. Andrei, S. Ataman, M. O. Cernaianu, L. D'Alessi, I. Dancus, B. Diaconescu, N. Djourelov, D. Filipescu, P. Ghenuche, D. G. Ghita, C. Matei, K. Seto, M. Zeng, and N. V. Zamfir, Rep. Prog. Phys. **81**, 94301 (2018).

[19] Z. Li, K. Tsubakimoto, H. Yoshida, Y. Nakata, and N. Miyanaga, Appl. Phys. Express **10**, 102702 (2017).

[20] E. S. Sarachik and G. T. Schappert, Phys. Rev. D **1**, 2738 (1970).

[21] S.-y. Chen, A. Maksimchuk, and D. Umstadter, Nature (London) **396**, 653 (1998).

[22] C. Z. He, A. Longman, J. A. Pérez-Hernández, M. de Marco, C. Salgado, G. Zeraouli, G. Gatti, L. Roso, R. Fedosejevs, and W. T. Hill, Opt. Express **27**, 30020 (2019).

[23] O. Har-Shemesh and A. D. Piazza, Opt. Lett. **37**, 1352 (2012).

[24] W. Yan, C. Fruhling, G. Golovin, D. Haden, J. Luo, P. Zhang, B. Zhao, J. Zhang, C. Liu, M. Chen, S. Chen, S. Banerjee, and D. Umstadter, Nat. Photonics **11**, 514 (2017).

[25] O. E. Vais, S. G. Bochkarev, and V. Y. Bychenkov, Plasma Phys. Rep. **42**, 818 (2016).

[26] F. Mackenroth, A. R. Holkundkar, and H.-P. Schlenvoigt, New J. Phys. **21**, 123028 (2019).

[27] O. E. Vais, S. G. Bochkarev, S. Ter-Avetisyan, and V. Y. Bychenkov, Quantum Electron. **47**, 38 (2017).

[28] O. E. Vais, A. G. R. Thomas, A. M. Maksimchuk, K. Krushelnick, and V. Y. Bychenkov, New J. Phys. **22**, 023003 (2020).

[29] O. E. Vais and V. Y. Bychenkov, Plasma Phys. Controlled Fusion **63**, 014002 (2021).

[30] K. Krushelnick, E. Clark, Z. Najmudin, M. Salvati, M. I. K. Santala, M. Tatarakis, A. E. Dangor, V. Malka, D. Neely, R. Allott, and C. Danson, Laser Part. Beams **18**, 595 (2000).

[31] A. Link, E. A. Chowdhury, J. T. Morrison, V. M. Ovchinnikov, D. Offermann, L. Van Woerkom, R. R. Freeman, J. Pasley, E. Shipton, F. Beg, P. Rambo, J. Schwarz, M. Geissel, A. Edens, and J. L. Porter, Rev. Sci. Instrum. **77**, 10E723 (2006).

[32] M. F. Ciappina, S. V. Popruzhenko, S. V. Bulanov, T. Ditmire, G. Korn, and S. Weber, Phys. Rev. A **99**, 043405 (2019).

[33] M. F. Ciappina and S. V. Popruzhenko, Laser Phys. Lett. **17**, 025301 (2020).

[34] M. F. Ciappina, E. E. Peganov, and S. V. Popruzhenko, Matter Radiat. Extremes **5**, 044401 (2020).

[35] J. Derouillat, A. Beck, F. Pérez, T. Vinci, M. Chiaramello, A. Grassi, M. Flé, G. Bouchard, I. Plotnikov, N. Aunai, J. Dargent, C. Riconda, and M. Grech, Comput. Phys. Commun. **222**, 351 (2018).

[36] I. Y. Kostyukov and A. A. Golovanov, Phys. Rev. A **98**, 043407 (2018).

[37] A. A. Golovanov and I. Y. Kostyukov, Quantum Electron. **50**, 350 (2020).

[38] D. Bauer and P. Mulser, Phys. Rev. A **59**, 569 (1999).

[39] E. A. Chowdhury, C. P. J. Barty, and B. C. Walker, Phys. Rev. A **63**, 042712 (2001).

[40] E. A. Chowdhury and B. C. Walker, J. Opt. Soc. Am. B **20**, 109 (2003).

[41] S.-F. Zhao, L. Liu, and X.-X. Zhou, Opt. Commun. **313**, 74 (2014).

[42] Y. H. Lai, J. Xu, U. B. Szafruga, B. K. Talbert, X. Gong, K. Zhang, H. Fuest, M. F. Kling, C. I. Blaga, P. Agostini, and L. F. DiMauro, Phys. Rev. A **96**, 063417 (2017).

[43] K. Yamakawa, Y. Akahane, Y. Fukuda, M. Aoyama, N. Inoue, and H. Ueda, Phys. Rev. A **68**, 065403 (2003).

[44] L. V. Keldysh, J. Exp. Theor. Phys. (U.S.S.R.) **47**, 1945 (1964) [Sov. Phys. JETP **20**, 1370 (1964)].

[45] V. S. Popov, Phys. Usp. **47**, 855 (2004).

[46] A. M. Perelomov, V. S. Popov, and M. V. Terent'Ev, J. Exptl. Theoret. Phys. (U.S.S.R.) **50**, 1393 (1966) [JETP **23**, 924 (1966)].

[47] M. V. Ammosov, N. B. Delone, and V. P. Krainov, Zh. Eksp. Teor. Fiz. **91**, 2008 (1986) [Sov. Phys. JETP **64**, 1191 (1966)].

[48] S. Augst, D. D. Meyerhofer, D. Strickland, and S. L. Chin, J. Opt. Soc. Am. B **8**, 858 (1991).

[49] X. M. Tong and C. D. Lin, J. Phys. B: At., Mol. Opt. Phys. **38**, 2593 (2005).

[50] S. Rae, Opt. Commun. **97**, 25 (1993).

[51] T. Auguste, P. Monot, L.-A. Lompré, G. Mainfray, and C. Manus, Opt. Commun. **89**, 145 (1992).

[52] M. Tipping, J. Machine Learn. Res. **1**, 211 (2001).

[53] Our inclusion of the three-piece field-ionization rate from [37] into the SMILEI PIC code is freely available from https://github.com/iouatu/mySmilei.

[54] https://www.archer2.ac.uk.

[55] K. Atkinson and W. Han, Elementary Numerical Analysis Course, 3rd Ed., https://homepage.math.uiowa.edu/~whan.

[56] T. Chen and C. Guestrin, *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Association for Computing Machinery, New York, 2016), pp. 785–794.

[57] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '19 (Association for Computing Machinery, New York, 2019), pp. 2623–2631.

[58] A. P. Dempster, N. M. Laird, and D. B. Rubin, J. R. Statist. Soc. B **39**, 1 (1977).

[59] C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, New York, 2006).

[60] E. Fong and C. C. Holmes, Biometrika **107**, 489 (2020),.