

## Deep learning for retrieval of the internuclear distance in a molecule from interference patterns in photoelectron momentum distributions

N. I. Shvetsov-Shilovski \* and M. Lein*Institut für Theoretische Physik, Leibniz Universität Hannover, 30167 Hannover, Germany* (Received 18 August 2021; revised 15 December 2021; accepted 24 January 2022; published 7 February 2022)

We use a convolutional neural network to retrieve the internuclear distance in the two-dimensional  $H_2^+$  molecule ionized by a strong few-cycle laser pulse based on the photoelectron momentum distribution. We show that a neural network trained on a relatively small dataset consisting of a few thousand images can predict the internuclear distance with an absolute error less than 0.1 a.u. Deep learning allows us to retrieve more than one parameter from a given momentum distribution. Specifically, we used a convolutional neural network to retrieve both the internuclear distance and the laser intensity. We study the effect of focal averaging, and we find that the convolutional neural network trained using the focal averaged electron momentum distributions also shows a good performance in reconstructing the internuclear distance.

DOI: [10.1103/PhysRevA.105.L021102](https://doi.org/10.1103/PhysRevA.105.L021102)

### I. INTRODUCTION

Development of techniques aimed at visualization of electronic and molecular dynamics in real time will open new horizons in many branches of modern science and technology. Many different techniques for time-resolved molecular imaging have been proposed thus far (see Ref. [1] for a review). The emergence and availability of table-top intense femtosecond laser systems has led to several new time-resolved imaging techniques using the highly nonlinear phenomena originating from interaction of strong laser pulses with atoms and molecules (see, e.g., Refs. [2,3] for review of these phenomena and the whole field of strong-field physics). Examples are laser-assisted electron diffraction [4,5], laser induced Coulomb explosion imaging [6–9], high-order harmonic orbital tomography [10,11], laser-induced electron diffraction (LIED) [12–15], and strong-field photoelectron holography (SFPH) [16]. The two latter methods analyze momentum distributions of electrons from strong-field ionization. The recent experimental achievements in LIED and SFPH (see, e.g., Refs. [17,18]) suggest that future experiments will aim at extracting the information about nuclear motion in a molecule from electron momentum distributions.

The understanding of the outcomes of these forthcoming experiments requires thorough theoretical studies of the effects of nuclear motion on the photoelectron momentum distributions. Such theoretical studies are already on the way. For instance, it was shown in Ref. [17] that the different nuclear wave packet dynamics in hydrogen and deuterium molecules leads to a difference in bond length, which, in turn, transforms into a shift of the holographic fringe at certain electron momenta. Before analyzing the imprints of the nuclear motion in momentum distributions, it is useful to study the distributions for fixed nuclei with varying internuclear distance. In the present paper we address this problem using methods of

machine learning, which is “a subfield of Computer Science wherein machines learn to perform tasks for which they were not explicitly programmed” [19].

Machine learning, and more specifically deep learning (see, e.g., Refs. [20,21] for textbook treatments), has been successfully applied to the prediction of the flux of high-order harmonics for different experimental parameters [22], the prediction of the ground-state energy of an electron in various two-dimensional (2D) confining potentials [23], and the reconstruction of the intensity and the carrier-envelope phase (CEP) of ultrashort laser pulses from 2D images, namely from frequency-resolved optical gating traces [24] and from dispersion scan traces [25]. Recently a deep neural network was also applied to develop an efficient numerical implementation of the trajectory-based Coulomb-corrected strong-field approximation (TCSFA) (see Refs. [26,27] for the foundations of the TCSFA method) [28]. In all these examples the application of machine learning allowed the avoidance of heavy computational costs that would be inevitable when solving these problems using traditional ways.

Very recently the convolutional neural networks (CNN) were used to predict high-order harmonic generation (HHG) spectra for model di- and triatomic molecules for randomly chosen parameters. The latter include laser intensity, internuclear distance, and orientation of the molecule [29]. Furthermore, it was shown in Ref. [29] that the CNN can be used for solving inverse problems: determination of molecular and laser parameters, as well as classification of molecules based on their HHG spectrum (or time-dependent dipole acceleration) alone. These problems are hard to solve by manually inspecting a variety of complex spectra. On the other hand, classification is one of the typical tasks of machine learning. A similar situation is found for the problem of the present work. In this paper we train a CNN to predict the internuclear distance in the 2D model  $H_2^+$  molecule from a given photoelectron momentum distribution (PMD). Application of a CNN is not the only possible way to retrieve the internuclear distance from electron momentum distributions. This problem

\*n79@narod.ru

can be also solved by directly comparing the given PMD with a precalculated set of momentum distributions corresponding to various internuclear distances. However, the direct comparison is not expected to perform well when used on momentum distributions that it has not explicitly trained for. Therefore, we focus on the CNN in this paper. Although the CNN will also face problems when tested on distributions it has not been trained for, it is of interest to address this question quantitatively.

The momentum distributions that are needed to train the CNN are calculated from the direct numerical solution of the time-dependent Schrödinger equation (TDSE). We show that a good accuracy of the predictions can be achieved even for relatively small sets of training data. We then study the effect of the focal averaging on the retrieval of the internuclear distance with the CNN.

## II. MODEL

### A. Solution of time-dependent Schrödinger equation

For the calculations we use a few-cycle linearly polarized laser pulse that is defined in terms of the vector-potential and present between  $t = 0$  and  $t_f = (2\pi/\omega)n_p$ :

$$\vec{A}(t) = (-1)^{n_p} \frac{F_0}{\omega} \sin^2\left(\frac{\omega t}{2n_p}\right) \sin(\omega t + \varphi) \vec{e}_x. \quad (1)$$

Here  $\vec{e}_x$  is a unit vector in the polarization direction ( $x$  axis),  $n_p$  is the number of optical cycles within the pulse, and  $\varphi$  is the CEP. The electric field is to be obtained from Eq. (1) as  $\vec{F}(t) = -d\vec{A}/dt$ . We do our simulations for  $\varphi = 0$ .

In the velocity gauge, the 2D TDSE for an electron interacting with the laser pulse is given by

$$i \frac{\partial}{\partial t} \Psi(x, y, t) = \left\{ -\frac{1}{2} \left( \frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2} \right) - iA_x(t) \frac{\partial}{\partial x} + V(x, y) \right\} \Psi(x, y, t), \quad (2)$$

where  $\Psi(x, y, t)$  is the coordinate-space wave function and  $V(x, y)$  is the soft-core binding potential of the model  $\text{H}_2^+$  molecular ion in the frozen nuclei approximation:

$$V(x, y) = -\frac{1}{\sqrt{(x - R/2)^2 + y^2 + a}} - \frac{1}{\sqrt{(x + R/2)^2 + y^2 + a}}. \quad (3)$$

Here,  $R$  is the internuclear distance and  $a = 0.64$  is the soft-core parameter. We use the Feit-Fleck-Steiger split-operator method [30] to solve the TDSE Eq. (2). The ground-state wave function was obtained by imaginary time propagation. Our computational box was centered at  $(x = 0, y = 0)$  and extends over  $x \in [-400, 400]$  a.u. and  $y \in [-200, 200]$  a.u. We use equal grid spacings for  $x$  and  $y$  coordinates,  $\Delta x = \Delta y = 0.1954$  a.u.

The wave function was propagated from the beginning of the laser pulse  $t = 0$  to  $t = 4t_f$  with the time step  $\Delta t = 0.0184$  a.u. We apply absorbing boundaries to prevent unphysical reflections of the wave packet from the boundary of the computational grid, i.e., at every time step the wave function

is multiplied by the mask:

$$M(x, y) = \begin{cases} 1 & \text{for } r \leq r_b \\ \exp[-\beta(r - r_b)^2] & \text{for } r > r_b \end{cases}. \quad (4)$$

Here,  $r = \sqrt{x^2 + y^2}$ ,  $r_b = 150$  a.u. and  $\beta = 10^{-4}$ . We note that at the intensity of  $4.0 \times 10^{14}$  W/cm<sup>2</sup> and for the wavelength of 800 nm, the characteristic amplitude of the laser-induced electron quiver motion is  $F_0/\omega^2 = 32.8$  a.u. Therefore, the position of the absorbing boundary  $r_b$  exceeds this value by a factor of 4.5. The photoelectron momentum distributions are calculated by using the mask method [12,31].

### B. Architecture of convolutional neural network

The choice of architecture of a neural network should account for the structure of the data used for learning and the desired output. In our case the data used for learning are the pairs consisting of the PMD (image) and the corresponding internuclear distance  $R$  (label). Bearing in mind that the  $\text{H}_2^+$  molecule is ionized by a laser field linearly polarized along the internuclear axis, we assume that every PMD has the aspect ratio 2 : 1. The details of the necessary image preprocessing are given in Sec. III. We train the neural network to solve the regression problem, i.e., to predict the internuclear distance  $R$  from a given PMD.

The deep neural network that we use for the problem at hand consists of five nonreducing convolutional layers, each followed by a reducing average pooling layer. Each of the nonreducing convolutional layers operates with 32 filters with sizes of  $3 \times 3$  pixels. These convolution layers produce new images “feature maps” (see, e.g., Ref. [21]). The number of these new images equals the number of filters. The values of the filter matrices are to be determined through the training process, i.e., they play the same role as the trainable weights of an ordinary artificial neural network. After performing the convolution operation, all the convolutional layers apply the rectified linear unit (ReLU) activation function, which is defined as  $\text{ReLU}(x) = \max(0, x)$ . The average pooling layers divide the images they get into pooling regions with sizes of  $2 \times 2$  pixels and calculate averaged values in every region. Therefore, each average pooling layer reduces the size of the image by a factor of 2. The last average pooling layer is connected to the dropout layer that randomly sets its input elements, i.e., output of the preceding layer, to zero with a certain probability. This probability is chosen to be equal to 0.2. The dropout layer allows us to avoid overfitting. The output of the dropout layer is fed to a fully connected layer that produces only one single value: the internuclear distance  $R$ . This is the output value of the whole neural network.

## III. NUMERICAL EXPERIMENTS AND RESULTS

The photoelectron momentum distributions calculated from the solution of the TDSE (2) for three different internuclear distances are shown in Figs. 1(a), 1(c), and 1(e). It is seen that the shape of the distribution changes considerably with increasing  $R$ . However, the quantification of the corresponding changes in the PMDs is a nontrivial task. This makes application of neural networks particularly appropriate. In order to train the neural network, we first need to produce a

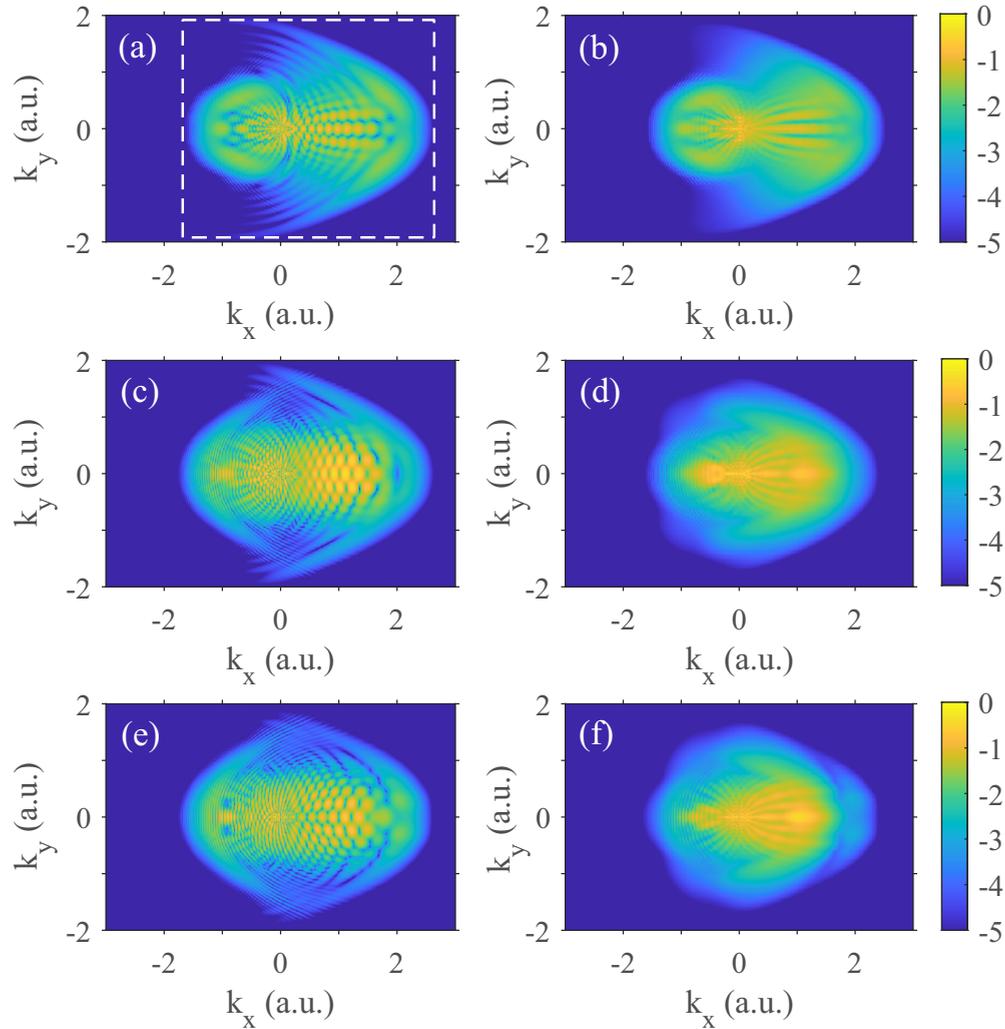


FIG. 1. Electron momentum distributions for ionization of the  $\text{H}_2^+$  molecule by a laser pulse with a duration of  $n_p = 2$  cycles and wavelength of  $\lambda = 800$  nm obtained from numerical solution of the TDSE. (a), (b) correspond to the internuclear distance 2.0 a.u.; (c), (d) correspond to the internuclear distance 5.0 a.u.; (e), (f) correspond to the internuclear distance 6.0 a.u. The left column [(a), (c), and (e)] show the distributions calculated at fixed intensity of  $4.0 \times 10^{14}$  W/cm<sup>2</sup>. The right column [(b), (d), and (f)] displays the distributions averaged over the focal volume for the same peak intensity of  $4.0 \times 10^{14}$  W/cm<sup>2</sup>. The laser field is linearly polarized along the  $x$  axis. The distributions are normalized to the maximum value. Shown is the decimal logarithm of the distribution, see text.

set of training data. To this end, we solve the TDSE, Eq. (2), for  $N$  random internuclear distances  $R_k \in [1.0, 8.0]$  a.u. and peak laser intensities  $I_k \in [1.0, 4.0] \times 10^{14}$  W/cm<sup>2</sup>, where  $k = 1, \dots, N$ , and we calculate the corresponding electron momentum distributions. Since the solution of the 2D TDSE takes a few hours on four to eight modern cores working in parallel, the formation of a large training set is computationally expensive. Here we use  $N = 3000$ . About one week is needed to create such a data set using a computer cluster. We note that our data set is relatively small (compared with  $N = 200000$  and  $N = 30000$  used in Refs. [23] and [29], respectively). Nevertheless, for the problem at hand even such a modest data set allows us to obtain satisfactory results.

The PMD calculated from the solution of the TDSE is a matrix of size  $4096 \times 2048$ . The usage of matrices of such sizes as an input for a convolutional neural network will lead to a slow training process. For this reason, we first modify the matrix of the PMD as follows.

We find the absolute maximum  $\text{PMD}_{\max}$  of the distribution and calculate the decimal logarithm of the normalized PMD:  $W = \log_{10}(\text{PMD}/\text{PMD}_{\max})$ . We set  $W = -5$  for all values that are smaller than  $-5$ . We note that in doing so we consider not only the low-energy part of the distribution created by the electrons that do not experience hard recollisions with their parent ions, but also the beginning of the high-energy part of the PMD. This high-energy part is formed due to electrons that are driven back by the laser field to their parent ions and rescatter from them. Classically, the boundary between low- and high-energy parts of the PMD corresponds to the momentum  $k = 2\sqrt{U_p}$ , where  $U_p$  is the ponderomotive potential. For the parameters of Fig. 1, this estimate yields  $k \approx 1.87$  a.u. Then we find a rectangular area such that the values of  $W$  at the boundary of the rectangle are just above  $-5$ . This rectangle is shown by the dashed lines in Fig. 1(a). The image within the rectangle is resized to  $256 \times 128$  by using bicubic interpolation. Finally, all the elements of the matrix

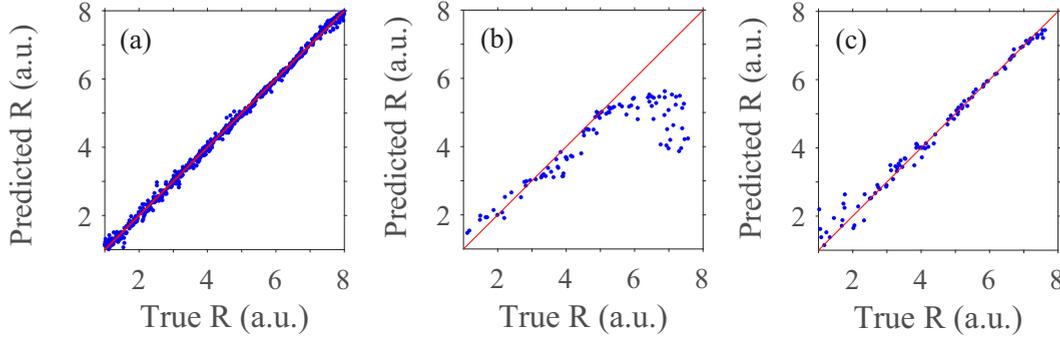


FIG. 2. Plot of predicted vs true internuclear distances illustrating the performance of neural networks. (a) Neural network trained on a set of distributions that were calculated with fixed laser intensities (not focal averaged), see text. (b) The same neural network as in (a), but receiving focal averaged momentum distributions as test images. (c) The neural network trained on a set of focal averaged distributions and tested on another independent set of focal averaged distributions.

are rescaled so that the minimum value corresponds to zero and the maximum one is mapped to 255. The resulting matrix of the size of  $256 \times 128$  is used as an input for the CNN.

We split our data set into training and test sets in the ratio 0.75:0.25. Only data from the training set were used for training of the neural network. The goal of training is the minimization of the loss function, i.e., the measure of deviation (in our case the mean squared error) between predictions of the neural network and expected outcomes for the training set. The MATLAB package [32] is used for the calculations. The training of a neural network is performed on a modern PC using a graphic processing unit and takes only a few minutes. The prediction of the internuclear distance from one single image when the CNN is fully operational takes us about 0.024 sec on a PC. Therefore, the validation procedure for a test set consisting of 750 images requires only 18 sec.

The results of the application of the trained neural network to the test data are presented in Fig. 2(a). It is clearly seen that the neural network can successfully predict the internuclear distance. We characterize the quality of the neural network by the mean absolute error (MAE) between the predicted and true values of  $R$  over the test data set - a measure, which is different from the loss function (mean squared error) used in the training process. The neural network predicts the internuclear distance with the MAE of 0.07 a.u. We have found that another neural network that uses solely the low-energy part of the PMDs ( $0 < k_x < \sqrt{2U_p}$ ,  $|k_y| < \sqrt{U_p}$ ) shows slightly worse results: the corresponding MAE of  $R$  is equal to 0.12 a.u. This implies that the recognition of the internuclear distance with the neural network relies mostly on the interference patterns in the low-energy part of the momentum distributions, i.e., on the holographic patterns [16], but the accuracy can be enhanced by including high-energy electrons.

It is clear that the shape of the PMDs depends not only on  $R$ , but also on the laser parameters, especially on the intensity  $I$ . Deep learning allows us to retrieve more than one parameter from a given PMD. Using the same training data set, we have trained another CNN that is able to retrieve both the internuclear distance and the laser intensity. The MAEs provided by this neural network for  $R$  and  $I$  are equal to 0.07 a.u. and  $0.05 \times 10^{14}$  W/cm<sup>2</sup>, respectively. It is seen that the error for  $R$  coincides with the one obtained using the very first neural network aimed at the retrieval of only the internuclear

distance. We therefore conclude that the ability of the CNN to retrieve both parameters does not affect the accuracy with which the internuclear distance is retrieved.

It is well-known that the intensity fluctuates in an experiment. This raises the question: How vulnerable is the performance of the trained neural network to the effect of focal averaging? To answer this question, we calculate a number of electron momentum distributions averaged over the focal volume and use them to test our neural network trained on the distributions obtained for *fixed* laser intensities. For a peak intensity  $I_0$ , the focal-volume averaged distribution  $dP/d\vec{k}$  can be calculated as [35]

$$\frac{dP}{d^3k} = \int_0^{I_0} \frac{dP(I)}{d^3k} \left( -\frac{\partial V}{\partial I} \right) dI, \quad (5)$$

where  $dP(I)/d^3k$  is the momentum distribution for a fixed intensity  $I$ , and  $(\partial V/\partial I)dI$  is the focal volume element that corresponds to intensities between  $I$  and  $I + dI$ . We assume that the laser beam has Lorentzian spatial distribution of the intensity along the propagation direction and Gaussian intensity profile in the transverse direction (see, e.g., Refs. [3,33,34]). The focal volume element for such a beam is given by [35]:

$$\left( -\frac{\partial V}{\partial I} \right) dI \sim \frac{I_0}{I} \left( \frac{I_0}{I} + 2 \right) \sqrt{\frac{I_0}{I} - 1} dI. \quad (6)$$

Obviously, the calculation of the focal-volume averaged distribution requires a number of TDSE solutions for different intensities  $I < I_0$ , and therefore is computationally demanding. For this reason, we calculate only  $N_a = 100$  focal volume averaged PMDs for random internuclear distances  $R_k \in [1.0, 8.0]$  a.u. and peak intensities  $I_k \in [1.0, 4.0] \times 10^{14}$  W/cm<sup>2</sup> ( $k = 1, \dots, N_a$ ).

Figure 2(b) illustrates the performance of the neural network on this test set. We see that the performance of the CNN for the averaged PMDs is not as good as for the non-averaged PMDs. The MAE on this test set reaches the value of 0.83 a.u., which is still better than the value 1.4 a.u. that we find in an approach based on direct comparison with the training set. However, it is seen that the neural network works relatively well for PMDs that correspond to the internuclear distances less than 5.0 a.u. Indeed, the MAE calculated for the

focal-volume averaged PMDs with  $R < 5.0$  a.u. is equal to 0.24 a.u. The unsatisfactory performance of the CNN for focal averaged PMDs with  $R > 5.0$  a.u. can be understood from a close inspection of Figs. 1(a)–1(f). It is seen that the averaged distributions for  $R = 5.0$  a.u. and  $R = 6.0$  a.u. shown in Figs. 1(d) and 1(f), respectively, are similar to each other, especially in their low-energy parts, i.e., for  $0 < k_x < \sqrt{2U_p}$  and  $|k_y| < \sqrt{U_p}$ . Simultaneously, these distributions are not too similar to their counterparts calculated for fixed laser intensity [cf. Figs. 1(c) and 1(d), as well as Figs. 1(e) and 1(f)]. In contrast to this, the averaged distributions corresponding to smaller values of  $R$  resemble the PMDs for same internuclear distances and fixed intensities [cf. Figs. 1(a) and 1(b)]. All this explains why the CNN trained with the distributions for fixed intensities underestimates large internuclear distances by treating them as  $R \lesssim 5.0$  a.u. Aiming at a CNN that is able to map different PMDs (averaged and unaveraged) to the same  $R$  would indeed imply a non-bijective mapping.

In order to understand whether the internuclear distance can be reliably retrieved from the focal averaged momentum distributions using deep learning, we train another CNN. This second neural network has the same architecture as the first one, but it is trained on a set of averaged PMDs. A set of  $N_a = 100$  distributions is too small to train a neural network. Therefore, the data set should be augmented. To this end, we apply 2D interpolation on an irregular grid (see, e.g., Ref. [36]) in the  $(R, I_0)$  plane formed by the  $N_a$  points. As a result, we produce a set of 6000 focal averaged electron momentum distributions and use them to train our new CNN. In order to have a test set independent of the initial  $N_a$  focal averaged momentum distributions, we produce another  $N_a = 100$  of averaged PMDs for random internuclear distances and peak laser intensities using direct numerical solution of the TDSE and Eq. (5). We find that the CNN trained on the set of intensity averaged PMDs shows a rather good performance, see Fig. 2(c). It is also seen from Fig. 2(c) that the performance of the CNN is slightly worse for small internuclear distances as compared to larger values of  $R$ . We attribute this to the fact that the interpolation accuracy for the focal volume averaged momentum distributions is slightly worse for small internuclear distances. Nevertheless, the MAE on the independent test set is about 0.14 a.u. This result clearly shows that neural networks can be used to retrieve the unknown internuclear distance from a given electron momentum distribution even if the latter is affected by focal averaging.

When applied to a real experimental situation it would be desirable for a neural network to possess some transferability, i.e., to predict correct results even for PMDs obtained at parameters that are beyond the range of the training data. The neural networks trained here show only limited transferability. This is seen from the example of our first neural network trained on nonaveraged PMDs. The application of this neural network to a new set of PMDs obtained for internuclear distances  $8.0 < R \leq 12.0$  a.u., i.e., outside of the training range  $1.0 \leq R \leq 8.0$  a.u., leads to an MAE of 3.0 a.u. If our neural network is applied to images reflected about the vertical axis (corresponding to a change of the CEP by  $\pi$ ), we find an MAE of 1.4 a.u. Slightly better results are achieved in the case where the neural network is applied to distributions obtained for a nonzero angles between the molecular axis and polarization

direction. The corresponding MAE is 0.9 a.u. We note that an approach based on the direct comparison of a given PMD with a precalculated set of the distributions shows worse results in terms of transferability. Here we find MAEs of 5.1, 1.6, and 1.4 a.u., respectively. The transferability problem of the neural network can be solved by the transfer learning technique (see, e.g., Ref. [37] for details). This approach was successfully used in Ref. [29]. The application of this technique to the retrieval of the internuclear distance will be a subject of further studies.

It should also be noted that the application of the neural network to the problem at hand is not without shortcomings. Since any neural network works as a “blackbox”, i.e., it is not clear how the neural network takes its decisions, it is often difficult to assess whether it works properly for a given input image. On the other hand, the so-called visualization methods of deep neural networks that allow to explain the decisions of the CNNs are being actively developed nowadays (see, e.g., Ref. [38] for a review). Overall, there are reasons to believe that CNNs have a high potential for extracting various molecular properties from the electron momentum distributions produced by strong-field ionization.

#### IV. CONCLUSIONS AND OUTLOOK

In conclusion, we have investigated the capabilities of deep learning for retrieval of the internuclear distance in the  $H_2^+$  molecule from a given 2D electron momentum distribution generated by a strong laser pulse. We have shown that the neural network trained on a few thousand images is able to predict the internuclear distance with a MAE less than 0.1 a.u. In addition to this, the CNN can be trained to retrieve more than one parameter from a given PMD. We have used the neural network to predict both the internuclear distance and the intensity of the laser pulse. Furthermore, we have studied the effect of focal averaging on the retrieval of the internuclear distance with a neural network. It is shown that the CNN trained on a set of focal averaged distributions also performs well.

The electron momentum distributions are sensitive not only to intensity fluctuations, but also to the changes of other laser parameters. For short laser pulses the variations of the CEP can change the resulting PMDs significantly. Therefore, the effect of the CEP on the retrieval of the internuclear distance needs to be studied. Moreover, it is of interest to look “inside” the CNN and analyze what features of the holographic structures allow the network to classify the images. This can be done by application of the visualization methods developed for the CNNs. Finally, the transferability of the neural networks designed for the problem at hand should be improved. These questions will be the subject of further studies. Progress in these directions is important for the development of SFPH and for the whole field of time-resolved molecular imaging.

#### ACKNOWLEDGMENTS

We are grateful to S. Brennecke, F. Oppermann, and S. Yue for continued interest in this work and stimulating discussions. This work was supported by the Deutsche Forschungsgemeinschaft (Grant No. SH 1145/1-2).

- [1] J. Xu, C. I. Blaga, P. Agostini, and L. F. DiMauro, Time-resolved molecular imaging, *J. Phys. B* **49**, 112001 (2016).
- [2] M. Kitzler and S. Gräfe (Eds.), *Ultrafast Dynamics Driven by Intense Light Pulses. From Atoms to Solids, From Lasers to Intense X-rays* (Springer, Cham, 2016).
- [3] C. D. Lin, A.-T. Le, C. Jin, and H. Wei, *Attosecond and Strong-Field Physics. Principles and Applications* (Cambridge University Press, Cambridge, 2018).
- [4] R. Kanya, Y. Morimoto, and K. Yamanouchi, Observation of Laser-Assisted Electron-Atom Scattering in Femtosecond Intense Laser Fields, *Phys. Rev. Lett.* **105**, 123202 (2010).
- [5] Y. Morimoto, R. Kanya, and K. Yamanouchi, Laser-assisted electron diffraction for femtosecond molecular imaging, *J. Chem. Phys.* **140**, 064201 (2014).
- [6] L. J. Frasinski, K. Codling, P. Hatherly, J. Barr, I. N. Ross, and W. T. Toner, Femtosecond Dynamics of Multielectron Dissociative Ionization by use of a Picosecond Laser, *Phys. Rev. Lett.* **58**, 2424 (1987).
- [7] C. Cornaggia, J. Lavancier, D. Normand, J. Morellec, P. Agostini, J. P. Chambaret, and A. Antonetti, Multielectron dissociative ionization of diatomic molecules in an intense femtosecond laser field, *Phys. Rev. A* **44**, 4499 (1991).
- [8] J. H. Posthumus, L. J. Frasinski, A. J. Giles, and K. Codling, Dissociative ionization of molecules in intense laser fields: a method of predicting ion kinetic energies and appearance intensities, *J. Phys. B* **28**, L349 (1995).
- [9] C. Cornaggia, M. Schmidt, and D. Normand, Laser-induced nuclear motions in the Coulomb explosion of  $C_2H_2^+$  ions, *Phys. Rev. A* **51**, 1431 (1995).
- [10] J. Itatani, J. Levesque, D. Zeidler, H. Niikura, H. Pépin, J. C. Kieffer, P. B. Corkum, D. M. Villeneuve, Tomographic imaging of molecular orbitals, *Nature (London)* **432**, 867 (2004).
- [11] S. Haessler, J. Caillat, W. Boutu, C. Giovanetti-Teixeira, T. Ruchon, T. Auguste, Z. Diveki, P. Breger, A. Maquet, B. Carrère, R. Taïeb, and P. Salières, Attosecond imaging of molecular electronic wave packets, *Nat. Phys.* **6**, 200 (2010).
- [12] M. Lein, J. P. Marangos, P. L. Knight, Electron diffraction in above-threshold ionization of molecules, *Phys. Rev. A* **66**, 051404(R) (2002).
- [13] M. Meckel, D. Comtois, D. Zeidler, A. Staudte, D. Pavičić, H. C. Bandulet, H. Pépin, J. C. Kieffer, R. Dörner, D. M. Villeneuve, and P. B. Corkum, Laser-Induced electron tunneling and diffraction, *Science* **320**, 1478 (2008).
- [14] C. I. Blaga, J. Xu, A. D. DiChiara, E. Sistrunk, K. Zhang, P. Agostini, T. A. Miller, L. F. DiMauro, and C. D. Lin, Imaging ultrafast molecular dynamics with laser-induced electron diffraction, *Nature (London)* **483**, 194 (2012).
- [15] M. G. Pullen, B. Wolter, A. T. Le, M. Baudisch, M. Hemmer, A. Senftleben, C. D. Schröter, J. Ullrich, R. Moshhammer, C. D. Lin, J. Biegert, Imaging an aligned polyatomic molecule with laser-induced electron diffraction, *Nat. Commun.* **6**, 7262 (2015).
- [16] Y. Huismans, A. Rouzée, A. Gijsbertsen, J. H. Jungmann, A. S. Smolkowska, P. S. W. M. Logman, F. Lépine, C. Cauchy, S. Zamith, T. Marchenko *et al.*, Time-Resolved Holography with Photoelectrons, *Science* **331**, 61 (2011).
- [17] M. Haertelt, X.-B. Bian, M. Spanner, A. Staudte, and P. B. Corkum, Probing Molecular Dynamics by Laser-Induced Backscattering Holography, *Phys. Rev. Lett.* **116**, 133001 (2016).
- [18] S. G. Walt, N. Ram, M. Atala, N. I. Shvetsov-Shilovski, A. von Conta, D. Baykusheva, M. Lein, and H. J. Wörner, Dynamics of valence-shell electrons and nuclei probed by strong-field holography and rescattering, *Nat. Commun.* **8**, 15651 (2017).
- [19] A. W. Trask, *Grokking Deep Learning* (Manning, Shelter Island, NY, 2019).
- [20] S. Raschka, *Python Machine Learning* (Packt, Birmingham, UK, 2016).
- [21] P. Kim, *MATLAB Deep Learning: With Machine Learning, Neural Networks and Artificial Intelligence* (Apress, Seoul, South Korea, 2017).
- [22] A. M. M. Gherman, K. Kováčh, M. V. Cristea, and V. Toşa, Artificial neural network trained to predict high-harmonic flux, *Appl. Sci.* **8**, 2106 (2018).
- [23] K. Mills, M. Spanner, and I. Tamblyn, Deep learning and the Schrödinger equation, *Phys. Rev. A* **96**, 042113 (2017).
- [24] T. Zahavy, A. Dikopoltsev, D. Moss, G. I. Haham, O. Cohen, S. Mannor, and M. Segev, Deep learning reconstruction of ultrashort pulses, *Optica* **5**, 666 (2018).
- [25] S. Kleinert, A. Tajalli, T. Nagy, and U. Morgner, Rapid phase retrieval of ultrashort pulses from dispersion scan traces using deep neural networks, *Opt. Lett.* **44**, 979 (2019).
- [26] T.-M. Yan, S. V. Popruzhenko, M. J. J. Vrakking, and D. Bauer, Low-Energy Structures in Strong Field Ionization Revealed by Quantum Orbits, *Phys. Rev. Lett.* **105**, 253002 (2010).
- [27] T.-M. Yan and D. Bauer, Sub-barrier Coulomb effects on the interference pattern in tunneling-ionization photoelectron spectra, *Phys. Rev. A* **86**, 053403 (2012).
- [28] X. Liu, G. Zhang, J. Li, G. Shi, M. Zhou, B. Huang, Y. Tang, X. Song, and W. Yang, Deep Learning for Feynman's Path Integral in Strong-Field Time-Dependent Dynamics, *Phys. Rev. Lett.* **124**, 113202 (2020).
- [29] M. Lytova, M. Spanner, and I. Tamblyn, Deep learning and high harmonic generation, [arXiv:2012.10328](https://arxiv.org/abs/2012.10328) v2 (2021).
- [30] M. D. Feit, J. A. Fleck, and A. Steiger, Solution of the Schrödinger equation by a spectral method, *J. Comput. Phys.* **47**, 412 (1982).
- [31] X. M. Tong, K. Hino, and N. Toshima, Phase-dependent atomic ionization in few-cycle intense laser pulse, *Phys. Rev. A* **74**, 031405(R) (2006).
- [32] MATLAB, version 9.10.0 (R2021a). The MathWorks Inc., Natick, Massachusetts, 2021.
- [33] A. E. Siegman, *Lasers* (University Science Books, Sausalito, CA, 1986).
- [34] S. Augst, D. D. Meyerhofer, D. Strickland, and S. L. Chin, Laser ionization of noble gases by Coulomb-barrier suppression, *J. Opt. Soc. Am. B* **8**, 858 (1991).
- [35] T. Morishita, Z. Chen, S. Watanabe, and C. D. Lin, Two-dimensional electron momentum spectra of argon ionized by short intense lasers: Comparison of theory with experiment, *Phys. Rev. A* **75**, 023407 (2007).
- [36] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes. The Art of Scientific Computing*, 3rd ed. (Cambridge University Press, Cambridge, 2007).
- [37] I. Goodfellow, Y. Benqio, and A. Courville, *Deep Learning* (The MIT Press, Cambridge, MA, 2016).
- [38] A. Shahroudjed, A survey on understanding, visualizations, and explanation of deep neural networks, [arXiv:2102.01792](https://arxiv.org/abs/2102.01792) (2021).