Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2018 doi:10.20944/preprints201809.0563.v1

Peer-reviewed version available at Appl. Sci. 2018, 8, 2106; doi:10.3390/app8112106

Article

Artificial neural network trained to predict high-harmonic flux

Ana Maria Mihaela Gherman[1,2], Katalin Kovács[1]*, Mircea Vasile Cristea[2], and Valer Toșa[1]

[1]National Institute for R&D of Isotopic and Molecular Technologies, Donat str. 67-103, 400293 Cluj-Napoca, Romania
[2]Chemical Engineering Department, Babes-Bolyai University, Arany Janos str. 11, 400028 Cluj-Napoca, Romania
* Correspondence: kkovacs@itim-cj.ro; Tel.: +40-264-584037

Featured Application: We present a method based on artificial neural networks – which is an efficient prediction procedure – with the main role to provide quick help for attosecond science laboratories, aimed to predict the expected outcome of high-order harmonic generation experiments to be performed with previously unexplored experimental input parameter combinations.

Abstract: In this work we present the results obtained with an artificial neural network (ANN) which we trained to predict the expected output of high-order harmonic generation (HHG) process, while exploring a multi-dimensional parameter space. We argue on the utility and efficiency of the ANN model and demonstrate its ability to predict the outcome of HHG simulations. In this case study we present the results for a loose focusing HHG beamline, where the changing parameters are: the laser pulse energy, gas pressure, gas cell position relative to focus and gas cell length. The physical quantity which we predict here using ANN is directly related to the total harmonic yield in a specified spectral domain (20-40 eV). We discuss the versatility and adaptability of the presented method.

Keywords: high-order harmonic generation, 3D non-adiabatic model, simulation, artificial neural network, prediction

1. Introduction and motivation

High-order harmonic generation (HHG) is a well-established experimental technique for obtaining coherent radiation in the XUV to soft X-ray spectral domain on the attosecond timescale [1,2]. The process of HHG is highly nonlinear and has very low efficiency [3]. Despite the fact that high-harmonic radiation is the result of the highly nonlinear interaction between the atom and the driving laser pulse, this method is reliable and provides coherent secondary sources in the UV–XUV–soft-X-ray spectral domain with a table-top setup. The ELI-ALPS pan-European research infrastructure is also based on HHG sources [4], demonstrating that the increasing research efforts dedicated to this field are fully justified. Given the advantages of exploring HHG in the emerging field of Attosecond Science and Technology, considerable research effort is being dedicated to find ways for increasing high-harmonic efficiency by implementing loose focusing configurations [4,5], inventing different quasi-phase-matching techniques [6–13], multi-color and/or multi-jet configurations [11,14–16], but also to extend the available photon energy to reach the water-window [16–18] region.

Numerical approaches are unavoidable in this research for several reasons: (i) often, new experiments are designed based on predictions resulting from simulations; (ii) interpretation of existing experimental results cannot be done without additional results from models; (iii) when

© 2018 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2018 doi:10.20944/preprints201809.0563.v1

Peer-reviewed version available at Appl. Sci. 2018, 8, 2106; doi:10.3390/app8112106

looking for the best possible outcome within given experimental possibilities, reliable simulation tools can considerably contribute in finding the best solution.

Increasing efficiency and/or finding the optimum photon flux obtainable within a restricted parameter space is crucial. Experimental scan along many tunable parameters is very time- and resource consuming, additionally the experimental uncertainties result in variations from one measurement to the next. Simulation with a complete 3D model seems a more convenient route. However, as we will show in the following section, scanning along multiple parameters, typically 4-5, implies the execution and interpretation of hundreds to thousands of simulation runs. Since the amount of raw data and computational effort is very difficult to be managed, we propose an additional method to extract the needed information from the data pool, and to predict possible outcomes for parameter combinations which were not included in the original simulation sets.

In this respect, artificial neural networks (ANN) are known to be high speed mathematical models that are capable of solving linear and non-linear multivariate regression problems [19,20], being generally used for their excellent prediction capabilities [21,22] but also for their capacity to solve problems such as classification [23,24], pattern recognition [25] or system control [26,27]. Inspired by the human brain, these numerical modelling techniques are able to gain high efficiency and accuracy in the presence of uncertainties due to their ability to learn from experience [28]. Based on these properties and also on their effectiveness in processing large amount of data, the ANNs have applicability in various fields such as medicine [29], engineering [30,31], science technology [32], nanotechnology [33,34] or physics [35].

In this particular paper the physical quantity we want to predict is the total harmonic flux in a predefined spectral region of the high-harmonic spectrum, which we define as yield:

$$Y = \int_{\omega_1}^{\omega_2} \int_0^R [E_h(\omega, r)]^2 \, r dr d\omega, \qquad (1)$$

where $[\omega_1; \omega_2]$ is the spectral domain of interest, R is the radial extent of the interaction region, $E_h(\omega,r)$ is the harmonic field in spectral representation.

Alternatively, one can identify different physical quantities which we want to predict using ANN, for example the existence and bandwidth / duration of isolated attosecond pulses. One can search to maximize the brightness of a single harmonic as function of different experimental parameters, as it has already been done experimentally [36].

We propose here a combined method to predict the photon flux in the 20-40 eV spectral range while varying four experimental parameters: laser pulse energy, gas pressure, gas cell position and gas cell length. For this prediction an artificial neural network is designed, trained and tested. The input data are the results obtained with the 3D non-adiabatic model (see next section) from a 4-dimensional parameter scan [37].

The paper has the following structure: in section 2.1 we briefly describe the 3D non-adiabatic model for laser pulse propagation and high-harmonic generation, while in section 2.2 the artificial neural network method is presented. In section 3 we present and discuss the results obtained with the ANN, we test its robustness and reliability. Possible applications are suggested as well, since the major role of this prediction procedure is to provide a quick help for HHG laboratories: predict the expected outcome of HHG experiments to be performed with previously unexplored experimental parameter combinations.

2. Theoretical models and numerical methods for high-harmonic generation and artificial neural networks

2.1. High-harmonic generation

In order to have sufficient input data for the training of the ANN, extensive simulation of the complete HHG process has been performed using a 3D non-adiabatic model [38] while the theoretical basis implemented in the simulation code has been described in [39]. According to the physical reality, this model has three main parts: (A) solves the propagation equation of the fundamental laser pulse in the generating medium, which is a noble gas in our case; (B) quantify the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2018 doi:10.20944/preprints201809.0563.v1

Peer-reviewed version available at Appl. Sci. 2018, 8, 2106; doi:10.3390/app8112106

response of the individual atoms to the external laser electric field; and (C) solves the propagation equation of the dipole radiation generated in step (B) within the same medium.

(A) The propagation equation for the fundamental laser pulse is the Maxwell wave equation. This is solved with a self-consistent iterative method in frequency domain and using the paraxial approximation. The method is implemented in a way that the full electric field (carrier and envelope) is calculated, we do not use the slowly varying envelope approximation. The interaction configuration is considered to have cylindrical symmetry being z the propagation direction, and r the radial coordinate. The pulse experiences an effective refractive index – function of time and spatial coordinates – which accounts for the linear refraction and absorption, the intensity-dependent optical Kerr effect and the dispersion due to the free electrons. In this way the modifications of the generating pulse during propagation are included: self-phase modulation, pulse lengthening, intensity clamping, possible self-guiding. The propagation part of the simulation code is partially parallelized. For detailed mathematical formulation we refer to [38].

(B) The elemental interaction of a laser pulse with an atom is treated within the strong-field approximation and the generated dipole is calculated by solving the Lewenstein integral [40]. This part of the simulation code has been also parallelized, but it is still the most demanding part of the calculation in terms of CPU time and memory use.

(C) The total harmonic field results from the propagation and coherent addition of the atomic polarizations calculated in step (B). The wave equation describing the propagation of the harmonic field is similar to the wave equation for the fundamental, but the source term here is the non-linear polarization calculated in step (B).

This simulation code has several advantages which make it a very useful tool in the design and interpretation of HHG experiments. When we construct one specific case all the needed input data are experimentally measurable (or can be estimated). These are information about the fundamental laser pulse like beam type, pulse energy, duration, CEP, chirp; geometrical arrangement like aperture, focusing, cell position and length, far field detection configuration; interaction medium like gas type and pressure profile.

One typical full calculation yields a complex set of quantities in the grid points built over the interaction domain such as: laser pulse shape before and after propagation in time and frequency representation, peak intensity and ionization fraction, harmonic spectrum in near- and far-field (spectral domain), harmonic bursts in temporal domain, spatial build-up of selected harmonic orders. The most important outcome is the harmonic power spectrum which can be directly compared to the experimentally measured spectra. In addition, the simulation results provide a detailed insight into the physics of the whole process (laser and harmonic field propagation), at sub-optical-cycle temporal resolution, while phase-matching conditions can be examined as well.

The model is flexible in the sense that it can be modified according to diverse experimental conditions. It has been extended to be able to treat multiple electric fields (two-color [41–47] and three- or four-color field coming from a light-field synthesizer [48]) or multiple jets configurations. At very high laser intensities the atoms become highly charged ions, therefore we implemented also the multiple ionization scenario [49, 50]. Propagation in a fiber is essentially different from free space propagation, therefore this scenario has also been recently introduced. In attosecond laboratories very diverse experimental configurations are tried, for example Airy beams obtained with drilled mirror – also implemented in one version of the code.

The theoretical model implemented in the complete 3D simulation code is suitable for exploring a multi-dimensional parameter space and searching for optimum based on a given criterion. We can search for the maximum harmonic yield in a given spectral domain, or can look for the existence of a single attosecond pulse, or maximum cutoff energy within given parameter ranges. Although, such a scan along multiple parameters is possible to be performed also experimentally, this process would require much more resources than usually affordable, therefore it is recommended to perform a scan with “in silica” experiments.

Depending on the complexity of the possible experimental situation to be simulated, one whole calculation typically takes between few hours to 1-2 days. This estimation is done based on the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2018 doi:10.20944/preprints201809.0563.v1

Peer-reviewed version available at Appl. Sci. 2018, 8, 2106; doi:10.3390/app8112106

typical cases implemented so far on a cluster IBM System iDataPlex M4 dx360 (in Data Center of INCDTIM Cluj-Napoca).

When we need to explore an unknown parameter space, and we have to scan along several experimental parameters, the number of cases to be run is on the order of hundreds to thousands. Even though, one can afford only sparse sampling along each parameter. Therefore, if we want to search for an optimum output, for example harmonic yield, this sparse sampling is far not enough and cannot guarantee that an optimum configuration can be found.

2.2. Artificial neural networks

In an ANN, neurons are the ones responsible for processing the information. Each computing unit receives a weighted signal from all the previous neighboring neurons, sums the received signals and its own bias, processes the information via a transfer function and generates the output [22]. The neurons are organized in input, output and one or several hidden layers. As general feature, for prediction purposes, the connections between the neurons that belong to the same layer are not allowed, the information being directed from the input towards the output layer [51]. This kind of ANN is named feed-forward neural network and, in most cases, it is associated with the back-propagation learning algorithm which modifies the weights and biases starting from the output towards the input layer in order to minimize the network error [21].

Regarding the HHG process, it is known that the harmonic yield is non-linearly influenced by several parameters such as the fundamental laser pulse characteristics, focusing geometry, gas type and pressure which characterize the medium where HHG takes place. For this particular study we chose as varying parameters the laser pulse energy (E), gas pressure (p), gas cell position (z_i) and gas cell length (L_{med}). Consequently, these four process parameters are chosen to be the ANN inputs, whereas the harmonic yield is the model output. Based on the 3D non-adiabatic model, a total number of 453 input-output data sets were obtained and, subsequently, were divided in: i) training data which represent 77.9% of the total amount of input-output data sets, ii) validation data (15.7%) and iii) testing data (6.4%). In order to improve the accuracy of the model, all input data are mapped to the [-1,1] range.

As a result of the complexity of the HHG process, the multilayer feed-forward network is structured in one input layer, two hidden layers and one output layer, the appropriate number of neurons in the hidden layers being found by systematically testing the ANN performance. The information from the hidden and output layers is processed with two different transfer functions: the hidden layer uses a tansigmoid function, whereas a linear transfer function is applied for the output layer. Furthermore, the network biases and weights are computed using the backpropagation algorithm, and the ANN is trained with the Levenberg-Marquardt algorithm.

Several networks are trained and the best is chosen such that the ANN performance is maximized. The performance of the model is statistically evaluated using the Pearson correlation factor (R) and the relative errors (RE). Consequently, the best ANN has the highest correlation factor and the lowest relative errors.

$$R = \frac{\sum_{i=1}^{n}\left(X_{s_i} - \overline{X_s}\right)\left(X_{ANN_i} - \overline{X_{ANN}}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_{s_i} - \overline{X_s}\right)^2}\sqrt{\sum_{i=1}^{n}\left(X_{ANN_i} - \overline{X_{ANN}}\right)^2}} \quad (2)$$

where X_{si} is the simulated output value, X_{ANNi} is the model predicted output value, $\overline{X_S}$ and $\overline{X_{ANN}}$ are the arithmetic means of the experimental and respectively, of the predicted output values.

$$RE = 100\frac{X_{ANN_i} - X_{s_i}}{X_{s_i}} \quad (3)$$

3. Results and discussion

The main goal of this research is to explore a possible loose focusing HHG configuration and find a combination of input experimental parameters which enable the generation of XUV attosecond pulses at high flux. We have built up one scenario which is experimentally feasible in

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2018 doi:10.20944/preprints201809.0563.v1

Peer-reviewed version available at Appl. Sci. 2018, 8, 2106; doi:10.3390/app8112106

currently operating large attosecond science facilities, for example the ELI-ALPS. All input parameters are experimentally measurable, from which we fix all but 4 parameters, these are the scanned parameters.

Fixed input data are: pulse duration 10 fs, central wavelength 800 nm, focal length 21 m, radius ($1/e^2$ intensity) of the beam at the focusing element is 30 mm, total length of the interaction region is 20 cm, gas medium is argon. The scanned parameters with the values chosen for the full 3D calculations are the following: input laser pulse energy E (0.77; 3; 6; 9; 11.6) mJ; argon pressure p (0.1; 0.38; 1.41; 5.31; 20) mbar; cell entrance position relative to the nominal focus z_i (-50; -25; -20; -10; 0; 10; 25; 50) cm, medium length L_{med} (4; 8; 12; 16; 20) cm. Although the parameters are very sparsely spaced, scanning along all these parameters and performing complete 3D simulations, while also to individually analyze and interpret the results means an enormous data pool and working time and effort, which would be very inefficient. In Figure 1 we represent in a very schematic manner the cases for which the complete 3D simulations have been performed, checked, validated and analyzed.

The unprecedented power of the ANN comes in exactly at this stage of the research. (1) We set the ranges of all the moving parameters; (2) we perform – in a carefully chosen way – a sufficient number of complete simulations, such as to have representative points from every part of the multi-dimensional parameter space; (3) use these simulation results as inputs for the ANN; (4) train, test and validate the ANN as it was presented in section 2.2. The expected results are the estimated high-harmonic yields in the whole multi-dimensional parameter space with arbitrarily fine resolution along each parameter. We emphasize that the ANN method in the present stage of implementation is not yet a parameter optimization method, but an efficient prediction method. However, the trained ANN is also suitable to search for optimum, and the maximum yield within this 4-dimensional parameter space can be found by the user.

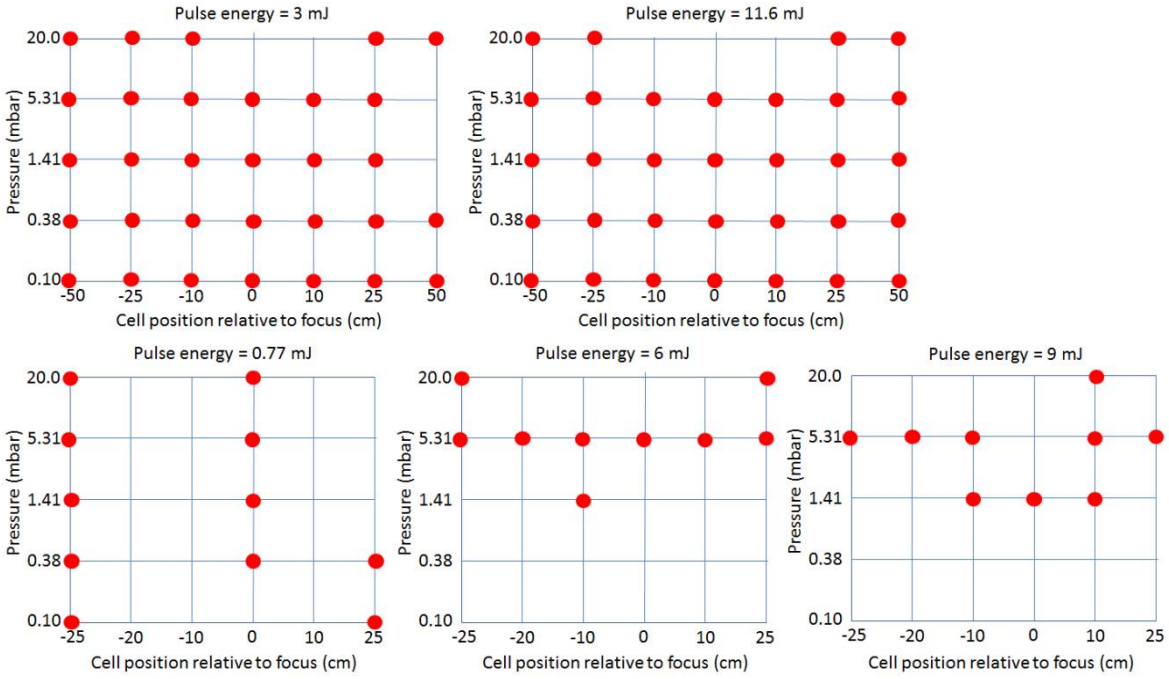

Figure 1. Schematic representation of the parameter space where the full 3D calculations have been performed. For each pulse energy full 3D calculations were performed for several cell positions relative to focus (x axis) and gas pressures (y axis). Red bullets represent the particular cases solved by full 3D simulations and used by ANN for training, validation and testing. The forth scanned parameter is the medium length which is included in every simulation with values 4, 8, 12, 16, 20 cm.

The results obtained with the complete 3D code are discussed in detail elsewhere [37]. There the main result was that the highest harmonic yield around 40 eV can be obtained with the parameter combination: E = 11.6 mJ; p = 5.31 mbar; z_i = -25 cm; L_{med} = 16 cm. In this study we also test this result,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2018 doi:10.20944/preprints201809.0563.v1

Peer-reviewed version available at Appl. Sci. 2018, 8, 2106; doi:10.3390/app8112106

and present other parameter combinations which guarantee high yield comparable to the one found with the multi-parameter scan. We also propose experimentally feasible configurations for existing attosecond science laboratories.

3.1. Testing the ANN against the full 3D simulation results

The optimum topology of the ANN is found to have 11 neurons in the first hidden layer and 17 neurons in the second hidden layer. For this best trained ANN the Pearson coefficient and mean relative errors are calculated for the training and validation data sets and its prediction capabilities are tested considering a set of data that were not yet seen by the model (i.e. the testing data set). The excellent correlation between the experimental results (from 3D simulations, i.e. in silica experiments) and the predicted values is demonstrated by the regression plots and the high Pearson coefficients (Figure 2).

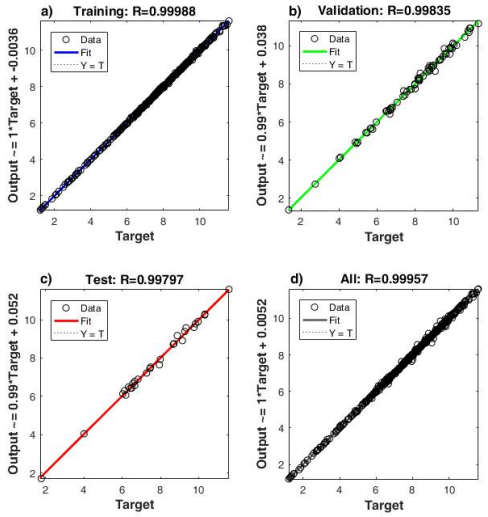

Figure 2. Plot of ANN calculated $\log_{10}$(harmonic yield) against the values obtained with the complete 3D model for the following data sets: (a) training, (b) validation, (c) testing and (d) overall data.

From Figure 2 one can observe that the correlation factor has values greater than 0.99 for all data sets: 0.9998 for the training data set, 0.9983 for the validation data set, 0.9979 for the testing data and 0.9996 for the overall data set. As a consequence, the best linear fit of the points and the dotted line that represents the perfect correlation (R=1) overlap, indicating an extremely reduced deviance of the ANN predictions from the 3D model data. The high performance of the ANN is also demonstrated by the low relative errors which are calculated for the three data sets and are presented in Table 1.

Table.1. The mean and maximum values of the relative error

	|Mean Relative Error| [%]	**|Maximum Relative Error| [%]**
Training Data Set	0.33	1.80
Validation Data Set	1.04	4.48
Testing Data Set	1.41	3.94

The maximum value of the relative error is 4.48% and is obtained for the validation data set. The testing data set has a similar maximum relative error of 3.94 %, whereas the training data set has a value of 1.8%. In case of the mean relative errors, the testing data set has the highest relative errors among the three data sets, with a fairly reduced value of 1.41%. This is expected, as the input testing set of data was processed by the ANN only after the training-validation step, in order to prove its prediction capacity for the newly met testing data.

As a result of the excellent ANN prediction capabilities (extremely low relative errors and high correlation factors), the model was further used to predict the harmonic yield for various input

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2018 doi:10.20944/preprints201809.0563.v1

Peer-reviewed version available at Appl. Sci. 2018, 8, 2106; doi:10.3390/app8112106

variables. In further applications it is very important to have sufficient flux in a given spectral domain, therefore increasing the harmonic yield is of great interest.

We here show that the correctly trained ANN predicts the expected harmonic yield of a possible HHG experiment with the parameters changing in the given range. The ANN works best if the new parameters take values in the range of those used for training and validation. At first glance it may look as a typical interpolation in the multivariable space, however if we consider the complexity of the HHG process given by the highly nonlinear nature of the laser-atom interaction, entangled in the macroscopic effects like propagation, phase matching, etc., then we can understand that the potential of a good prediction method goes far beyond the capabilities of an interpolation.

3.2. Prediction potential of the ANN

The quantity we propose to predict as the main outcome of a virtual HHG experiment is the logarithm of the harmonic yield (Eq. (1)) in the specified spectral domain, namely in the 20-40 eV region (between harmonic orders H13 – H27). We have chosen this quantity because it can be directly related to the total photon flux in the specified spectral domain which is of central interest in any attosecond science experiment.

One very important parameter in a HHG experiment is the optical density of the generation medium. In the usual cases the length of the interaction medium is constant, thus the optical density is proportional to the gas pressure. However, since we have as scanning parameter the medium length as well, we here use p x L_{med} as parameter, a quantity which is proportional to the optical density. In the absorption limited HHG regime the harmonic yield (i.e. the total photon number) should grow with the square of the optical density [52]. In order to have better insight into the physics of how the harmonic yield evolves with the optical density of the medium, we show the results as function of the $\log_{10}(p[\text{mbar}] \times L_{med}[\text{cm}])$.

In Figure 3 we show the prediction capability of the ANN. Dots represent full 3D calculations, while the "+" symbols are the ANN predicted yields for two intermediate pulse energies, namely 6 mJ and 9 mJ. In Figure 3 we represent the expected evolution of the harmonic yield as a function of the gas cell position relative to focus for different laser pulse energy values, while the gas pressure and medium length are fixed: (a) p = 0.38 mbar and L_{med} = 20 cm; (b) p = 5.31 mbar and L_{med} = 20 cm. Comparing the data in (a) and (b), we expect that the ratio of the yields for the same pulse energy should be close to the ratio of the square of the pressures (around 200) which is clearly not the case. This is because the cell position is a critical factor in every HHG process, influencing the yield via propagation of the fields, phase-matching effects and absorption of the harmonics.

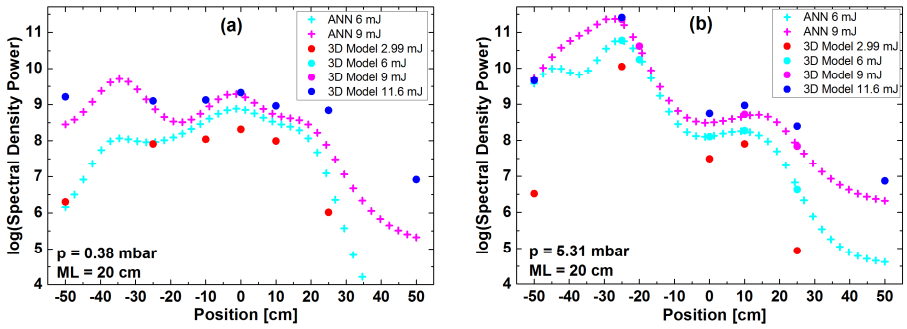

Figure 3. Harmonic yield as function of the gas cell position relative to the nominal focus. Dots are data obtained from the complete 3D simulations, "+" symbols are the results of the ANN predictions. Panel (a) synthesizes the results obtained for different laser pulse energies at the fixed parameter values p = 0.38 mbar and L_{med} = 20 cm, i.e. at the cell end. Panel (b) shows the results obtained with p = 5.31 mbar and L_{med} = 20 cm. Colors of the symbols dots and "+" are the same for the same energy. At the intermediate energies, 6 and 9 mJ, we have performed merely a few but computationally demanding 3D simulations and only with p = 5.31 mbar.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2018 doi:10.20944/preprints201809.0563.v1

Peer-reviewed version available at Appl. Sci. 2018, 8, 2106; doi:10.3390/app8112106

The prediction potential of the ANN simulation proves to be excellent, and captures the highly nonlinear dependence of the macroscopic harmonic yield on many parameters. Especially important conclusion can be drawn from the position-dependence of the yield, namely that very high input pulse energy is not in all cases necessary in order to obtain high harmonic efficiency. Harmonic yield predicted with E = 9 mJ pulse is almost as high as those obtained from 3D simulations with E = 11.6 mJ pulse energy. This result can be good news for attosecond science infrastructures where the other parameters – including cell position and length – are tunable, but the available laser power is limited.

Particularly, the ANN simulation was able to reproduce the transition between two different regimes of laser pulse propagation and consequently different conditions for HHG. At low pulse energy (i.e. starting intensity) and also low pressure conditions we find two comparable maxima of conversion efficiency – see data in Figure 3(a) with the cell positions before and after the focus. In these mild macroscopic conditions the fundamental pulse does not suffer serious distortions during propagation, therefore we find a confirmation of the classical phase-matching theory [53,54]: in the converging beam harmonics find favorable phase-matching conditions off-axis; while in the diverging beam collinear phase-matching on-axis helps the constructive build-up of the generated harmonic field. The situation is different at higher pressure as seen in Figure 3(b). The maximum conversion efficiency found in the whole explored multi-dimensional parameter space has been found at E = 11.6 mJ, p = 5.31 mbar, z_i = -25 cm, L_{med} = 16 cm. The maximum harmonic yield in Figure 3(b) (p = 5.31 mbar) is two orders of magnitude higher than the maximum yield in Figure 3(a) (p = 0.38 mbar). The physical reason for this behavior is that the conditions for the formation of a self-guided beam propagation are fulfilled. A stable "working intensity" is maintained in a large interaction volume (both axially and radially) which creates good phase-matching conditions for a range of harmonics [55].

The ANN has no knowledge about the physics of HHG, however it has been trained with data that contain synthetically the whole nonlinear process. It is highly recommended to perform the complete 3D simulation in cases when ANN predicts less intuitive output. In Figure 4 we present the main results obtained with the full 3D model for the parameter combination (9 mJ; 5.31 mbar; -25 cm; 20 cm). The power spectrum in Figure 4(a) confirms that the highest yield is expected towards the end of the interaction region. In Figure 4(b) we show the spatial (r,z) evolution of the laser pulse's peak intensity, which confirms the formation of a self-guided propagation regime with stable peak intensity 2.5-3·10^{14} W/cm^2. Figure 4(c) shows the spatial (r,z) evolution of the harmonic order H25. In accordance with the power spectrum, it is confirmed that H25 builds up constructively and attains maximum yield towards the end of the interaction region, predominantly off-axis. In correlation with the evolution of the fundamental pulse (Figure 4(b)) the spatial region of maximum H25 intensity is the same where the off-axis refocalization of the fundamental beam happens.

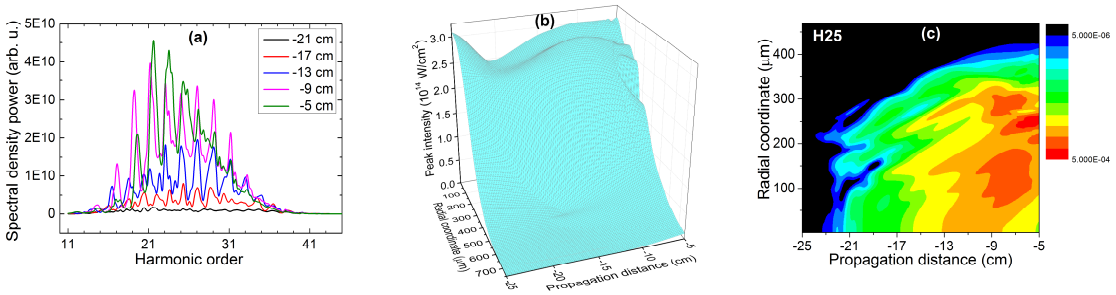

Figure 4. Results of the complete 3D simulation performed with the following values of the parameters: E = 9mJ; p = 5.31 mbar; z_i = -25 cm; L_{med} = 20 cm. Other parameters are specified in the main text at the beginning of Section 3.

In the absorption limited HHG regime the harmonic yield (i.e. the output photon number) should grow quadratically with the optical density [52]. In Figure 5 we represent the harmonic yield as function of the optical density of the medium. Symbols are data coming from the full 3D simulations, lines are the results predicted by the ANN. With dashed line we show as guide to the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2018 doi:10.20944/preprints201809.0563.v1

Peer-reviewed version available at Appl. Sci. 2018, 8, 2106; doi:10.3390/app8112106

eye the line with slope = 2. In Figures 5(a) and 5(b) the 3D simulations provide sufficient data and they confirm that under low optical density conditions the quadratic growth of the yield is valid. Data in Figure 5(c) have been obtained for an intermediate pulse energy E = 9 mJ and the cell positioned in the diverging beam at z_i = 10 cm. In this parameter combination there are few 3D simulation results available, being a good opportunity to test the accuracy of the ANN against the basic physics of the HHG process. Figure 5(c) shows that the ANN works correctly in this respect, predicting quadratic growth of the yield in the low density regime. Moreover, the end of the region for the absorption limited HHG is also reproduced in a similar manner by the ANN as by the complete 3D simulations. It is not possible to show all results, but the main conclusion is that the ANN is capable to predict regions in the multi-dimensional parameter space where the absorption limited HHG can take place. Furthermore, ANN also predicts for which parameter combinations there is a transition from the initially absorption limited HHG to a saturation, then to a drop in obtainable harmonic yield. This dynamics is represented for example in Figures 5(b) and 5(c) at pressure values 0.38 mbar (green symbols and lines) and 1.41 mbar (blue symbols and lines).

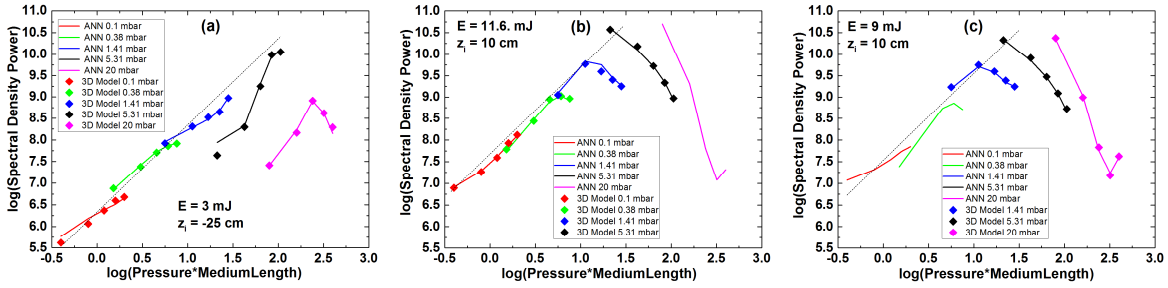

Figure 5. Harmonic yield as function of the optical density of the medium. We represent the logarithm of both quantities, the units are arbitrary. Slope = 2 indicates the quadratic growth of the harmonic yield with the optical density.

4. Conclusions

In this work we presented preliminary results for the prediction of the harmonic yield in different HHG configurations. The results have been obtained with the combination of a complete 3D simulation model and an artificial neural network model. The main result of this work is that we succeeded to build a complex computational tool which is able to predict the expected outcome of a HHG experiment covering conditions when several experimental parameters are varied in a given range. Since this is a usual situation in most attosecond science laboratories, we consider our achievement being a useful and very practical numerical instrument. We need to perform sufficient number of full 3D simulations designed such as to cover (if possible uniformly) the whole explored parameter space. The next step is to introduce these results as input into the ANN, train it, validate and test. A correctly trained ANN is capable to predict the expected outcome of a possible HHG experiment to be performed with parameter combinations that were not implemented in the 3D simulations.

In general terms: the 3D model for HHG combined with the ANN model provide a smooth scan of the parameter space along any variable parameter, and can indicate combinations of possible experimental circumstances which are favorable for the targeted outcome.

In particular: within the present study we targeted to obtain enhanced harmonic yield in the 20-40 eV spectral domain. The variable experimental parameters were: input laser pulse energy E = [0.77; 11.6] mJ; argon pressure p = [0.1; 20] mbar; cell entrance position relative to the nominal focus z_i = [-50; 50] cm, medium length L_{med} = [0; 20] cm, while we kept unchanged the pulse duration, wavelength, focal length, beam size, gas type.

The main message of the obtained results is that a correctly trained ANN fed with reliable input data originating from the full 3D calculations is capable to predict the outcome of a HHG experiment in the unexplored parts of the multi-dimensional parameter space.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2018 doi:10.20944/preprints201809.0563.v1

Peer-reviewed version available at Appl. Sci. 2018, 8, 2106; doi:10.3390/app8112106

In this particular case we only demonstrated the capabilities of the prediction method which uses as input the results of full 3D simulations taking as the target quantity the harmonic yield. However, other physical quantities of interest can be defined and predicted if the ANN is modified accordingly. We can define for example a spectral domain where we search for the existence of isolated attosecond pulse; if it has satellites, then we can predict the contrast relative to the main pulse; we can look for the configuration which provides highest brightness for a specified harmonic order, i.e. high yield combined with low divergence.

ANN as a numerical instrument for prediction can use as input data either data originating from complex simulations (like in our case), or real experimental data (if available), or even a combination of experimental and simulation results. Obviously in the latter case one has to check for the agreement between them, and to perform a calibration of the simulated against the experimental data.

We foresee that this prediction method can be extended to become an easy to use optimization tool, which would mean an essential progress in the way to solve the problem of low conversion efficiency in HHG.

Author Contributions: Conceptualization, V. Tosa; Software, V. Tosa (3 D model), M. V. Cristea (ANN model).; Validation, K. Kovacs and A.M.M. Gherman; Investigation, K. Kovacs and A.M.M. Gherman; Running and analysis of the 3D simulations, K. Kovacs; Running and analysis of the ANN simulations, A.M.M. Gherman; Writing-Original Draft Preparation, K. Kovacs and A.M.M. Gherman; Writing-Review & Editing, K. Kovacs, A.M.M. Gherman and V. Tosa; Visualization, K. Kovacs and A.M.M. Gherman; All authors discussed the results and contributed to the final manuscript. K. Kovacs and A.M.M. Gherman contributed equally to this work.

Funding: This research was funded by the Romanian Ministry of Research and Innovation (MCI) Core Program project no. PN 18 03 02/01.

Acknowledgments: We acknowledge the Data Center of INCDTIM Cluj-Napoca where the 3D simulations have been performed.

Conflicts of Interest: The authors declare no conflict of interest.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2018 doi:10.20944/preprints201809.0563.v1

Peer-reviewed version available at Appl. Sci. 2018, 8, 2106; doi:10.3390/app8112106

References

1. Krausz, F.; Ivanov, M. Attosecond physics. *Rev. Mod. Phys.* **2009**, *81*, 163–234, doi:10.1103/RevModPhys.81.163.
2. Calegari, F.; Sansone, G.; Stagira, S.; Vozzi, C.; Nisoli, M. Advances in attosecond science. *J. Phys. B At. Mol. Opt. Phys.* **2016**, *49*, 062001, doi:10.1088/0953-4075/49/6/062001.
3. Tate, J.; Auguste, T.; Muller, H. G.; Salières, P.; Agostini, P.; Dimauro, L. F. Scaling of wave-packet dynamics in an intense midinfrared Field. *Phys. Rev. Lett.* **2007**, *98*, 1–4, doi:10.1103/PhysRevLett.98.013901.
4. Kühn, S.; Dumergue, M.; Kahaly, S.; Mondal, S.; Füle, M.; Csizmadia, T.; Farkas, B.; Major, B.; Várallyay, Z.; Calegari, F.; Devetta, M.; Frassetto, F.; Mansson, E.; Poletto, L.; Stagira, S.; Vozzi, C.; Nisoli, M.; Rudawski, P.; Maclot, S.; Campi, F.; Wikmark, H.; Arnold, C. L.; Heyl, C. M.; Johnsson, P.; L'Huillier, A.; Lopez-Martens, R.; Haessler, S.; Bocoum, M.; Boehle, F.; Vernier, A.; Iaquaniello, G.; Skantzakis, E.; Papadakis, N.; Kalpouzos, C.; Tzallas, P.; Lépine, F.; Charalambidis, D.; Varjú, K.; Osvay, K.; Sansone, G. The ELI-ALPS facility: the next generation of attosecond sources. *J. Phys. B At. Mol. Opt. Phys.* **2017**, *50*, 132002, doi:10.1088/1361-6455/aa6ee8.
5. Rudawski, P.; Heyl, C. M.; Brizuela, F.; Schwenke, J.; Persson, A.; Mansten, A.; Rakowski, R.; Rading, L.; Campi, F.; Kim, B.; Johnsson, P.; L'Huillier, A. A high-flux high-order harmonic source. *Rev. Sci. Instruments* **2013**, *84*, 073103, doi:10.1063/1.4812266.
6. Gibson, E. A.; Paul, A.; Wagner, N.; Tobey, R.; Gaudiosi, D.; Backus, S.; Christov, I. P.; Aquila, A.; Gullikson, E. M.; Attwood, D. T.; Murnane, M. M.; Kapteyn, H. C. Coherent Soft X-ray Generation in the Water Window with Quasi – Phase Matching. *Science (80-.).* **2003**, *302*, 95–98, doi:10.1126/science.1088654.
7. Sidorenko, P.; Kozlov, M.; Bahabad, A.; Popmintchev, T.; Murnane, M.; Kapteyn, H.; Cohen, O. Sawtooth grating-assisted phase-matching. *Opt. Express* **2010**, *18*, 22686–22692, doi:10.1364/OE.18.022686.
8. Serrat, C.; Biegert, J. All-regions tunable high harmonic enhancement by a periodic static electric field. *Phys. Rev. Lett.* **2010**, *104*, 1–4, doi:10.1103/PhysRevLett.104.073901.
9. Biegert, J.; Austin, D. R. Attosecond pulse shaping using partial phase matching. *New J. Phys.* **2014**, *16*, 113011, doi:10.1088/1367-2630/16/11/113011.
10. O'Keeffe, K.; Lloyd, D. T.; Hooker, S. M. Quasi-phase-matched high-order harmonic generation using tunable pulse trains. *Opt. Express* **2014**, *22*, 7722–7732, doi:10.1364/OE.22.007722.
11. Ganeev, R. A.; Suzuki, M.; Kuroda, H. Quasi-phase-matching of high-order harmonics in multiple plasma jets. *Phys. Rev. A - At. Mol. Opt. Phys.* **2014**, *89*, 2–7, doi:10.1103/PhysRevA.89.033821.
12. Ganeev, R. A.; Toşa, V.; Kovács, K.; Suzuki, M.; Yoneya, S.; Kuroda, H. Influence of ablated and tunneled electrons on quasi-phase-matched high-order-harmonic generation in laser-produced plasma. *Phys. Rev. A - At. Mol. Opt. Phys.* **2015**, *91*, 1–8, doi:10.1103/PhysRevA.91.043823.
13. Strelkov, V. V.; Ganeev, R. A. Quasi-phase-matching of high-order harmonics in plasma plumes: theory and experiment. *Opt. Express* **2017**, *25*, 21068–21083, doi:10.1364/OE.25.021068.
14. Cormier, E.; Lewenstein, M. Optimizing the effciency in high order harmonic generation

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2018 doi:10.20944/preprints201809.0563.v1

Peer-reviewed version available at Appl. Sci. 2018, 8, 2106; doi:10.3390/app8112106

optimization by two-color fields. *Eur. Phys. J. D* **2000**, *12*, 227.

15. Schütte, B.; Weber, P.; Kovács, K.; Balogh, E.; Major, B.; Tosa, V.; Han, S.; Vrakking, M. J. J.; Varjú, K.; Rouzée, A. Bright attosecond soft X-ray pulse trains by transient phase-matching in two-color high-order harmonic generation. *Opt. Express* **2015**, *23*, 33947, doi:10.1364/OE.23.033947.
16. Luo, J.; Hong, W.; Zhang, Q.; Liu, K.; Lu, P. Dramatic cutoff extension and broadband supercontinuum generation in multi-cycle two color pulses. *Opt. Express* **2012**, *20*, 9801, doi:10.1364/OE.20.009801.
17. Chipperfield, L. E.; Robinson, J. S.; Tisch, J. W. G.; Marangos, J. P. Ideal waveform to generate the maximum possible electron recollision energy for any given oscillation period. *Phys. Rev. Lett.* **2009**, *102*, 2–5, doi:10.1103/PhysRevLett.102.063003.
18. Haessler, S.; Balčiunas, T.; Fan, G.; Andriukaitis, G.; Pugžlys, A.; Baltuška, A.; Witting, T.; Squibb, R.; Zaïr, A.; Tisch, J. W. G.; Marangos, J. P.; Chipperfield, L. E. Optimization of Quantum Trajectories Driven by Strong-Field Waveforms. *Phys. Rev. X* **2014**, *4*, 021028, doi:10.1103/PhysRevX.4.021028.
19. Xi, J.; Xue, Y.; Xu, Y.; Shen, Y. Artificial neural network modeling and optimization of ultrahigh pressure extraction of green tea polyphenols. *Food Chem.* **2013**, *141*, 320–326, doi:10.1016/j.foodchem.2013.02.084.
20. Neocleous, C.; Schizas, C. Artificial Neural Network Learning: A Comparative Review. In *Methods and Applications of Artificial Intelligence, SETN 2002, Lecture Notes in Computer Science*; 2002; Vol. 2308, pp. 300–313.
21. Maleki, N.; Kashanian, S.; Maleki, E.; Nazari, M. A novel enzyme based biosensor for catechol detection in water samples using artificial neural network. *Biochem. Eng. J.* **2017**, *128*, 1–11, doi:10.1016/j.bej.2017.09.005.
22. Belalia Douma, O.; Boukhatem, B.; Ghrici, M.; Tagnit-Hamou, A. Prediction of properties of self-compacting concrete containing fly ash using artificial neural network. *Neural Comput. Appl.* **2017**, *28*, 707–718, doi:10.1007/s00521-016-2368-7.
23. Sitton, J. D.; Zeinali, Y.; Story, B. A. Rapid soil classification using artificial neural networks for use in constructing compressed earth blocks. *Constr. Build. Mater.* **2017**, *138*, 214–221, doi:10.1016/j.conbuildmat.2017.02.006.
24. Cimpoiu, C.; Cristea, V. M.; Hosu, A.; Sandru, M.; Seserman, L. Antioxidant activity prediction and classification of some teas using artificial neural networks. *Food Chem.* **2011**, *127*, 1323–1328, doi:10.1016/j.foodchem.2011.01.091.
25. Durodola, J. F.; Li, N.; Ramachandra, S.; Thite, A. N. A pattern recognition artificial neural network method for random fatigue loading life prediction. *Int. J. Fatigue* **2017**, *99*, 55–67, doi:10.1016/j.ijfatigue.2017.02.003.
26. Cristea, M. V.; Roman, R.; Agachi, S. P. Neural Networks Based Model Predictive Control of the Drying Process. *Eur. Symp. Comput. Aided Process Eng.* **2003**, *13*, 389–394.
27. Bouhoune, K.; Yazid, K.; Boucherit, M. S.; Chériti, A. Hybrid control of the three phase induction machine using artificial neural networks and fuzzy logic. *Appl. Soft Comput. J.* **2017**, *55*, 289–301, doi:10.1016/j.asoc.2017.01.048.
28. Erzin, Y.; Ecemis, N. The use of neural networks for the prediction of cone penetration resistance of silty sands. *Neural Comput. Appl.* **2017**, *28*, 727–736,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2018 doi:10.20944/preprints201809.0563.v1

Peer-reviewed version available at Appl. Sci. 2018, 8, 2106; doi:10.3390/app8112106

doi:10.1007/s00521-016-2371-z.

29. Bhatikar, S. R.; DeGroff, C.; Mahajan, R. L. A classifier based on the artificial neural network approach for cardiologic auscultation in pediatrics. *Artif. Intell. Med.* **2005**, *33*, 251–260, doi:10.1016/j.artmed.2004.07.008.
30. Hakeem, M. A.; Kamil, M. Analysis of artificial neural network in prediction of circulation rate for a natural circulation vertical thermosiphon reboiler. *Appl. Therm. Eng.* **2017**, *112*, 1057–1069, doi:10.1016/j.applthermaleng.2016.10.119.
31. Chen, Z.; Ma, W.; Wei, K.; Wu, J.; Li, S.; Xie, K.; Lv, G. Artificial neural network modeling for evaluating the power consumption of silicon production in submerged arc furnaces. *Appl. Therm. Eng.* **2017**, *112*, 226–236, doi:10.1016/j.applthermaleng.2016.10.087.
32. Oladunjoye, A. O.; Oyewole, S. A.; Singh, S.; Ijabadeniyi, O. A. Prediction of Listeria monocytogenes ATCC 7644 growth on fresh-cut produce treated with bacteriophage and sucrose monolaurate by using artificial neural network. *LWT - Food Sci. Technol.* **2017**, *76*, 9–17, doi:10.1016/j.lwt.2016.10.042.
33. Youshia, J.; Ali, M. E.; Lamprecht, A. Artificial neural network based particle size prediction of polymeric nanoparticles. *Eur. J. Pharm. Biopharm.* **2017**, *119*, 333–342, doi:10.1016/j.ejpb.2017.06.030.
34. Gherman, A. M. M.; Tosa, N.; Cristea, M. V.; Tosa, V.; Porav, S. Artificial neural networks modeling of the parameterized gold nanoparticles generation through photo-induced process. *Mater. Reserch Express* **2018**, *5*, doi:10.1088/2053-1591/aad0d5.
35. Lagaris, I. E.; Likas, A.; Fotiadis, D. I. Artificial neural network methods in quantum mechanics. *Comput. Phys. Commun.* **1997**, *104*, 1–14, doi:10.1016/S0010-4655(97)00054-4.
36. Kim, H. T.; Kim, I. J.; Tosa, V.; Lee, Y. S.; Nam, C. H. High brightness harmonic generation at 13 nm using self-guided and chirped femtosecond laser pulses. *Appl. Phys. B Lasers Opt.* **2004**, *78*, 863–867, doi:10.1007/s00340-004-1456-z.
37. Kovács, K.; Major, B.; Kőrös, C. P.; Rudawski, P.; Heyl, C. M.; Johnsson, P.; Arnold, C. L.; L'Huillier, A.; Toşa, V.; Varjú, K. Multi-parameter optimization of a loose focusing high flux high-harmonic beamline. *Prep.*
38. Tosa, V.; Takahashi, E.; Nabekawa, Y.; Midorikawa, K. Generation of high-order harmonics in a self-guided beam. **2003**, 1–4, doi:10.1103/PhysRevA.67.063817.
39. Priori, E.; Cerullo, G.; Nisoli, M.; Stagira, S.; De Silvestri, S.; Villoresi, P.; Poletto, L.; Ceccherini, P.; Altucci, C.; Bruzzese, R.; de Lisio, C. Nonadiabatic three-dimensional model of high-order harmonic generation in the few-optical-cycle regime. *Phys. Rev. A* **2000**, *61*, 1–8, doi:10.1103/PhysRevA.61.063801.
40. Lewenstein, M.; Balcou, P.; Ivanov, M. Y.; L'Huillier, A.; Corkum, P. B. Theory of high-harmonic generation by low-frequency laser fields. *Phys. Rev. A* **1994**, *49*, 2117–2132.
41. Tosa, V.; Kovacs, K.; Altucci, C.; Velotta, R. Generating single attosecond pulse using multi-cycle lasers in a polarization gate. *Opt. Express* **2009**, *17*, 17700–17710, doi:10.1364/OE.17.017700.
42. Balogh, E.; Kovacs, K.; Dombi, P.; Fulop, J. A.; Farkas, G.; Hebling, J.; Tosa, V.; Varju, K. Single attosecond pulse from terahertz-assisted high-order harmonic generation. *Phys. Rev. A - At. Mol. Opt. Phys.* **2011**, *84*, 1–9, doi:10.1103/PhysRevA.84.023806.
43. Negro, M.; Vozzi, C.; Kovacs, K.; Altucci, C.; Velotta, R.; Frassetto, F.; Poletto, L.; Villoresi, P.;

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2018 doi:10.20944/preprints201809.0563.v1

Peer-reviewed version available at Appl. Sci. 2018, 8, 2106; doi:10.3390/app8112106

de Silvestri, S.; Tosa, V.; Stagira, S. Gating of high-order harmonics generated by incommensurate two-color mid-IR laser pulses. *Laser Phys. Lett.* **2011**, *8*, 875–879, doi:10.1002/lapl.201110082.

44. Tosa, V.; Altucci, C.; Kovcs, K.; Negro, M.; Stagira, S.; Vozzi, C.; Velotta, R. Isolated attosecond pulse generation by two-mid-ir laser fields. *IEEE J. Sel. Top. Quantum Electron.* **2012**, *18*, 239–247, doi:10.1109/JSTQE.2011.2118193.

45. Balogh, E.; Kovács, K.; Toşa, V.; Varjú, K. A case study for terahertz-assisted single attosecond pulse generation. *J. Phys. B At. Mol. Opt. Phys.* **2012**, *45*, 074022, doi:10.1088/0953-4075/45/7/074022.

46. Kovacs, K.; Tosa, V.; Major, B.; Balogh, E.; Varju, K. High-Efficiency Single Attosecond Pulse Generation with a Long-Wavelength Pulse Assisted by a Weak Near-Infrared Pulse. *IEEE J. Sel. Top. Quantum Electron.* **2015**, *21*, doi:10.1109/JSTQE.2015.2411580.

47. Major, B.; Balogh, E.; Kovács, K.; Han, S.; Schütte, B.; Weber, P.; Vrakking, M. J. J.; Tosa, V.; Rouzée, A.; Varjú, K. Spectral shifts and asymmetries in mid-infrared assisted high-order harmonic generation. *J. Opt. Soc. Am. B* **2018**, *35*, A32, doi:10.1364/JOSAB.35.000A32.

48. Balogh, E.; Bódi, B.; Tosa, V.; Goulielmakis, E.; Varjú, K.; Dombi, P. Genetic optimization of attosecond-pulse generation in light-field synthesizers. *Phys. Rev. A - At. Mol. Opt. Phys.* **2014**, *90*, 1–7, doi:10.1103/PhysRevA.90.023855.

49. Tosa, V.; Kovács, K.; Ursescu, D.; Varjú, K. Characteristics of femtosecond laser pulses propagating in multiply ionized rare gases. *Nucl. Instruments Methods Phys. Res. B* **2017**, *408*, 271–275, doi:10.1016/j.nimb.2017.05.031.

50. Tosa, V.; Kovács, K.; Major, B.; Balogh, E.; Varjú, K. Propagation effects in highly ionised gas media. *Quantum Electron.* **2016**, *46*, 321–326, doi:10.1070/QEL16039.

51. Cristea, M. V.; Varvara, S.; Muresan, L.; Popescu, I. C. Neural networks approach for simulation of electrochemical impedance diagrams. *Indian J. Chem.* **2002**, *42*, 764–768.

52. Constant, E.; Garzella, D.; Breger, P.; Mével, E.; Dorrer, C.; Le Blanc, C.; Salin, F.; Agostini, P. Optimizing High Harmonic Generation in Absorbing Gases: Model and Experiment. *Phys. Rev. Lett.* **1999**, *82*, 1668–1671, doi:10.1103/PhysRevLett.82.1668.

53. Balcou, P.; Salieres, P.; L'Huillier, A.; Lewenstein, M. Generalized phase-matching conditions for high harmonics: The role of field-gradient forces. *Phys. Rev. A* **1997**, *55*, 3204–3210, doi:10.1103/PhysRevA.55.3204.

54. Salières, P.; L'Huillier, A.; Lewenstein, M. Coherence control of high-order harmonics. *Phys. Rev. Lett.* **1995**, *74*, 3776–3779, doi:10.1103/PhysRevLett.74.3776.

55. Major, B.; Kovács, K.; Tosa, V.; Rudawski, P.; L'Huillier, A.; Varjú, K. The effect of plasma-core induced self-guiding on phase matching of high-order harmonic generation in gases. *Prep.*