Deep-learning approach for large atomic structure calculations

Pavlo Bilous,[1,2,*] Adriana Pálffy,[3,2] and Florian Marquardt[1,4]

[1] *Max Planck Institute for the Science of Light, Staudtstraße 2, 91058 Erlangen, Germany*
[2] *Max Planck Institute for Nuclear Physics, Saupfercheckweg 1, 69117 Heidelberg, Germany*
[3] *Institute of Theoretical Physics and Astrophysics, University of Würzburg, Am Hubland, 97074 Würzburg, Germany*
[4] *Department of Physics, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany*

(Dated: September 14, 2022)

High-precision atomic structure calculations require accurate modelling of electronic correlations involving large multiconfiguration wave function expansions. Here we develop a deep-learning approach which allows to preselect the most relevant configurations out of large basis sets until the targeted precision is achieved. Our method replaces a large multiconfiguration Dirac-Hartree-Fock computation by a series of smaller ones performed on an iteratively expanding basis subset managed by a convolutional neural network. The results for several examples with many-electron atoms show that deep learning can significantly reduce the required computational memory and running time and renders possible large-scale computations on otherwise unaccessible basis sets.

The precise knowledge of atomic structure is indispensable for frequency standards in metrology, spectral analysis in astrophysics, understanding of nuclear phenomena involving atomic electrons, or investigations of physics beyond the standard model, e.g space and time variation of fundamental constants [1]. *Ab initio* atomic structure calculations are the scope of high performance codes that provide a wide range of electronic properties of atoms and ions, such as energy levels, radiative transition rates, g-factors or hyperfine structure constants. The practical difficulty arises from many-body effects when considering atoms or ions with high atomic number Z and many electrons. The electronic correlations are typically tackled by multiconfiguration wave function expansions combined with the configuration interaction (CI) method. The size of the involved basis set can easily become challenging even for state-of-the-art parallelized codes running on supercomputer systems, see e.g. recent calculations of electronic energy levels in Th^{35+} [2–4], Ir^{17+} [5, 6] or Fe^{16+} [5, 7].

The ability to preselect the basis states which have large relative weights would allow us to construct a compact wave function that delivers accurate observables without requiring the computational effort on the full basis. Selected CI methods were applied to atomic and molecular systems using selection criteria based on perturbation theory [8, 9] or the Monte-Carlo approach [10, 11]. However, these methods become inefficient for large basis sets, since perturbation theory still requires computations on the entire basis, whereas the random selection completely disregards the properties of its states. Recently in quantum chemistry selected CI based on machine learning has been demonstrated, for instance using supervised active learning [12–14], or reinforcement learning [15]. Neural networks (NN) have so far proved their versatility for the active learning iterations in selected CI [13], although for small electronic systems like H_4, H_2C or H_2O, support vector machines can be preferable [14]. Unfortunately, shallow NNs as applied in Ref. [12] as well as NNs of straightforward dense structure with more layers [13] are not sufficient for high-performance calculations for heavy atoms and ions. Apart from the strong electronic correlations, another reason is the more complicated labelling of the basis states due to presence of the rotational symmetry in atomic calculations as opposed to the quantum chemistry examples considered so far.

In this Letter, we develop an efficient deep-learning approach to iteratively construct an approximative wave function that delivers atomic level energies for high-Z atoms and ions with many electrons. The crucial step is application of a convolutional NN after appropriate encoding of basis states. This significantly improves the performance for atomic systems as compared to similar-sized NN of a simple dense type as applied e.g. in Refs. [12, 13]. The NN is trained on-the-fly and asked to select the important states for specific electronic levels. The atomic structure calculations are addressed employing the multiconfiguration Dirac-Hartree-Fock (MCDHF) method with the General Relativistic Atomic Structure Package GRASP2018 [16].

For benchmarking and demonstration of our approach we choose the case of neutral ^{187}Re and ^{187}Os atoms and calculate the respective ground state energies. These values have been just recently evaluated with GRASP2018 to extract the β-decay energy of ^{187}Re from experimentally determined masses of $^{187}Re^{29+}$ and $^{187}Os^{29+}$ [17]. Due to the prohibitively large basis set size, the authors of Ref. [17] had to preselect the most important basis states by evaluating transition and ionization energies and fitting them to experimental values [18]. In contrast, the NN allows to avoid *ad hoc* methods and carry out large computations from first principles by replacing them by a series of smaller ones performed on an iter-

atively expanding basis subset until the required energy precision is achieved.

In MCDHF, the wave function of the considered state Ψ is represented as an expansion $\Psi = \sum_\alpha c_\alpha \Phi_\alpha$ in the basis of configuration state functions (CSF) Φ_α, where α stands for all quantities uniquely characterizing a CSF. For a given electronic configuration $n_1 l_1^{a_1} \dots n_k l_k^{a_k}$ (here a is the occupation of the orbital with the principal and orbital quantum numbers n and l, respectively), CSFs are constructed from individual electronic wave functions by subsequently considering (1) different distributions of electrons over relativistic suborbitals; (2) all possible angular momenta couplings within each suborbital; (3) all possible couplings between the suborbitals to the total angular momentum J. The unknown coefficients c_α and the energy E of the electronic state are then obtained by solving the diagonalization problem $\hat{H}\Psi = E\Psi$, where $\hat{H}$ is the Dirac-Hartree-Fock Hamiltonian [19]. The diagonalization problem is solved iteratively with simultaneous evaluation of the radial electronic wave functions. When the latter are fixed, MCDHF reduces to the well-known configuration interaction (CI) [1].

A single electronic configuration $n_1 l_1^{a_1} \dots n_k l_k^{a_k}$ yields usually only a few to a few tens of CSFs. However, for accurate atomic structure data, correlation effects [19] need to be taken into account by additionally including CSFs originating from other configurations with the same good quantum numbers J and parity π. Thereby the targeted accuracy determines the size of the required basis set. These additional configurations are obtained from the initial one by moving ("exciting") electrons from occupied orbitals to vacancies in partially occupied or empty (virtual) orbitals up to some fixed orbital $n_v l_v$. In practice, the innermost core electronic shells are kept frozen and only excitations from higher occupied orbitals down to a fixed orbital $n_c l_c$ are allowed. The main difficulty stems from the very fast growth of the set size and the increased computational costs per matrix element. This is why often only single (S) and double (D) excitations are considered. We denote the obtained set of CSFs as $X(n_c l_c, n_v l_v)$, where the generic notation X stands for S, D or their combination.

For example, the set SD$(3p, 9h)$ for the ground state of Re and Os atoms considered here and in Ref. [17] consists of about 100 million CSFs and lies far beyond the computational capabilities of GRASP2018. Ref. [17] restricted the basis set to the approx. 5 million most important CSFs selected by reproducing experimental data for excitation and ionization energies [18]. However, such experimental data are often not available for highly charged ions and higher excited states of high-Z atoms. Other *ab initio* approximation methods such as e.g. the many-body perturbation theory [9, 20] are still computationally demanding for large sets and often do not provide reasonable precision. This renders the MCDHF problem with a large number of CSFs generally very challenging.

In this work, we develop a systematic and general deep-learning-based approach allowing for approximate MCDHF computations on large basis sets without the need to actually perform any atomic calculations on the entire CSF set. In our method, the diagonalization problem is solved iteratively on a relatively small subset expanded in each iteration until the required precision of the state energy is achieved. The WF set expansion is managed by a NN trained on the CSF weights $w_\alpha = |c_\alpha|^2$ obtained at the diagonalization stage in the current and previous iterations. Following Refs. [12, 13], we consider a binary output: a CSF is either important or unimportant, in the sense that w_α exceeds or not some cutoff value w^0 chosen in advance. However, instead of using a fixed cutoff as in Refs. [12, 13], we use a running cutoff taking at the i-th iteration a new value $w^i < w^{i-1} < \ldots < w^1 < w^0$. This choice is crucial for obtaining convergence of the energy to its true value, a point which was also observed in Ref. [14]. The final subset of CSFs together with their expansion coefficients c_α constitute an approximate state wave function which can be used for evaluation of further electronic properties of interest. We refer to it henceforth as the wave function (WF) subset.

Some important CSFs for a particular electron configuration are known from the start and should always be included in the WF set. They form the primary subset and we do not expose them to the NN at any stage. At first, the NN is trained on a random selection of CSFs from the considered set $X(n_c l_c, n_v l_v)$ (excluding the primary subset). Throughout this work we use 1% of CSFs for the initial training. In the i-th iteration, the WF set is enriched with the CSFs identified by the NN as important with respect to the cutoff w^i. Then the diagonalization problem is solved on the WF subset with GRASP2018. The obtained weights are used for re-training of the NN and to exclude from the WF set those CSFs which turned out to be unimportant.

In contrast to the earlier approaches [12, 13] which dealt with relatively simple input of spin-orbital occupations fed directly into usual dense NNs, we employ a NN of the convolutional type well known from image recognition applications [21, 22] combined with proper restructuring of the basis state labels. As input, the NN receives the information uniquely characterizing CSFs (see Fig. 1): (1) populations of the relativistic suborbitals n_k; (2) the total angular momenta for each suborbital J_k; (3) the intermediate angular momenta for couplings between them J_k^{cpl}. The orbitals are ordered in natural succession of quantum numbers. We normalize the populations of the orbitals to their maximal capacity and the angular momenta to the total angular momentum of the state. The three classes of input data are then interpreted as color channels of a 1D convolutional input layer.

A network architecture that we found to work effi-

ciently and which we focus on here is shown in Fig. 1. The input layer (A) consists of 3 channels (see color encoding in Fig. 1) of size N. The input is processed with a filter kernel of size 3 (B) resulting in 96 feature maps (C) each of size $N-2$. The latter are mapped to 16 feature maps (D) of size $N-2$ by application of a filter kernel of size 1 (thus representing a purely local transformation). The obtained output of $16\times(N-2)$ values is then flattened and forwarded to a network of 4 dense layers (E) with 150, 120, 90 and 2 neurons, respectively. The ReLU activation function was used throughout the NN apart from the two-neuron output layer (F), where the softmax function is applied yielding the probabilities of the CSF to be important or unimportant. The categorical cross-entropy was chosen as the loss function. The NN is trained on batches of size 32 using the Adam optimization algorithm with early stopping. The described NN functionality was implemented using the Python library Keras [23, 24] which is a high-level interface of the well-known TensorFlow library [25].

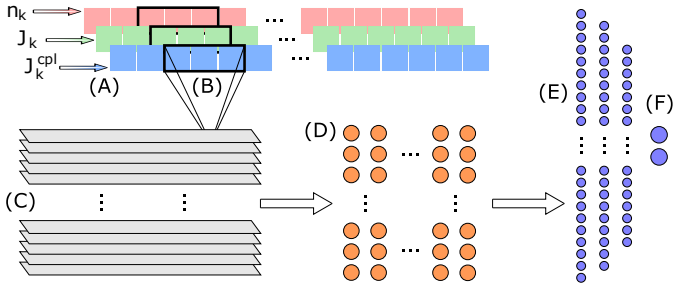

FIG. 1. NN architecture. See text for further explanations.

This NN architecture, together with the strategy of CSF selection and iterative re-training described above, constitutes our new approach to atomic structure calculations. For benchmarking we first consider a basis set of a moderate size, for which a direct full GRASP2018 calculation is possible. We choose the set SD$^*(3p,9h)$ for the Re atom ground state (4,267,362 CSFs characterized by 64×3 quantum numbers), where D^* denotes the subset of D with the additional restriction that each virtual orbital can be either doubly excited or empty. We simplify the GRASP2018 runs by considering at this stage just the Coulomb interaction and by keeping the radial wave functions fixed. These are obtained with GRASP2018 in advance using the layer-by-layer procedure as described in Ref. [26] and including SD excitations within the main configuration and S excitations to the virtual orbitals. The primary subset used in all iterations consists of 37,220 CSFs constructed based on SD excitations to the valence orbitals and S excitations to the virtual orbitals.

Following this procedure, our results show that the convolutional NN tends to treat both completely filled as well as empty orbitals as a common "background". Similarly to image recognition, this background is stripped off by the NN such that the remaining "image" represents the partially filled (open) orbitals. The input data undergos this transformation when moving from part (A) to part (D) of the NN (see Fig. 1). In the 16 feature maps (D), non-zero values tend to appear only at those of the 62 positions which correspond to the partially filled orbitals. This "background elimination" is observed for *all* CSFs in the set already after a few first training epochs, whereas the whole computation consists of a few tens of epochs. In this way, the NN discovers in the data the well-known physical fact that the CSF properties in a basis set are determined by the open orbitals. We attribute the better performance of our NN in comparison to NNs of a usual dense architecture to this CSF processing feature.

Table I shows the results obtained in each iteration: the energy $E^{\rm WF}$ on the current WF set with respect to the exact value $E^{\rm all}=-454,655.537$ eV (obtained separately in a direct calculation), the number of CSFs in the GRASP2018 run, and the time taken by diagonalization. We note that for fixed radial wave functions the energies always satisfy $E^{\rm WF}>E^{\rm all}$ [27]. The iterations are labelled by $\log_{10}w^i$ where w^i is the running cutoff value at the i-th iteration. The initial NN training on 1% randomly chosen CSFs is also represented in Table I in the row labelled as "Init.". After the very last iteration, CSFs unimportant with respect to the value $\log_{10}w^i=-11.6$ as calculated by GRASP2018 are excluded from the WF set yielding the final WF set with 784,446 CSFs instead of 1,083,274. The latter step is important if the obtained WF set is intended for further calculations on the state, such as e.g. improving the radial wave functions or evaluation of QED corrections and isotope shifts.

The accuracy of 7.6 meV was achieved on 180 cores in 85 min, of which 67 min were taken by the atomic calculations and the rest was used mostly for training of the NN. The peak memory consumed by the program was 177 GB. In contrast, the direct GRASP2018 computation on the same computer hardware configuration took 6.4 hours with the peak memory consumption of 1.3 TB. This comparison shows the clear advantages of the presented approach.

Fig. 2 illustrates graphically the growth of the WF set for the considered example of SD$^*(3p,9h)$. Plotted is the number of CSFs not (yet) included in the WF set as a function of the weight $\log_{10}w_\alpha$ for each iteration. The distribution is normalized with respect to the total size of the SD$^*(3p,9h)$ set and the weights w_α are taken from the full GRASP2018 calculation. In each iteration, CSFs are included in the WF set (and thus removed from the depicted distributions) from the right. The right edge of the distributions is not completely sharp, meaning that not all CSFs important with respect to the current cutoff are included in the WF set. The NN selection ensures

$\log_{10} w^i$	$E^{\rm WF} - E^{\rm all}$, meV	CSFs in GRASP	Time, min
Init.	11,180.2	79,521	1.6
-8.6	3,590.8	244,405	3.5
-9.2	900.1	396,378	6.2
-9.8	114.1	597,313	11.0
-10.4	32.6	818,002	17.3
-11.0	7.6	1,083,274	27.6

TABLE I. Results of approximate energy calculations on the SD$^*(3p, 9h)$ basis set for the Re atom ground state using our deep-learning-based approach. Iterations are labelled by $\log_{10} w^i$ where w^i is the running cutoff value at the i-th iteration. The row labelled as "Init." represents the initial NN training on 1% randomly chosen CSFs.

that the slope becomes stable in the first iterations and moves then from right to left. As alternative, we have checked the energy convergence for the case that CSFs were randomly picked instead. In this case, the chosen CSF come mostly from the maximum region of the distribution, and only a small fraction of them is important, leading to a very slow expansion of the WF set and impractical slow convergence of the calculated energy.

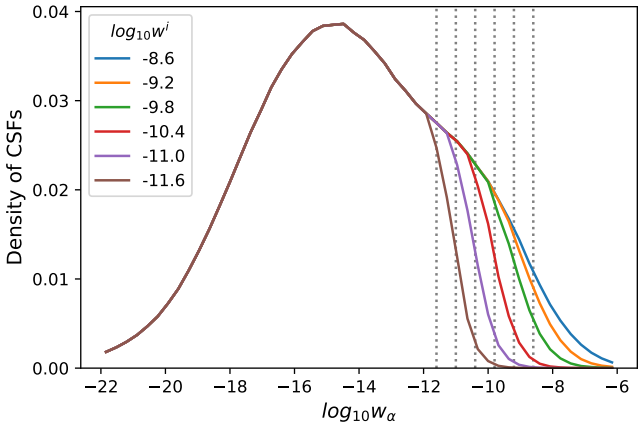

FIG. 2. (Color online) The distribution of CSF from the set SD$^*(3p, 9h)$ not (yet) included in the WF set as a function of their respective weights $\log_{10} w_\alpha$ for each iteration. The distributions are normalized with respect to the total size of the SD$^*(3p, 9h)$ set. The iterations are labelled by the cutoff values $\log_{10} w^i$ which are additionally illustrated by the vertical dotted lines.

The WF set expansion is a stochastic process implying different final energy values $E^{\rm WF}$ for separate runs. To check the stability, we have run the same calculation 30 times. This yielded for the (positive) quantity $E^{\rm WF} - E^{\rm all}$ the average value of 7.8 meV with the mean squared deviation of approx. 40%. At the same time, the WF set size in the last and most computationally demanding iteration deviated only 5% from its average value of 1,065,268 CSFs, ensuring also stability of the required resources. Further checks showed that variation of the NN hyperparameters influences the result smoothly and mostly within the stochastic distribution for $E^{\rm WF} - E^{\rm all}$. Since the convolutional network can deal with arbitrary input sizes, the NN parameters applied here can be used as a reasonable starting point for other computations.

Finally, we have performed the described calculations replacing the convolutional NN by a NN of a usual dense architecture with similar number of trainable parameters. The concrete shape of the latter was varied. As input, we used both the (flattened) representation applied in this work, and a binary encoding of the input data favoured by dense NNs. In all considered cases, the obtained energy deviation from $E^{\rm all}$ was several times larger than the average value 7.8 meV achieved using a convolutional NN. This advantage over simple dense NNs as applied in Refs. [12, 13] is especially important for large CSF sets as the one considered in the following.

We switch now to calculations on the large basis sets SD$(3p, 9h)$ for the Re and Os neutral atoms relevant for the determination of the ^{187}Re β-decay energy in Ref. [17]. These calculations involve basis sets of over 90 million CSFs and cannot be performed directly using GRASP2018. However, it is sufficient to retain for each set only the CSFs that deliver a 1 eV precision for the energy. Using our deep-learning approach, we could achieve the targeted accuracy by performing partial GRASP2018 runs on up to about 4 million CSFs. The CSFs included in the primary subset were constructed as in the previous example resulting in 37,220 and 32,660 CSFs for the Re and Os atom, respectively. The radial wave functions were obtained in advance in the layer-by-layer procedure with GRASP2018 [26] by including SD excitations to the valence orbitals and S excitations to the virtual orbitals, and kept fixed in the CSF selection process.

We started from the cutoff value $\log_{10} w^0 = -8.0$ and performed iterations with decreasing it in steps from 0.3 to 0.5 until the energy converged. In order to make sure that the converged value represents the correct energy, we carried out a few separate runs choosing each time different cutoff paths. Each computation could be performed on 500–600 cores within one day (a half of this time was taken by the NN training) with the peak memory consumption of 4 TB. The approximative WF sets for neutral Re and Os atoms were then used to refine the atom energies by including QED corrections and isotope shifts and by improving the radial wave functions for the virtual orbitals. In order to compare our results with Ref. [17] where the electronic binding energy difference δE between a neutral atom and a 29+ ion for Re and Os was provided, we evaluate the energies of the Re^{29+} and Os^{29+} ions on the basis sets SD$(3p, 9h)$. Since these consist of only 53,885 and 2,455,449 CSFs, respectively, we carry out MCDHF calculations directly taking into account also QED corrections and isotope shifts using

GRASP2018.

Based on the calculated atom and ion energies, we find $\delta E_{\rm Re} = -10,905.80$ eV and $\delta E_{\rm Os} = -10,958.94$ eV, which agree with the values $\delta\widetilde{E}_{\rm Re} = -10,894.5 \pm 25.9$ eV and $\delta\widetilde{E}_{\rm Os} = -10,947.9 \pm 24.6$ eV from Ref. [17]. For the quantity ΔE we obtain $\Delta E = 53.14$ eV, whereas in Ref. [17] the value $\Delta\widetilde{E} = 53.4 \pm 1.0$ eV was reported. Our results are thus in good agreement. We note that although the set SD$(3p, 9h)$ accounts for a large part of the correlation corrections, its choice affects the individual atom energies significantly. By carrying out analogous calculations for "adjacent" sets, e.g., SD$(3d, 9h)$ or SD$(3d, 10h)$, we have observed that whereas a change of the lowest core orbital influences the binding energy at the level of 1 eV, expansion to the valence orbitals with $n = 10$ leads to a variation of the result of the order of 10 eV. The small error of ΔE with respect to $\delta E_{\rm Re}$ and $\delta E_{\rm Os}$ is the result of cancellation of the correlation effects due to the similarity of the Re and Os neutral atoms. This calls for further investigations of the Re and Os electronic structure, e.g. using our deep-learning approach.

In conclusion, our deep-learning-based approach offers a crucial simplification of MCDHF calculations on large basis sets, which are otherwise computationally demanding and often not feasible. The NN allows to replace a large computation by a series of smaller ones performed on an iteratively expanding basis subset until the targeted precision is achieved. A benchmark has been performed on intermediate-size basis for the examples of Re and Os, and energy predictions were made on large basis sets which are at present not accessible with GRASP2018. The developed approach can be used with other atomic codes, which do not necessarily employ the rotational symmetry and are based on an uncoupled basis of Slater determinants (see for instance Ref. [5]).

We are thankful to Marianna Safronova, Sergey Porsev and Charles Cheung from University of Delaware for fruitful discussions. AP gratefully acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG) in the framework of the Heisenberg Program.

REFERENCES

* pavlo.bilous@mpl.mpg.de

[1] C. F. Fischer, M. Godefroid, T. Brage, P. Jönsson, and G. Gaigalas, Journal of Physics B: Atomic, Molecular and Optical Physics **49**, 182004 (2016).

[2] P. V. Bilous, H. Bekker, J. C. Berengut, B. Seiferle, L. von der Wense, P. G. Thirolf, T. Pfeifer, J. R. C. López-Urrutia, and A. Pálffy, Phys. Rev. Lett. **124**, 192502 (2020).

[3] S. G. Porsev, C. Cheung, and M. S. Safronova, Quantum Science and Technology **6**, 034014 (2021).

[4] E. Peik, T. Schumm, M. S. Safronova, A. Pálffy, J. Weitenberg, and P. G. Thirolf, Quantum Science and Technology **6**, 034002 (2021).

[5] C. Cheung, M. Safronova, and S. Porsev, Symmetry **13** (2021), 10.3390/sym13040621.

[6] M. G. Kozlov, M. S. Safronova, J. R. Crespo López-Urrutia, and P. O. Schmidt, Rev. Mod. Phys. **90**, 045005 (2018).

[7] S. Kühn, C. Shah, J. R. C. López-Urrutia, K. Fujii, R. Steinbrügge, J. Stierhof, M. Togawa, Z. Harman, N. S. Oreshkina, C. Cheung, M. G. Kozlov, S. G. Porsev, M. S. Safronova, J. C. Berengut, M. Rosner, M. Bissinger, R. Ballhausen, N. Hell, S. Park, M. Chung, M. Hoesch, J. Seltmann, A. S. Surzhykov, V. A. Yerokhin, J. Wilms, F. S. Porter, T. Stöhlker, C. H. Keitel, T. Pfeifer, G. V. Brown, M. A. Leutenegger, and S. Bernitt, Phys. Rev. Lett. **124**, 225001 (2020).

[8] R. J. Harrison, The Journal of Chemical Physics **94**, 5021 (1991), https://doi.org/10.1063/1.460537.

[9] V. A. Dzuba, V. V. Flambaum, and M. G. Kozlov, Phys. Rev. A **54**, 3948 (1996).

[10] J. C. Greer, The Journal of Chemical Physics **103**, 7996 (1995), https://doi.org/10.1063/1.470218.

[11] J. Greer, Journal of Computational Physics **146**, 181 (1998).

[12] J. P. Coe, Journal of Chemical Theory and Computation **14**, 5739 (2018), pMID: 30285426, https://doi.org/10.1021/acs.jctc.8b00849.

[13] W. Jeong, C. A. Gaggioli, and L. Gagliardi, Journal of Chemical Theory and Computation **17**, 7518 (2021).

[14] S. D. Pineda Flores, Journal of Chemical Theory and Computation **17**, 4028 (2021), pMID: 34125549, https://doi.org/10.1021/acs.jctc.1c00196.

[15] J. J. Goings, H. Hu, C. Yang, and X. Li, Journal of Chemical Theory and Computation **17**, 5482 (2021), pMID: 34423637, https://doi.org/10.1021/acs.jctc.1c00010.

[16] C. Froese Fischer, G. Gaigalas, P. Jönsson, and J. Bieroń, Computer Physics Communications **237**, 184 (2019).

[17] P. Filianin, C. Lyu, M. Door, K. Blaum, W. J. Huang, M. Haverkort, P. Indelicato, C. H. Keitel, K. Kromer, D. Lange, Y. N. Novikov, A. Rischka, R. X. Schüssler, C. Schweiger, S. Sturm, S. Ulmer, Z. Harman, and S. Eliseev, Phys. Rev. Lett. **127**, 072502 (2021).

[18] C. Lyu, private communication (2022).

[19] I. P. Grant, *Relativistic calculation of molecular properties* (Springer New York, New York, NY, 2007).

[20] I. Lindgren and J. Morrison, *Atomic Many-Body Theory* (Springer Berlin Heidelberg, Berlin, Heidelberg, 1986).

[21] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, 2016) `http://www.deeplearningbook.org`.

[22] A. Krizhevsky, I. Sutskever, and G. E. Hinton, in *Advances in Neural Information Processing Systems*, Vol. 25, edited by F. Pereira, C. Burges, L. Bottou, and K. Weinberger (Curran Associates, Inc., 2012).

[23] F. Chollet *et al.*, "Keras," `https://keras.io` (2015).

[24] A. Gulli and S. Pal, *Deep learning with Keras* (Packt Publishing Ltd, 2017).

[25] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard,

et al., in *12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16)* (2016) pp. 265–283.

[26] J. Bieroń, C. F. Fischer, G. Gaigalas, I. Grant, and P. Jönsson, "A practical guide to grasp2018 — a collection of fortran 95 programs with parallel computing using mpi," (2018).

[27] A. Szabo and N. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover Books on Chemistry (Dover Publications, 2012).