1. PHD PROJECT DESCRIPTION (4000 characters max., including the aims and work plan)

Project title: Machine Learning *ab initio* DFT.

1.1. Project goals

The density functional theory (DFT) is a workhorse method in computational chemistry and materials science that is applied to studying phenomena in biochemical processes, solid-state physics, organic and inorganic substances, etc. Despite the widespread use of DFT, the accuracy of its predictions is often hard to ensure because no exact exchange-correlation functional was found that applies to practical systems of interest. On the other hand, *the ab initio route of development DFT, analogously to the wavefunction-based theory, can systematically control* the accuracy of the resulting *ab initio* DFT. The orbital-dependent Kohn-Sham DFT functionals follow this route. Unfortunately, *ab initio* DFT has several technical bottlenecks preventing it from becoming a mainstream simulation tool. In this project, the Ph.D. student will be working on breaking through these bottlenecks by leveraging the power of machine learning (ML), which has already helped to break through the limitations of more approximate semi-empirical quantum mechanical methods and is used to improve non-*ab initio* DFT. We will develop ML methods for tuning the optimized effective potential for obtaining more numerically stable *ab initio* DFT. ML will also be applied to improve the correlation functionals. These will enable the practical use of the *ab initio* DFT for routine computational chemistry and materials science simulations.

1.2. Outline

We will use ML to make *ab initio* DFT methods both numerically stable and approaching the accuracy of the high-level wavefunction correlated theories (at least of CCSD(T)/CBS level) while retaining the affordable cost based on the moderate-size basis sets (double and triple-zeta). ML will be used to tune the optimized effective potential and improve the correlation functional in the orbital-dependent DFT.

1.3. The methods will be implemented in the state-of-the-art ML-enhanced computational chemistry software MLatom, possibly via interfaces to PySCF and other required DFT libraries. The training of ML will be done with MLatom and its available databases used to develop other ML-enhanced quantum chemical methods. We will test the performance of the resulting methods for relative energies, enthalpies of formation, and other electronicn properties, and check their adherence to correct physical behavior.Work plan

i. Derivation of the ML-integrated solution for tuning optimized effective potential.
ii. Implementation of the derived solutions.
iii. Training and testing ML models.
iv. Design, implementation, and testing of ML-improved correlation functionals.
v. Analyzing the adherence of the implemented models to required physical behavior, testing performance on energies and other electronic properties.

1.4. **Literature** *(max. 10 listed, as a suggestion for a PhD candidate)*

[1] R. J. Bartlett , I. Grabowski, S. Hirata, S. Ivanov, J. Chem. Phys. **122**, 034104, (2005)
[2] I. Grabowski, E. Fabiano, S. Śmiga, A. Buksztel, A. M. Teale,, F. Della Sala *J. Chem. Phys.* **141** (2014), 024113-1
[3] S. Śmiga, F. Della Sala, A. Buksztel, I. Grabowski, E. Fabiano *J Comput Chem* **37** (2016), 2081-2090
[4] P. O. Dral, O. A. von Lilienfeld, W. Thiel. *J. Chem. Theory Comput.* **11** (2015), 2120-2125.
[5] *Quantum Chemistry in the Age of Machine Learning*. P. O. Dral, ed. Elsevier: Amsterdam, Netherlands (2023).

1.5. Required initial knowledge and skills of the PhD candidate

- Basic knowledge of quantum chemistry density functional theory.
- Fundamentals of machine learning.
- Programming skills (Python).
- Good spoken and written English
- Basic experience with research work

1.6. Expected development of the PhD candidate's knowledge and skills

- Gaining in-depth experience with machine learning
- Advanced understanding of DFT methods
- Solid grip on effective potentials and basis sets
- Advanced knowledge of the DFT functionals.
- Independent work on research projects involving all stages from theoretical design to implementation and tests to analysis of results.
- Computational chemistry skills