

# Quantum chemistry and machine learning in computational materials and interface chemistry

Axel Groß  
Institute of Theoretical Chemistry  
Ulm University

## Abstract

Due to the ever-increasing improvement in computer power and the development of more efficient codes, atomistic first-principles calculations have become an indispensable tool in materials and interface chemistry. They are no longer limited to explanatory purposes but have gained predictive power, so that computational modelling and experiment can collaborate on the same footing. Still, quantum chemical first-principles methods are computationally expensive, and there are certain limitations as far as their scaling behavior in high-performance computing is concerned. Hence computational methods based on machine learning have become increasingly popular as an alternative approach to study materials and interfaces. Here some success stories of both approaches will be presented and their respective advantages and disadvantages will be critically discussed.

## 1 Introduction

In recent years, first-principles quantum chemical calculations have tremendously contributed to further our understanding of structures and processes in materials [10], at surfaces and interfaces [8], in catalysis [22], in electrochemical energy storage and conversion [11], just to name a few. They have become an indispensable tool in research, but they also play an increasingly important role for example in the materials development. Still, in spite of the ever-increasing computer power and the development of more efficient codes, quantum chemical studies are still restricted to at most a few thousand atoms, either in total or per unit cell in periodic codes, because of the exponential scaling of these codes with the number of considered atoms.

With the advent of methods and techniques based on artificial intelligence (AI) and machine learning (ML), there have been also numerous attempts to use these techniques in computational materials and interface chemistry [5]. One of the earliest applications has been to use artificial neural networks (NNs) in the construction of chemical interaction potentials [14] as a reliable substitute for classical force fields, allowing to perform simulations at much longer time and length scales compared to first-principles simulations. Significant progress has been made in the meantime, also allowing to reproduce long-range charge transfer employing advanced NN potentials [4].

Another more recent application of AI and ML techniques to materials and interface problems is the identification of so-called descriptors [25] which correspond to a correlation between a fundamental materials property and a functional property in the sense of structure-property relationships. Here, we will present recent examples of studies in computational materials and interface chemistry and compare first-principles with AI-based approaches.

## 2 Modelling of electrochemical interfaces

Electrochemical interfaces between electrodes and electrolytes play an essential role in the electrochemical energy conversion and storage [11]. However, their theoretical modelling is computationally rather expensive, in particular when liquid electrolytes are involved, as their proper theoretical treatment requires demanding statistical sampling of complex and extended systems [19]. This numerical demand can be addressed by

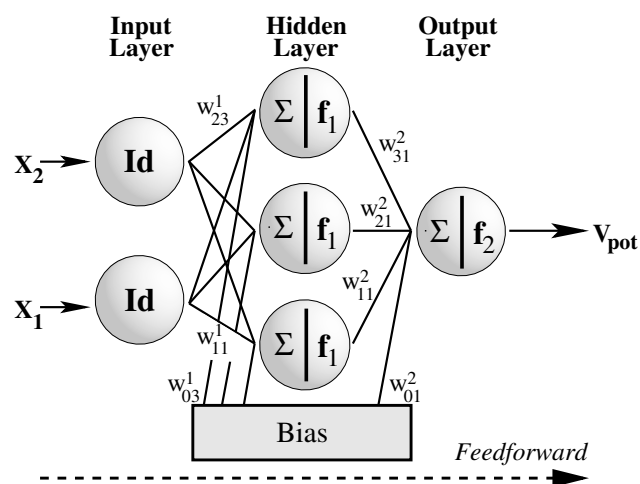


Figure 1: Schematic architecture of a feed-forward neural network with the neurons arranged in layers. The output of this NN is given by an expansion in non-linear basis functions  $f_{1,2}(x)$  is:  $V_{pot}(x_1, x_2) = f_2\left(w_{01}^{(2)} + \sum_{j=1}^3 w_{j1}^{(2)} f_1\left(w_{0j}^{(1)} + \sum_{i=1}^2 w_{ij}^{(1)} x_i\right)\right)$  (from Ref. [14] under the terms of the Creative Commons Attribution 3.0 License).

39 representing the chemical interaction between the atomic species in the system through adjusting fitting  
40 functions.

41 Artificial neural networks have become rather popular in this regard as they can in principle approximate  
42 any continuous function to arbitrary accuracy. Neural Networks can be regarded as an expansion in general,  
43 non-linear fitting functions that do not require any assumptions about the functional form of the underlying  
44 problem [4, 14]. The functions are chosen to mimic the functioning of neurons in the brain. A schematic  
45 presentation of an artificial NN is shown in Fig. 1. It is constructed from a number of artificial neurons or  
46 nodes that are typically arranged in layers and interconnected via a set of links. The particular functional form  
47 used in the construction of NNs is given in the captions of Fig. 1.

48 NN potentials constructed in this way from density functional theory input data have been used, among  
49 many other applications [5], to model structures and processes at the interface of water with low-index metal  
50 surfaces [16, 13]. The left panel of Fig. 2 illustrates these simulations. NN molecular dynamics simulations  
51 have been performed for several water layers on Cu surfaces including about 2000 atoms in the periodic cell  
52 for run times of about 1 ns. Figure 2 shows contour plots of the distribution of the oxygen atoms of the  
53 water molecules in the first hydration layer of Cu(111)–water interfaces [13]. In this particular study, the  
54 self-diffusion of surface defects such as adatoms and vacancies at copper–water interfaces was studied. This  
55 study demonstrated that water structures close to adatoms and vacancies differ strongly from those at flat  
56 metal surfaces. Systems sizes and time scales considered in this NN study can not be realized by first-principles  
57 simulations.

58 However, for electrochemical applications one has to take into account that the electrolytes contain ions  
59 that can strongly interact with electrodes, in particular as far as anions are concerned [7, 6]. Therefore, in order  
60 to realistically model electrochemical interfaces, additional elements have to be included in the NN potential.  
61 Unfortunately, the construction of NN potentials exhibits an unfavorable scaling with the number of elements,  
62 as all possible permutations of the elements have to be taken into account in the construction of the NN  
63 potentials [4]. Furthermore, the modelling of electrochemical interfaces often requires to take the electronic  
64 degrees of freedom explicitly into account. This is illustrated in the right panel of Fig. 2 where results of *ab*  
65 *initio* molecular dynamics (AIMD) simulations of a water film at Pt(111) [19] are illustrated. The water metal  
66 interface was modeled by six water layers within a  $6 \times 6$  surface unit cell corresponding to 144 water molecules  
67 on a five-layer Pt(111) slab, and the AIMD simulations could just been run for 40 fs.

68 Averaging the work function, i.e., the energy to extract an electron from the electrode to the vacuum

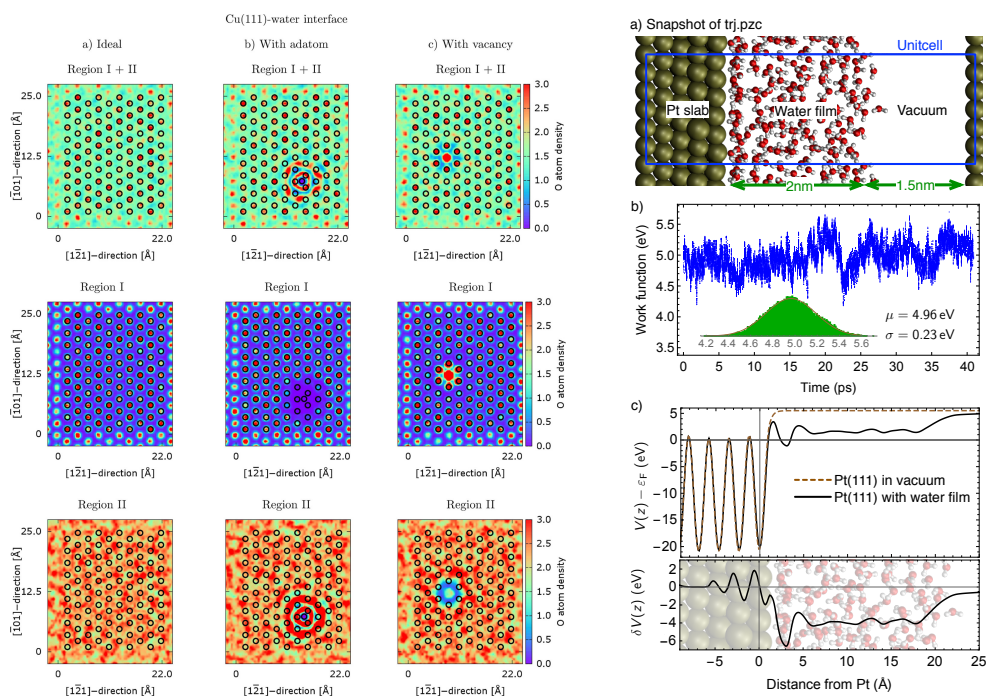


Figure 2: Simulation of water/metal interfaces. Left panel: Contour plots showing the distribution of the oxygen atoms of the water molecules of NN simulations of the first hydration layer of the Cu(111)-water interface for the ideal surface (a), the surface with an adatom (b), and a surface with a vacancy (c). (Reprinted with permission from Ref. [13]. Copyright 2017 American Chemical Society.) Right panel: (a) snapshot and (b) time evolution of the work function along an AIMD trajectory of the water/Pt(111) interface equilibrated at 298 K. (c) Electrostatic potential  $V(z)$  of Pt(111) in vacuum (dashed line) and averaged in water together with the difference  $\delta V(z)$ . (Reprinted from Ref. [19], with the permission of AIP Publishing.)

69 region through an ion-free water film along the simulations yields the so-called potential of zero charge of  
 70 the particular metal electrode. The results of the AIMD simulations have been in good agreement with  
 71 experimental results [19]. A closer analysis of the charge distribution furthermore shows that the resulting  
 72 dipole layer at the Pt/water interface dominating the work functions is not only governed by the reorientation  
 73 of the polar water molecules in the first water layer above the Pt electrode but also by a strong charge transfer  
 74 from the water layers to the Pt surface [19, 21] which is not trivial to reproduce using NN potentials.

### 75 3 Descriptors

76 Technological advances are often closely connected to the development of materials with improved properties.  
 77 Atomistic computational studies play an increasingly important role in the study of materials [9], insights from  
 78 such simulations and calculations can accelerate materials discovery. A particularly helpful concept to speed  
 79 up materials development is based on the concept of so-called descriptors.

80 Descriptors can be defined as a fundamental materials property that is directly correlated to a desired (or  
 81 undesired) functional property. As a very prominent example, the oxygen binding energy on metal catalysts  
 82 can be used as a descriptor for the activity of the oxygen reduction reaction on these catalysts [17, 20], but  
 83 also metal self-diffusion barriers have been proposed to act as a descriptor for the so-called dendrite growth  
 84 in batteries [12], which is supposed to be the origin for battery fires. The determination of descriptors can  
 85 considerably speed up the search for novel materials with desired functional or multifunctional properties,  
 86 because, once they are identified, only the particular descriptor needs to be optimized in a first step to  
 87 preselect promising candidate materials.

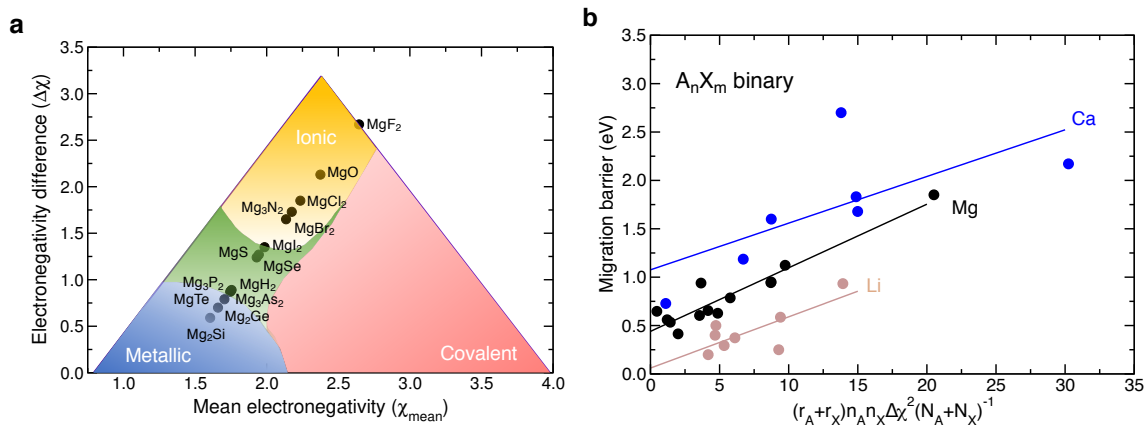


Figure 3: (a) Van Arkel-Ketelaar triangle for  $Mg_nX_m$  binaries plotted as a function of the mean electronegativity and the difference in the electronegativity of the two components. (b) Calculated migration barriers of  $A = \text{Li}, \text{Mg}, \text{and Ca}$  in  $A_nX_m$  binaries as a function of the migration number  $N_{\text{migr}}^{\text{AX}}$  for various elements  $X$  according to Eq. (1). Linear regressions of the results are plotted as solid lines. (From Ref. [24] under the terms of the CC BY 4.0 License.)

88 Traditionally, descriptors have been derived based on rational arguments and a concept of the principles  
 89 underlying the connection between structural and functional properties [17]. Here we will present a recent  
 90 example related to the ion mobility in crystalline electrodes and solid electrolytes [24]. Ion mobility is a critical  
 91 parameter significantly influencing the efficiency of batteries [3]. Typically, this mobility has been discussed  
 92 using the size of the migrating ions and the charge state, but no true understanding of the factors influencing  
 93 the ion mobility across different classes of materials has been derived yet [3]. However, periodic density  
 94 functional theory calculations of the ion site preference in spinel materials have shown that the charge carrier  
 95 mobility in such materials can not be fully understood solely based on ionic concepts [23]. These findings  
 96 have motivated more detailed investigations [24] of the ion mobility in a series of different materials. In fact,  
 97 the bonding characteristics in materials can be classified in so-called Van Arkel-Ketelaar triangles [1] in which  
 98 compounds are arranged according to the mean electronegativity  $\chi_{\text{mean}}$  ( $x$ -axis) and the electronegativity  
 99 difference  $\Delta\chi$  ( $y$ -axis) of the constituting elements (see Fig. 3a).

100 Based on this notion, the *migration* parameter or number  $N_{\text{migr}}$  was defined combining the square of the  
 101 electronegativity with the known dependence of the diffusion barriers on the ionic radii ( $r_i$ ) and oxidation  
 102 states ( $n_i$ ) of the involved compounds,

$$N_{\text{migr}}^{\text{AX}} = (r_A + r_X)n_A n_X \Delta\chi_{\text{AX}}^2 / (N_A + N_X), \quad (1)$$

103 as the product of these three quantities. Figure 3b shows the dependence of migration barriers as a function  
 104 of the migration parameter for the three migrating ions  $\text{Li}^+$ ,  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  in a number of binary materials,  
 105 yielding a satisfactory linear scaling in all three cases. Such linear scaling relations can be rather helpful because  
 106 once the scaling factor is known, promising candidates with a high ion mobility can be identified just based  
 107 on materials parameters that are readily available from openly accessible tables without further computations.

108 However, as such descriptors in principle correspond to a correlation between two properties, they can also  
 109 be identified by statistical methods based, e.g., on machine learning techniques. And indeed such methods  
 110 have been developed within the framework of compressed-sensing-based dimensionality reduction: The sure  
 111 independence screening and sparsifying operator (SISSO) gives stable results for descriptors, also with small  
 112 training sets [18]. This approach has been applied to the data presented in Fig. 3. Interestingly enough, this  
 113 machine learning approach yielded a descriptor with some other functional dependencies which in fact was  
 114 associated with a slightly larger root mean square error than the rationally derived descriptor  $N_{\text{migr}}$  [24]. In  
 115 order to understand this discrepancy, one has to take into account that the training set presented to the  
 116 SISSO operator was still rather small. This is a common problem in many areas of materials and interface  
 117 chemistry in which the number of reliable quantitative functional properties, derived either from experiment  
 118 or from calculations, is still rather low [2].

## 119 4 Discussion and Conclusions

120 In this contribution, quantum chemical and machine learning approaches in computational materials and inter-  
121 face chemistry were compared with each other using a few selected example. As far as machine learning  
122 potentials derived from quantum chemical calculations are concerned, they allow to perform simulations of  
123 complex systems at much larger length and time scales as quantum chemical calculations, but with a com-  
124 parable reliability. However, the construction of these potentials still scales rather unfavorably with the  
125 number of elements typically restricting these studies to systems with a small number of different elements.  
126 Furthermore, properties that explicitly depend on electronic degrees of freedom can still hardly be derived from  
127 these potentials. As an alternative approach to quantum chemical calculations, the implementation of the  
128 these methods on quantum computer is intensively discussed with first examples already available [15]. How-  
129 ever, these applications still suffer from the low fidelity of present quantum computers and the fact that still  
130 substantial input from quantum chemical calculations performed on classical computers is needed. Therefore,  
131 quantum chemical simulations on classical computers will most probably remain the basis for computational  
132 materials and interface chemistry studies for a long time.

133 Artificial intelligence and machine learning methods are typically rather capable to find correlations in large  
134 data sets. Hence they should for example be well-suited to identify descriptors in materials and interface  
135 chemistry. Yet, in these fields the number of really reliable quantitative data is still limited. Consequently, a  
136 blended learning approach combining rational and numerical approaches might correspond to the method of  
137 choice at the moment.

## 138 5 Acknowledgements

139 This research presented in this extended abstract has been supported by the German Research Foundation  
140 (DFG) through contract GR 1503/21-1 and through the Cluster of Excellence POLiS (EXC-2154, project  
141 ID 390874152) and by the Baden-Württemberg Foundation within the Network of Excellence Functional  
142 Nanostructures. Computer time was provided by the state of Baden-Württemberg through the bwHPC project  
143 and the Germany Research Foundation (DFG) through grant numbers INST 40/575-1 and INST 40/467-1  
144 FUGG (JUSTUS I and II clusters).

## 145 References

- 146 [1] L. C. Allen, J. F. Capitani, G. A. Kolks, and G. D. Sproul. Van Arkel—Ketelaar triangles. *J. Mol. Struct.*,  
147 300:647–655, 1993.
- 148 [2] M. Andersen, S. V. Levchenko, M. Scheffler, and K. Reuter. Beyond scaling relations for the description  
149 of catalytic materials. *ACS Catal.*, 9:2752–2759, 2019.
- 150 [3] J. C. Bachman, S. Muy, A. Grimaud, H.-H. Chang, N. Pour, S. F. Lux, O. Paschos, F. Maglia, S. Lu-  
151 part, P. Lamp, L. Giordano, and Y. Shao-Horn. Inorganic solid-state electrolytes for lithium batteries:  
152 Mechanisms and properties governing ion conduction. *Chem. Rev.*, 116:140–162, 2016.
- 153 [4] J. Behler. Four generations of high-dimensional neural network potentials. *Chem. Rev.*, 121:10037–10072,  
154 2021.
- 155 [5] M. Ceriotti, C. Clementi, and O. Anatole von Lilienfeld. Introduction: Machine learning at the atomic  
156 scale. *Chem. Rev.*, 121:9719–9721, 2021.
- 157 [6] F. Gossenberger, F. Juarez, and A. Groß. Sulfate, bisulfate, and hydrogen co-adsorption on Pt(111) and  
158 Au(111) in an electrochemical environment. *Front. Chem.*, 8:634, 2020.
- 159 [7] F. Gossenberger, T. Roman, and A. Groß. Equilibrium coverage of halides on metal electrodes. *Surf.*  
160 *Sci.*, 631:17 – 22, 2015.

- 161 [8] A. Groß. The virtual chemistry lab for reactions at surfaces: Is it possible? Will it be helpful? *Surf. Sci.*,  
162 500:347, 2002.
- 163 [9] A. Groß and L. A. Kibler. Elementary reaction steps in electrocatalysis: Theory meets experiment.  
164 *Electrocatal.*, 8:499–500, 2017.
- 165 [10] J. Hafner. Ab-initio simulations of materials using VASP: Density-functional theory and beyond. *J. Comp.*  
166 *Chem.*, 29:2044–2078, 2008.
- 167 [11] N. G. Hörmann, M. Jäckle, F. Gossenberger, T. Roman, K. Forster-Tonigold, M. Naderian, S. Sakong, and  
168 A. Groß. Some challenges in the first-principles modeling of structures and processes in electrochemical  
169 energy storage and transfer. *J. Power Sources*, 275:531 – 538, 2015.
- 170 [12] M. Jäckle and A. Groß. Influence of electric fields on metal self-diffusion barriers and its consequences  
171 on dendrite growth in batteries. *J. Chem. Phys.*, 151:234707, 2019.
- 172 [13] S. Kondati Natarajan and J. Behler. Self-diffusion of surface defects at copper–water interfaces. *J. Phys.*  
173 *Chem. C*, 121:4368–4383, 2017.
- 174 [14] S. Lorenz, M. Scheffler, and A. Groß. Descriptions of surface chemical reactions using a neural network  
175 representation. *Phys. Rev. B*, 73:115431, 2006.
- 176 [15] Y. Nam, J.-S. Chen, N. C. Pistenti, K. Wright, C. Delaney, D. Maslov, K. R. Brown, S. Allen, J. M.  
177 Amini, J. Apisdorf, K. M. Beck, A. Blinov, V. Chaplin, M. Chmielewski, C. Collins, S. Debnath, K. M.  
178 Hudek, A. M. Ducore, M. Keesan, S. M. Kreikemeier, J. Mizrahi, P. Solomon, M. Williams, J. D. Wong-  
179 Campos, D. Moehring, C. Monroe, and J. Kim. Ground-state energy estimation of the water molecule  
180 on a trapped-ion quantum computer. *npj Quantum Information*, 6:33, 2020.
- 181 [16] S. K. Natarajan and J. Behler. Neural network molecular dynamics simulations of solid-liquid interfaces:  
182 water at low-index copper surfaces. *Phys. Chem. Chem. Phys.*, 18:28704–28725, 2016.
- 183 [17] J. K. Nørskov, J. Rossmeisl, A. Logadottir, L. Lindqvist, J. R. Kitchin, T. Bligaard, and H. Jónsson. Origin  
184 of the overpotential for oxygen reduction at a fuel-cell cathode. *J. Phys. Chem. B*, 108:17886–17892,  
185 2004.
- 186 [18] R. Ouyang, S. Curtarolo, E. Ahmetcik, M. Scheffler, and L. M. Ghiringhelli. SISO: A compressed-sensing  
187 method for identifying the best low-dimensional descriptor in an immensity of offered candidates. *Phys.*  
188 *Rev. Materials*, 2:083802, 2018.
- 189 [19] S. Sakong and A. Groß. The electric double layer at metal-water interfaces revisited based on a charge  
190 polarization scheme. *J. Chem. Phys.*, 149:084705, 2018.
- 191 [20] S. Sakong, D. Mahlberg, T. Roman, M. Pandey, and A. Groß. Influence of local inhomogenities and the  
192 electrochemical environment on the oxygen reduction reaction on Pt-based electrodes: A DFT study. *J.*  
193 *Phys. Chem. C*, 124:27604–27613, 2020.
- 194 [21] S. Schnur and A. Groß. Properties of metal-water interfaces studied from first principles. *New J. Phys.*,  
195 11:125003, 2009.
- 196 [22] Z. W. Seh, J. Kibsgaard, C. F. Dickens, I. Chorkendorff, J. K. Nørskov, and T. F. Jaramillo. Combining  
197 theory and experiment in electrocatalysis: Insights into materials design. *Science*, 355:eaad4998, 2017.
- 198 [23] M. Sotoudeh, M. Dillenz, and A. Groß. Mechanism of magnesium transport in spinel chalcogenides. *Adv.*  
199 *Energy Sustainability Res.*, 2:2100113, 2021.
- 200 [24] M. Sotoudeh and A. Groß. Descriptor and scaling relations for ion mobility in crystalline solids. *Preprint*,  
201 <https://doi.org/10.21203/10.21203/rs.3.rs-309875/v2>, 2021.
- 202 [25] B. Wang and F. Zhang. Main descriptors to correlate structures with the performances of electrocatalysts.  
203 *Angew. Chem. Int. Ed.*, 60:DOI 10.1002/anie.202111026, 2021.