Machine Learning Prediction of Defect Structures in Amorphous Silicon Dioxide

Diego Milardovich[1], Markus Jech[1], Dominic Waldhoer[1,2], Al-Moatasem Bellah El-Sayed[1,3] and Tibor Grasser[1]
[1]Institute for Microelectronics, Technische Universität Wien,
Gußhausstraße 27–29, 1040 Vienna, Austria
[2]Christian Doppler Laboratory for Single-Defect Spectroscopy in Semiconductor Devices,
Gußhausstraße 27–29, 1040 Vienna, Austria
[3]Nanolayers Research Computing, Ltd.,
1 Granville Court, Granville Road, London N12 0HL, United Kingdom
E-mail: [milardovich | jech | waldhoer | el-sayed | grasser]@iue.tuwien.ac.at

ESSDERC 2021 - IEEE 51st European Solid-State Device Research Conference (ESSDERC) | 978-1-6654-3748-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/ESSDERC53440.2021.9631837

***Abstract*—Defects in insulators can have a highly detrimental impact on the performance of semiconductor devices. The study of defect formation in these amorphous insulating materials is a computationally challenging task, due to the relatively large model sizes required and their stochastic nature. Here, we propose a novel machine learning framework to predict the formation and structure of defects in amorphous materials. Our approach aims at significantly reducing the computational costs, while maintaining a high level of accuracy. We present the results of applying our workflow to the formation of hydroxyl E' center defects in amorphous silicon dioxide, which have recently been suggested to be responsible for random telegraph and 1/f noise, as well as the bias temperature instability. The process of predicting a particular defect structure is studied in full-detail and statistical results are presented for a testing data-set.**

I. Introduction

The study of point defects in amorphous materials plays a crucial role in the development of a wide range of modern microelectronic technologies (e.g., random access memory (RAM) devices) [1]–[3]. The majority of calculations required for defect studies are routinely performed with *ab initio* methods, in particular those based on density functional theory (DFT). However, these methods are computationally expensive and this disadvantage narrows their use to relatively small systems (on the order of a few hundred atoms) and particularly short time scales (on the order of tens of ps). This substantially limits their applicability for the calculation of large statistics, which is a crucial factor in the study of amorphous structures. In this context, computationally inexpensive machine learning (ML) based solutions offer a very promising approach for drastically reducing the computational efforts needed to study defects in complex atomistic structures. ML based solutions can either aid current *ab initio* based methods or even replace them entirely [4], [5].

In this work, a ML based solution was developed to predict the structure of hydroxyl E' center defects in amorphous silicon dioxide (a-SiO_2) models. The study of these defects is particularly important in the field of modern micro- and nano-electronic device reliability issues, since they are suspected to be responsible for bias temperature instability (BTI) and random telegraph noise (RTN) in MOS transistors [6]–[8].

We present a framework based on ML applied to predict a particular defect structure in a specific material. However, our approach is not limited to this particular test case and can be easily applied to other types of defects and/or other materials.

II. Methodology

The first step towards an efficient ML based prediction is a consistent training data-set. We prepared 16 different defect-free a-SiO_2 models, each containing a total of 216 atoms. These structures were created by utilizing the melt-quench technique [9] within the molecular dynamics engine LAMMPS [10] in conjunction with the classical force field ReaxFF [11]. Examples of these structures are shown in Fig. 1. A total number of 1271 unique hydroxyl E' centers were selectively created by placing a hydrogen atom in the vicinity of a bridging oxygen atom, followed by a subsequent geometrical relaxation. An elongated (weaker) Si-O bond will break and a hydroxyl (OH) group forms with a remaining defective silicon dangling bond in its neutral configuration, as shown in Fig. 2.

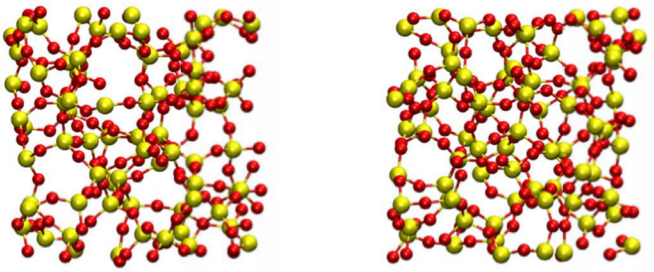
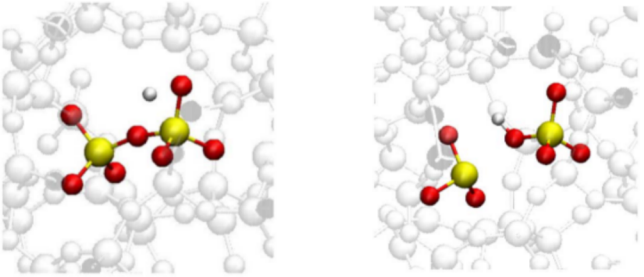


Fig. 1: Example of two a-SiO_2 structures containing 216 atoms each, used in this work to train and test the ML model. Yellow depicts silicon atoms, while red depicts oxygen atoms.



(a) (b)

Fig. 2: A hydrogen atom is placed in the vicinity of a bridging oxygen atom (a) and the structure is subsequently relaxed. Breaking of an elongated Si-O bond leads to the formation of a hydroxyl E' center (b), which is a silicon dangling bond facing a hydroxyl (OH) group.

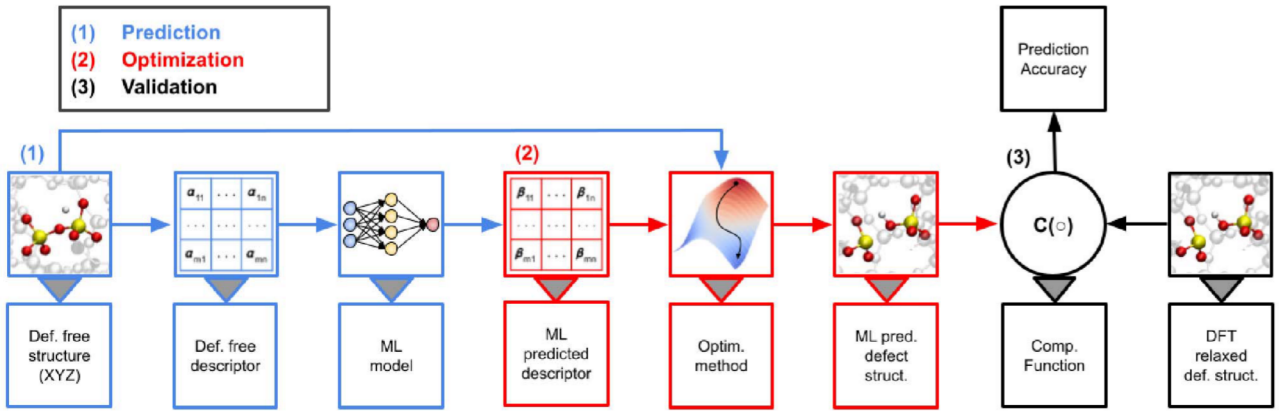


Fig. 3: Proposed workflow to predict the structure of defects in atomistic models. First, the local environment around the atom of interest in the defect free structure is represented by a descriptor. This matrix contains a vector for every atom near the site of interest and it is used as an input for a ML model, which then predicts the descriptor of a defect formed at that site (1). An optimization method is used to adjust the coordinates of the atoms in the defect free structure until their description matrix matches the predicted one (2). Finally, in order to validate the results, the ML predicted structure is compared to its equivalent DFT-relaxed structure (if available), by means of a comparison function (e.g., geometrical distance) (3).

Calculations were performed using the PBE functional [12] in the CP2K software package [13], which is a computationally expensive process. Typically, 4 nodes with 48 cores each require several hours to complete a single geometry relaxation. In total, the data-set is composed of 1271 hydroxyl E' centers and the respective defect free host structures. Once the data-set is created, the next step is to represent these structures in a way compatible with the selected ML algorithm. Such a mathematical representation of the structures (i.e., a vector or matrix) is called a *descriptor*. There is a wide range of available descriptors, ranging from local to global representations and including structural and/or electrostatic contributions, such as *atom-centered symmetry functions* (ACSF) [14], Ewald sum matrix [15] and *smooth overlap of atomic positions* (SOAP) [16]. Among the different options, we chose a structural representation using the SOAP descriptor, since this descriptor showed the best performance in our previous work, in which we used a ML model to predict the formation energies of hydroxyl E' centers in a-SiO_2 structures [17]. However, given the flexibility of our approach, any other descriptor, even individually designed ones, can be used within the proposed workflow.

Out of the 16 a-SiO_2 structures, 15 are used to train the model and one is left for testing purposes. This translates into 1188 hydroxyl E' center structures in the training set and 83 in the testing set. The atoms within a 3 Å cutoff radius at every defect site, which on average contains about 9 atoms, are used to construct a local descriptor of the defect environment, utilizing the SOAP descriptor, as implemented in the Python package DScribe [18]. To analyze the data, we employ a kernel ridge regression (KRR) model [19], as implemented in the scikit-learn package [20]. This is a cost-efficient model, which performed well in our previous study [17]. In this work, we extend its application to the prediction of defect structures, instead of just their formation energies.

Our ML model is trained to find the relationship between the SOAP descriptions of the hydroxyl E' centers and their host defect free a-SiO_2 structures. Once the ML model is trained, it is capable of predicting the formation of defects and their structural properties in new structures. The procedure is as follows:

1) *Prediction*: The environments of the surrounding atoms of interest in the defect-free structure are represented with the SOAP descriptor. This matrix is used as an input for the ML model, to predict the SOAP descriptor of the resulting defect complex.
2) *Optimization*: An optimization method is used to adjust the coordinates of the atoms in the defect free structure until its descriptor matrix matches the descriptor matrix predicted by the ML model for the defect structure. In this particular application, the Nelder-Mead method, with a convergence criteria of 5% of the initial loss function value.
3) *Validation*: The final step of the process is to validate the results. This is done by comparing the ML predicted defect structure with the equivalent target structure produced by DFT relaxation. We use the geometrical distance between the ML predicted and DFT relaxed defect site as a measure for the prediction quality.

The entire proposed workflow is graphically summarized in Fig. 3. As stated above, the same generic workflow could be used in the prediction of other defects in new materials.

III. Results and Discussion

In order to provide a clear illustration of the application of our proposed framework, the workflow shown in Fig. 3 is applied to the prediction of a particular defect, namely a hydroxyl E' center in an initially defect-free a-SiO_2 structure.

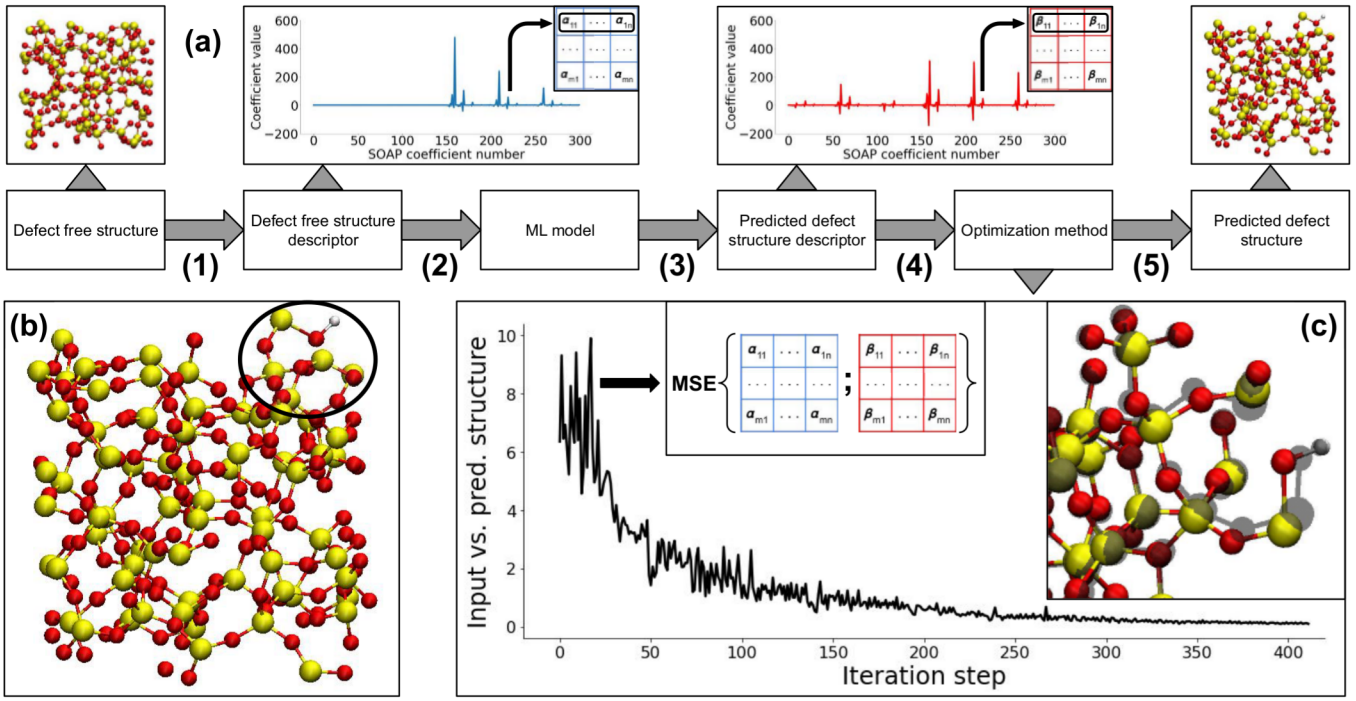


Fig. 4: (a) Prediction of the formation of a hydroxyl E' center in an a-SiO_2 structure. The local environments of the atoms around the oxygen atom of interest are described with a matrix of SOAP descriptors (1). This matrix is used as input for our ML model (2), to predict the SOAP description matrix of the defective site (3). An optimization method is used to adjust the coordinates of the atoms in the input structure, in order to minimize the mean squared error (MSE) between the optimized and predicted SOAP matrices, and hence forms the defect in the a-SiO_2 structure. The bottom right figure shows how the difference between the initial and predicted defect structure descriptor matrices reduces as the optimization progresses. The final result (5) is the predicted defect structure. (b) Zoom to the predicted structure. (c) Superposition between the ML predicted defect structure (color) and its DFT-relaxed equivalent (shadow) around the defect site.

Details of our application and the individual steps are shown in Fig. 4. A bridging O atom in the top right of the structure is selected as the precursor configuration for the hydroxyl E' center. Subsequently, the SOAP descriptor matrix of the surroundings of the O atom is computed with the parameters: $r_{cut} = 3.0$ Å and $n_{max} = l_{max} = 4$, as shown in Fig. 5.
The resulting matrix is used as an input for our ML model, which was previously trained with the training data-set. Our ML model predicts the expected SOAP description matrix of the defective structure at the site of interest.
Once this prediction is made, a loss function is defined as the mean squared error (MSE) between the predicted SOAP and the defect free structure with one additional H atom placed in the direct vicinity, 1 Å, of the bridging atom, to form the defect. An optimization method is used to adjust the coordinates of the respective atoms in the input structure until this loss function reduces below a certain threshold value. The optimization method selected in this case is the Nelder-Mead method, implemented in the Python package SciPy [21]. However, given the flexibility of our approach, other optimization methods, including optional bounds or (non-) linear constraints can be used instead. The final result is the predicted structure of the defect, as shown in the bottom-left of Fig. 4.

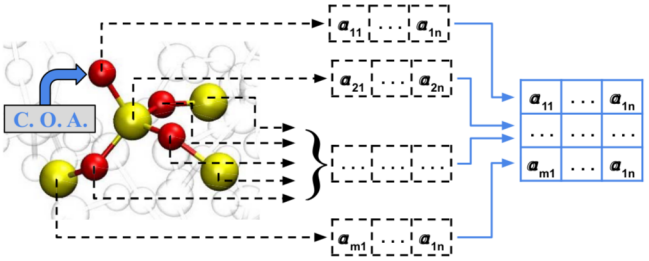


Fig. 5: Defect free structure descriptor matrix. Each row of this matrix is the SOAP vector representation of the local environment of a specific atom. Only those atoms within a certain cutoff distance from the central oxygen atom (C. O. A.) of interest, are considered.

The complete prediction and optimization process took 0.5 seconds on a typical desktop machine (Intel Core i7 2.2 GHz and 8 GB of RAM).
Given the stochastic nature of the amorphous network, the accuracy and efficiency of our proposed framework must be analyzed on a statistical scale. We therefore benchmarked it against a full set of data of an a-SiO_2 model, the testing data-set. Nevertheless, the respective DFT calculations are available and were used as a reference by computing the distance vector of atoms within 6 Å around the defect site. The results are

scaled w.r.t. the number of atoms involved, in order to obtain a normalized quantity.

The loss function to be minimized in the optimization process was defined as the MSE between the SOAP matrices of the input and predicted structures (loss function A)[1]. A second loss function was defined by adding a penalty if the distance between the hydrogen and central oxygen atoms deviated from 1 Å (loss function B)[2]. The process of predicting the hydroxyl E' center defect structures in the testing data-set was repeated under identical conditions for both loss functions.

The results can be seen in Fig. 6, which shows the distribution of the deviations of the ML based approach compared to the DFT results across the testing data-set. For the 83 structures in the testing set, the average distance between the ML and DFT structures is 0.461 Å/atom for loss function A and 0.285 Å/atom for loss function B. This shows that the framework works without any user knowledge about the system. However, with some detailed information and knowledge, the individual modules can be fine-tuned to provide even more accurate predictions.

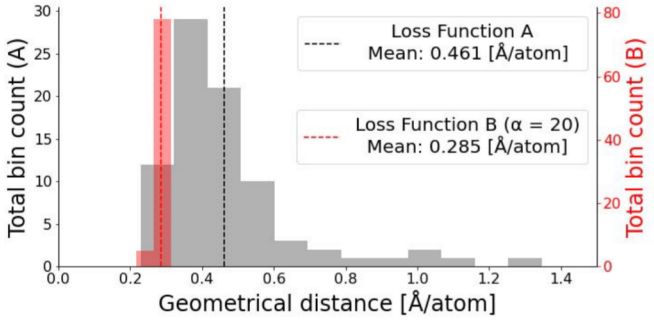


Fig. 6: Distribution of the deviations of the ML predicted defect structures compared to its equivalent DFT-relaxed structures, together with its mean value. Atoms within a 6.0 Å cutoff distance from the central oxygen atom are considered.

IV. Conclusions

There is a wide range of applications in which machine learning (ML) based techniques can provide computationally inexpensive, but accurate solutions, as shown by several examples in the literature [4], [5], [15]. In this work, a solution was developed to study the formation of hydroxyl E' center defects in amorphous silicon dioxide (a-SiO_2). The results clearly demonstrate a competitive level of accuracy, while being computationally inexpensive when compared to DFT.

The presented approach benefits from being highly modular, meaning that its components are interchangeable with other descriptors, ML models and optimization algorithms, depending on the individual problem. Hence, it can be easily extended and adapted to other defect species in a-SiO_2, or even different materials, such as hafnium dioxide (HfO_2).

Our approach is of particular importance for defect studies in situations where the application of DFT and other ab initio methods are too expensive, such as defect studies in large simulation cells and investigations on a large statistical scale. Furthermore, applications with a demand of on-the-fly calculations of certain data, such as kinetic Monte Carlo (kMC) simulations [22], benefit from the presented framework. Currently, reaction rates in such simulations have to be predefined. However, with our approach, structural and energetic information can be calculated almost instantaneously.

[1] $\mathrm{Loss_A} = \mathrm{MSE}(\mathrm{SOAP_{input}}, \mathrm{SOAP_{predicted}})$

[2] $\mathrm{Loss_B} = \mathrm{MSE}(\mathrm{SOAP_{input}}, \mathrm{SOAP_{predicted}}) + \alpha|\mathrm{OH_{BL}} - 1\ \text{Å}|$

V. Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 871813, within the framework of the project Modeling Unconventional Nanoscaled Device FABrication (MUNDFAB). The authors acknowledge support from the Christian Doppler Laboratory for Single-Defect Spectroscopy in Semiconductor Devices and the Vienna Scientific Cluster (VSC).

References

[1] A.-M. El-Sayed et al., "Hydrogen-Induced Rupture of Strained SiO Bonds in Amorphous Silicon Dioxide", Phys. Rev. Lett. 114, 115503.

[2] M. Jech et al., "Ab initio treatment of silicon-hydrogen bond rupture at Si/SiO_2 interfaces", Phys. Rev. B 100, 195302 (2019).

[3] P. V. Sushko et al., "Structure and properties of defects in amorphous silica: new insights from embedded cluster calculations", J. Phys.: Condens. Matter 17 S2115 (2005).

[4] S. Kiyohara, et al., "Prediction of interface structures and energies via virtual screening", Sci. Adv. 2016; 2:e1600746.

[5] A. Seko, et al., "Representation of compounds for machine-learning prediction of physical properties", Phys. Rev. B 95, 144110 (2017).

[6] T. Grasser et al., "On the Microscopic Structure of Hole Traps in pMOSFETs", IEEE Int. Electron Devices Meeting (2014) 21.1.1-21.1.4

[7] Y. Wimmer et al., "Role of hydrogen in volatile behaviour of defects in SiO_2-based electronic devices". Proc. R. Soc. A 472: 20160009.

[8] W. Goes et al., "Identification of oxide defects in semiconductor devices: A systematic approach linking DFT to rate equations and experimental evidence", Microelectronics Reliability, 87 (2018) 286-320.

[9] A.-M. El-Sayed et al., "Identification of intrinsic electron trapping sites in bulk amorphous silica from ab initio calculations", Microelectronic Engineering 109 (2013) 68–71.

[10] S. Plimpton, "Fast parallel algorithms for short–range molecular dynamics", J. of Comp. Phys., vol 117, p 1–19 (1995).

[11] J. C. Fogarty et al., "A reactive molecular dynamics simulation of the silica-water interface", J. Chem. Phys. 132, 174704 (2010).

[12] J. P. Perdew et al., "Generalized Gradient Approximation Made Simple", Phys. Rev. Lett. 77, 3865 (1996).

[13] J. VandeVondele, et al., "Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach", Comput. Phys. Commun. 167, 103 (2005).

[14] J. Behler, "Atom-centered symmetry functions for constructing high-dimensional NN potentials", J. Chem. Phys. 134, 074106 (2011).

[15] F. Faber et al., "Crystal structure representations for machine learning models of formation energies", International Journal of Quantum Chemistry 2015, 115, 1094–1101.

[16] A. P. Bartók, et al., "On representing chemical environments", Phys. Rev. B 87, 184115 (2013).

[17] D. Milardovich, et al., "Machine learning prediction of defect formation energies in a-SiO_2", Int. Conf. on Sim. of Semic. Proc. and Dev. (SISPAD) 15-2 (2020).

[18] L. Himanen et al., "DScribe: Library of descriptors for machine learning in materials science", Comp. Phys. Communications 247 (2020) 10694.

[19] Vovk V. (2013) Kernel Ridge Regression. In: Schölkopf B., Luo Z., Vovk V. (eds) Empirical Inference. Springer, Berlin, Heidelberg.

[20] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python", J. of Machine Learning Research 12 (2011) 2825-2830.

[21] P. Virtanen et al., "SciPy 1.0: fundamental algorithms for scientific computing in Python", 17 (2020) 261–272.

[22] A. La Magna et al., "Simulation of the Growth Kinetics in Group IV Compound Semiconductors", Phys. Status Solidi A 2019, 216, 1800597.