DISSERTATION

Bridging Molecular Dynamics and Neuroscience: Machine Learning for Efficient Simulations and Biomedical Signal Processing

zur Erlangung des akademischen Grades

Doktor der technischen Wissenschaften

eingereicht an der Technischen Universität Wien
Fakultät für Elektrotechnik und Informationstechnik
von

Ing. Diego Milardovich

[redacted]

unter der Betreuung von

Univ.Prof. Dipl.-Ing. Dr.techn. Tibor Grasser

Wien, im Jänner 2024

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Abstract

The pace of development in the microelectronics field is among the fastest across all industries. In line with Moore's Law, the number of transistors in a single integrated circuit has approximately doubled every two years since the early 1970s and, over the past two decades, the transistor count in commercially available central processing units (CPUs) has increased by approximately 1,000 times. This exponential growth has been allowed by a significant reduction in the size of individual components with each new generation. Due to this reduction in size, the necessity of an atomistic description becomes crucial in microelectronics, as the regions of interest contain only a few atoms. Consequently, accurate and efficient atomic-level simulations have become increasingly essential in the design and development processes.

Molecular dynamics (MD) has emerged as a widely used tool to meet a wide range of the atomistic simulation requirements of this industry. By evaluating the interactions among atoms in a system over time, MD enables the prediction of certain macroscopic properties of materials and sheds light on complex atom-level processes, such as the formation of defects that can impact the reliability of microelectronic devices. However, researchers employing MD face a long-standing challenge: the trade-off between accuracy and efficiency. On one hand, *ab initio* methods, such as density functional theory (DFT), provide high accuracy; nonetheless, their computational costs are prohibitively high, limiting their application to relatively small systems ($<10^3$ atoms) and short simulation times ($\sim 10^1$ ps). On the other hand, empirical potentials, which are parameterized mathematical models of the potential energy surface (PES), require fewer computational resources. However, their accuracy is not sufficient for many applications and they require expert parameterization for every new system of interest.

This thesis aims to develop machine learning (ML)-based solutions that enhance the trade-offs between accuracy and efficiency in MD simulations. As part of this work, four novel ML interatomic potentials were developed: (I) to model amorphous silicon-nitride, a material with applications ranging from dielectrics in microelectronic devices to the Space Shuttle's armor; (II) to simulate neutron irradiation effects in silicon-germanium, relevant to the study of microelectronic components exposed to radiative environments, like nuclear power plants and outer space; (III) to investigate the oxidation process of silicon, critical to the semiconductor industry; and (IV) to model amorphous silicon-dioxide, extensively used as an insulator in microelectronic devices. Moreover, two alternative ML-based approaches were developed to predict the formation energy and atomic structure of defects in amorphous materials.

During the aforementioned developments, two main bottlenecks were identified: The time required for building the training datasets, and the inadequate optimization of the hardware in which the models are deployed. To address these issues, the following solutions were investigated: (A) an active learning algorithm to expedite the training of ML interatomic potentials, and (B) the utilization of a novel neuromorphic-based hardware accelerator for conducting MD simulations. The latter solution, based on the work of Mo *et al.*, was successfully employed to simulate germanium-based memristors composed of over 100,000 atoms.

In this thesis, a dual-component approach, consisting of a descriptor and an ML model, has been successfully applied to the development of ML interatomic potentials for MD simulations. Building upon these advancements, a similar approach has been extended to address a distinct problem: the automatic annotation of motor evoked potentials (MEPs). These are electromyography (EMG) recordings obtained from muscles in response to transcranial magnetic stimulation (TMS), which play a crucial role in the treatment of neurological disorders, such as multiple sclerosis. MEPs are typically characterized by their latency, defined as the time interval between the TMS stimulation and the onset of the MEP, and the treatment of a single patient might require the analysis of thousands of MEPs. This is frequently done manually by a clinician, since the high variability and low signal-to-noise ratio of these signals make their automatic analysis a hard problem to solve. To mitigate these issues, an algorithm inspired by the developments in ML interatomic potentials is presented in this thesis. By automating and accelerating the MEPs annotation process, this algorithm reduces the workload for clinicians and minimizes the potential for human errors. Moreover, its low computational costs enable on-the-fly signal analysis, which opens up possibilities for future applications, such as closed-loop ML-based brain stimulation protocols.

Results show that the four ML interatomic potentials developed in this thesis presented a mean absolute error (MAE) $<$10 meV/atom when compared to *ab intio* methods in energy calculations, while being between three to four orders of magnitude faster. The active learning algorithm presented in this work was sucessfully employed in the development of the ML interatomic potential to study the oxidation of silicon, while the neuromorphic based hardware accelerator allowed for MD simulations about two orders of magnitude faster than on traditional scientific clusters. On the other hand, the transdiciplinary adaptation of these solutions resulted in annotations of MEPs with a MAE of 0.5 ms, making it virtually indistinguishible from the annotations of a medical expert.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Kurzfassung

Die Mikroelektronikbranche entwickelt sich mit rasanter Geschwindigkeit. Gemäß dem Mooreschen Gesetz hat sich die Anzahl der Transistoren in einem einzigen integrierten Schaltkreis seit den frühen 1970er Jahren ungefähr alle zwei Jahre verdoppelt. In den letzten zwei Jahrzehnten hat sich die Anzahl der Transistoren in kommerziell erhältlichen Mikroprozessoren um etwa das 1,000-fache erhöht. Diese exponentielle Wachstum wurde durch eine signifikante Reduktion der Größe einzelner Komponenten in jeder neuen Generation ermöglicht. Aufgrund dieser signifikanten Größenreduktion wird die Notwendigkeit einer atomistischen Beschreibung in der Mikroelektronik entscheidend, da die relevanten Bereiche nur eine geringe Anzahl von Atomen enthalten. Infolgedessen sind genaue und effiziente Simulationen auf atomarer Ebene zunehmend wichtiger geworden für den Entwurfs- und Entwicklungsprozess.

Molekulardynamik (MD) hat sich als ein weit verbreitetes Werkzeug etabliert, um eine Vielzahl der atomistischen Simulationserfordernisse dieser Branche zu erfüllen. Durch die Bewertung der Wechselwirkungen zwischen Atomen in einem System im Laufe der Zeit, ermöglicht MD die Vorhersage bestimmter makroskopischer Eigenschaften von Materialien und gibt Einblicke in komplexe atomare Prozesse, wie die Bildung von Defekten, die die Zuverlässigkeit von mikroelektronischen Geräten beeinträchtigen können. Forscher, die MD verwenden, stehen jedoch vor einer Herausforderung: dem Kompromiss zwischen Genauigkeit und Effizienz. Einerseits bieten *ab initio* Methoden wie Dichtefunktionaltheorie (DFT) eine hohe Genauigkeit; ihre Berechnungskosten sind jedoch prohibitiv hoch und beschränken ihre Anwendung auf relativ kleine Systeme ($<10^3$ Atome) und kurze Simulationszeiten ($\sim 10^1$ ps). Andererseits erfordern empirische Potentiale, die parametrisierte mathematische Modelle der potentiellen Energieoberfläche sind, weniger Rechenressourcen. Ihre Genauigkeit ist jedoch für viele Anwendungen nicht ausreichend und sie erfordern eine fachkundige Parameterisierung für jedes neue interessierende System.

Diese Dissertation konzentriert sich darauf, Maschinelles Lernen (ML) basierte Lösungen zur Verbesserung des Kompromisses zwischen Genauigkeit und Effizienz in MD Simulationen zu entwickeln. Im Rahmen dieser Arbeit wurden vier neue ML interatomare Potentiale entwickelt: (I) zur Modellierung von amorphem Siliziumnitrid, einem Material mit Anwendungen von der Dielektrikschicht von mikroelektronischen Geräten bis zur Panzerung des Space Shuttles; (II) zur Simulation von Neutronenbestrahlungseffekten in Siliziumgermanium, relevant für die Untersuchung mikroelektronischer Bauteile, die Strahlungsumgebungen wie in Kernkraftwerken und dem Weltraum ausgesetzt sind; (III) zur Untersuchung des Oxidationsprozesses von

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

Silizium, der für die Halbleiterindustrie von entscheidender Bedeutung ist; (IV) zur Modellierung von amorphem Siliziumdioxid, das als Isolator in mikroelektronischen Geräten weit verbreitet ist. Darüber hinaus wurden zwei alternative ML basierte Ansätze entwickelt um die Bildungsenergie und atomare Struktur von Defekten in amorphen Materialien vorherzusagen.

Während der genannten Entwicklungen wurden zwei thematische Hauptengpässe identifiziert: die Zeit, die für den Aufbau der Trainingsdatensätze benötigt wird, und die unzureichende Optimierung der Hardware, auf der die Modelle ausgeführt werden. Um diese Probleme zu lösen, wurden folgende Lösungsansätze untersucht: (A) ein aktiver Lernalgorithmus, um das Training von ML interatomaren Potentialen zu beschleunigen, und (B) die Verwendung eines neuromorphen Hardware Beschleunigers für MD Simulationen. Die letztere Lösung, basierend auf der Arbeit von Mo *et al.*, wurde erfolgreich eingesetzt, um einen Germanium-basierten Memristor mit über 100,000 Atomen zu simulieren.

In dieser Dissertation wurde ein zweikomponentiger Ansatz, bestehend aus einem Deskriptor und einem ML Modell, erfolgreich auf die Entwicklung von ML interatomaren Potentialen für MD Simulationen angewendet. Aufbauend auf diesen Erfolgen wurde ein ähnlicher Ansatz erweitert, um ein anderes Problem zu lösen: die automatische Annotation von motorisch evozierten Potenzialen (MEPs). Dabei handelt es sich um Elektromyographie Aufzeichnungen, die von Muskeln als Reaktion auf transkranielle Magnetstimulation (TMS) erhalten werden und eine wichtige Rolle bei der Behandlung neurologischer Erkrankungen wie Multipler Sklerose spielen. MEPs werden in der Regel anhand ihrer Latenz charakterisiert, definiert als Zeitintervall zwischen der TMS Stimulation und dem Beginn der MEPs, und die Behandlung eines einzigen Patienten erfordert möglicherweise die Analyse Tausender MEPs. Dies geschieht häufig manuell durch einen Mediziner, da die hohe Variabilität und das geringe Signal-Rausch-Verhältnis dieser Signale ihre automatische Analyse zu einem schwer zu lösenden Problem machen. Um diese Probleme zu mildern, wird in dieser Arbeit ein Algorithmus vorgestellt, der von den Entwicklungen von ML interatomaren Potentialen inspiriert ist. Durch Automatisierung und Beschleunigung des MEP Annotationsprozesses reduziert dieser Algorithmus die Arbeitsbelastung für Mediziner und minimiert das Potenzial menschlicher Fehler. Darüber hinaus ermöglichen seine geringen Berechnungskosten eine Echtzeitanalyse der Signale, was Möglichkeiten für zukünftige Anwendungen eröffnet.

Die Ergebnisse zeigen, dass die vier vorgestellten ML interatomaren Potentiale einen mittleren absoluten Fehler (MAE) $<$10 meV/Atom aufweisen, wenn sie mit *ab initio* Methoden in Energieberechnungen verglichen werden, während sie gleichzeitig zwischen drei bis vier Größenordnungen schneller sind. Der in dieser Arbeit vorgestellte aktive Lernalgorithmus wurde erfolgreich bei der Entwicklung des ML interatomaren Potentials zur Untersuchung der Oxidation von Silizium eingesetzt, während der auf neuromorpher Hardware basierende Beschleuniger für MD Simulationen etwa zwei Größenordnungen schneller ist als auf herkömmlichen wissenschaftlichen Clustern. Andererseits führte die transdisziplinäre Anpassung dieser Lösungen zu Annotationen von MEPs mit einer MAE von 0.5 ms, wodurch sie praktisch nicht von den Annotationen eines medizinischen Experten zu unterscheiden sind.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Resumen

La velocidad del desarrollo en la industria de la microelectrónica se encuentra entre las más rápidas en todos los sectores. Siguiendo la Ley de Moore, el número de transistores en circuitos integrados se ha duplicado aproximadamente cada dos años desde principios de la década de 1970 y, en las últimas dos décadas, la cantidad de transistores en procesadores disponibles comercialmente ha aumentado aproximadamente 1,000 veces. Este crecimiento exponencial ha sido producto de una reducción significativa en el tamaño de los componentes microelectrónicos con cada nueva generación. Debido a esta reducción significativa en tamaño, la necesidad de una descripción atomística se vuelve crucial en microelectrónica, ya que las regiones de interés contienen solo un pequeño número de átomos. En consecuencia, las simulaciones precisas y eficientes a nivel atómico se han vuelto cada vez más esenciales en los procesos de diseño y desarrollo.

La técnica de dinámica molecular (DM) ha surgido como una herramienta ampliamente utilizada para satisfacer muchos de los requisitos de simulación de esta industria. Al evaluar las interacciones entre los átomos de un sistema a lo largo del tiempo, la DM permite predecir algunas propiedades macroscópicas de los materiales y esclarece procesos complejos a nivel atómico, como la formación de defectos que pueden afectar la confiabilidad de los dispositivos microelectrónicos. Sin embargo, los investigadores que emplean simulaciones de DM se enfrentan a un desafío de larga data: el compromiso entre precisión y eficiencia. Por un lado, los métodos denominados *ab initio*, como la teoría del funcional de la densidad (TFD), proporcionan alta precisión; sin embargo, sus costos computacionales son prohibitivamente altos, limitando su aplicación a sistemas relativamente pequeños ($<10^3$ átomos) y tiempos de simulación cortos ($\sim 10^1$ ps). Por otro lado, los potenciales empíricos, que son modelos matemáticos parametrizados de la superficie de energía potencial, requieren menos recursos computacionales. Sin embargo, su precisión no es suficiente para muchas aplicaciones, y requieren de la parametrización de un experto para cada nuevo sistema de interés.

Esta tesis se centra en el desarrollo de soluciones basadas en machine learning (ML) para proporcionar mejores compromisos entre precisión y eficiencia en las simulaciones de DM. Como parte de este trabajo, se desarrollaron cuatro nuevos potenciales interatómicos basados en ML: (I) para modelar nitruro de silicio amorfo, un material con aplicaciones que van desde dieléctrico en dispositivos microelectrónicos hasta blindaje del Space Shuttle; (II) para simular los efectos de la radiación de neutrones en silicio-germanio, relevante para el estudio de componentes micro-

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

electrónicos expuestos a entornos radiactivos, como plantas de energía nuclear y el espacio exterior; (III) para investigar el proceso de oxidación del silicio, fundamental para la industria de semiconductores; (IV) para modelar dióxido de silicio amorfo, utilizado extensivamente como aislante en dispositivos microelectrónicos. Además, se desarrollaron dos técnicas alternativas basadas en ML para predecir la energía de formación y la estructura atómica de defectos en materiales amorfos.

Durante los desarrollos mencionados anteriormente, se identificaron dos obstáculos principales a nivel computacional: el tiempo requerido para construir los datasets de entrenamiento, y la inadecuada optimización del hardware en el que se implementan los modelos. Para abordar estos problemas, se investigaron las siguientes soluciones: (I) un algoritmo de aprendizaje activo para acelerar el entrenamiento de potenciales interatómicos basados en ML y (II) la utilización de un acelerador de hardware basado en computación neuromórfica para realizar simulaciones de DM. Esta última solución, basada en el trabajo de Mo *et al.*, se empleó con éxito para simular memristores de germanio, compuestos por más de 100,000 átomos.

En esta tesis, se ha aplicado con éxito un enfoque dual, que consta de un descriptor y un modelo de ML, para el desarrollo de potenciales interatómicos para simulaciones de DM. Sobre la base de estos avances, un enfoque similar se ha extendido para abordar un problema distinto: la anotación automática de potenciales evocados motores (PEMs). Estos son registros electromiográficos obtenidos de los músculos en respuesta a la estimulación magnética transcraneal (EMT), que desempeñan un papel crucial en el tratamiento de trastornos neurológicos, como la esclerosis múltiple. Los PEMs se caracterizan típicamente por su latencia, definida como el intervalo de tiempo entre la EMT y el inicio del movimiento del músculo. El tratamiento de un solo paciente puede requerir el análisis de miles de PEMs. Esto se realiza frecuentemente de forma manual por un médico, ya que la alta variabilidad y la baja relación señal-ruido de estas señales dificultan su análisis automático. Para mitigar estos problemas, se presenta en esta tesis un algoritmo inspirado en los desarrollos en potenciales interatómicos basados en ML. Al automatizar y acelerar el proceso de anotación de PEMs, este algoritmo reduce la carga de trabajo para los médicos y minimiza la posibilidad de errores humanos. Además, el bajo costo computacional de este algoritmo permite el análisis de señales en vivo, lo que abre posibilidades para futuras aplicaciones, como protocolos automáticos de estimulación cerebral en lazo cerrado basados en ML.

Los resultados muestran que los cuatro potenciales interatómicos basados en ML presentaron un error absoluto medio (EAM) inferior a 10 meV/átomo en comparación con métodos de tipo *ab initio* en cálculos de energía, al mismo tiempo que son entre tres y cuatro órdenes de magnitud más rápidos. El algoritmo de aprendizaje activo presentado en esta tesis se utilizó con éxito en el desarrollo del potencial interatómico para estudiar la oxidación del silicio, mientras que el acelerador de hardware basado en computación neuromórfica permitió simulaciones de DM aproximadamente dos órdenes de magnitud más rápidas que en servidores para cálculos científicos tradicionales. Por otro lado, la adaptación transdisciplinaria de estas soluciones dio como resultado anotaciones de PEMs con un EAM de 0.5 ms, lo que las hace prácticamente indistinguibles de las anotaciones de un experto médico.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Acknowledgement

The elaboration of a doctoral thesis is ultimately a collective effort, and I consider this to be particularly true in this case, given the interdisciplinary nature of this research and the extensive collaborations with foreign research groups. Doing justice to everyone who contributed to this work would be impossible in a few lines. However, I will do my best to acknowledge those who made the most remarkable contributions.

First and foremost, I would like to express my sincere appreciation to my supervisor, Prof. Tibor Grasser. I am very grateful for the complete scientific freedom he granted me during the planning and execution of my research. Moreover, this thesis would not have been possible without the help from the members of the Density Functional Theory group led by Prof. Grasser at the Institute for Microelectronics of TU Wien: Dominic Waldhör, Dr. Markus Jech, Christoph Wilhelmer, and Lukas Cvitkovich. Their inputs and comments in our lengthy meetings certainly helped polish and improve my writing.

Furthermore, my deepest thanks to my colleagues Francio Rodrigues and Dr. Luiz Felipe Aguinsky, my fellow Latin American expats at the Institute for Microelectronics, for making the integration into this working environment a very easy task. Without their coffee connoisseurship, this research would most certainly not have been possible.

I would like to express my heartfelt gratitude to Dr. Anne Hemeryck and her Modélisation Multi-niveaux des Matériaux group at LAAS-CNRS. Dr. Hemeryck's kindness and unwavering support made my stay in France a truly pleasant experience. Additionally, the daily technical discussions with the entire group greatly enhanced the scientific value of my research during my time there. I extend my sincerest appreciation to Prof. Dr. Mathieu Luisier and his Computational Nanoelectronics group at ETH Zürich for their wholehearted support and warm welcome. My time in Switzerland was the most productive period of my doctoral pursuit, and I am immensely grateful for the opportunity I had to work with such a renowned and experienced group of researchers.

Before, during, and after my doctoral pursuit, I switched fields more than once. However, I was consistently fortunate to have the unwavering support of accomplished researchers and exceptional mentors. I would like to extend my heartfelt gratitude to Dr. Victor Hugo Souza from Aalto University in Finland, Prof. Dr. Nima Taherinejad from Heidelberg University in Germany, and Dr. Ganesh Sivaraman from Argonne National Laboratory in the USA. Their engagement, active support, and remarkable expertise in their respective fields have been indispensable in making this doctoral thesis a reality. Without their guidance and encouragement, this journey would have been insurmountable.

In a different vein, my support network for this enterprise began a long time ago at a very distant place, the National University of Rosario, in Argentina. I would not have made it to Vienna without the constant support and encouragement of my colleagues. I am especially grateful to Alfonso Rafel, Rodolfo Paganini, and Lucas Grippo. Their collaborative spirit and camaraderie have not only shaped my work but also fostered personal growth, making this achievement all the more meaningful.

I would like to thank my lifelong and new Austrian friends in Vienna, in particular to David Hobel, Jasmin Arnetzeder, Simon Srienc, and Luca Mak, who made settling in this new home and integrating into Austrian society incredibly smooth and easy. From Karmelitermarkt to Cuba and Iceland, they made this journey remarkably more interesting and exciting. I must also thank my colleagues and Argentine compatriots exiled in Europe: Lucas Pérez and Federico Blanc. Leaving home to undertake this journey was immensely easier knowing that our destinies were shared.

Last, but certainly not least, I would like to thank my family. To my father, Natalio, a fellow electrical engineer with whom I share a passion for science. His constant support has been the cornerstone of my academic journey. To my mother, Liliana, for her love, unwavering support, and encouragement for me to always follow my passions, even when it meant prolonged periods of separation. To my brother, Sergio, regardless of the geographical distance, I always knew I could count on him in this journey. My accomplishments would not have been possible without their belief in me and their steadfast presence throughout this demanding endeavor.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Contents

Abstract i

Kurzfassung iii

Resumen v

List of Figures xi

List of Tables xxi

1 Introduction 1
- 1.1 Research Goals 3
- 1.2 Thesis Outline 4

2 State-of-the-Art and Methods 5
- 2.1 The Born-Oppenheimer Approximation 5
- 2.2 Molecular Dynamics Simulations 7
  - 2.2.1 Thermostats 8
  - 2.2.2 Barostats 9
- 2.3 Classical Potentials 10
  - 2.3.1 Lennard-Jones 11
  - 2.3.2 Tersoff 11
  - 2.3.3 ReaxFF 12
- 2.4 Machine Learning Interatomic Potentials 13
  - 2.4.1 Chemical Environment Descriptors 14
    - 2.4.1.1 Geometrical Descriptors 15
    - 2.4.1.2 Behler-Parrinello Descriptor 16
    - 2.4.1.3 Smooth Overlap of Atomic Positions (SOAP) 16
  - 2.4.2 The Gaussian Approximation Potential Method 18
  - 2.4.3 Artificial Neural Networks 23
  - 2.4.4 Moment Tensor Potentials 31

3 Gaussian Approximation Potential Results 33
- 3.1 GAP for Silicon-Nitride (Si_3N_4) 33
- 3.2 GAP for Silicon-Germanium (SiGe) 44
- 3.3 GAP for the Oxidation of Silicon (Si) 49

3.4 GAP for Silicon-Dioxide (SiO_2) 54

4 Machine Learning Alternatives to Interatomic Potentials 61
4.1 Prediction of Defect-Formation Energies 61
4.2 Prediction of Defect Structures 70

5 Neuromorphic Hardware Accelerator for Machine Learning-Based Molecular Dynamics 75
5.1 Framework 75
5.2 Use Case 77
5.3 Results 81

6 Transdisciplinary Application to Biomedical Signals Processing 85
6.1 Introduction and Motivation 86
6.2 Methodology 87
6.3 Results and Discussion 91

7 Summary and Outlook 99

Bibliography 101

List of Publications 117

Curriculum Vitae 119

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

List of Figures

2.1 PES for a hypothetical reaction system with two coordinates. (CC BY-NC; Ümit Kaya via LibreTexts) 6

2.2 Workflow to compute the energy of an atomic configuration with an ML model. The input atomic configuration is first represented by means of a descriptor. This descriptor is subsequently used as input to the ML model, which produces the energy as its output. 14

2.3 Two distinct atomic structures, constructed with the same set of six interatomic distances. 15

2.4 Left: Illustration of the process of extracting a SOAP matrix P for a single GB. Given a single atom in the GB, a Gaussian particle density function is placed at the location of each atom within a local environment sphere around the atom. Next, the total density function produced by neighbors is projected into a spherical basis consisting of radial basis functions and the spherical harmonics, as shown in the *boxed* region. Each basis function produces a single coefficient p_i in the SOAP vector $\vec{p}$ for the atom, the magnitude of which is represented in the figure by the colors of the arrays. Once a SOAP vector is available for all Q atoms in the GB, they are collected into a single matrix P that represents the GB. Figure reproduced from npj Computational Materials (2017) 3, 29, with permission from the authors. Right: Example of a trigonometric f_{cut} function for different cutoff parameters. 17

2.5 The three basic steps in GAP modeling: Representing the atomic structures by means of atomic descriptors, GPR and application. Original figure published under the CC BY 4.0 license. Credit: Deringer *et al.*, Chem. Rev. 121, 10073 (2021). 18

2.6 Descriptors for atomic structures commonly used in the GAP method: (a) conventional N-body terms (e.g., 2-body: interatomic distance), (b) General 3-body descriptor, and (c) SOAP. Modified from original figure published under the CC BY 4.0 license. Credit: Deringer *et al.*, Chem. Rev. 121, 10073 (2021). 19

2.7 Schematic of a GAP used as a surrogate model to a PES. The input atomic configuration (red) is compared to every atomic configuration available in the training data-set (blue) by means of a similarity measure or kernel function (K). The energy of the input atomic configuration is then computed as the sum of these similarity measures, weighed by the vector α. The values of this vector are learned in the training process. 20

2.8 Basic elements of the GPR: (1) Observations, in the form of a training dataset composed of multiple data points, (2) Basis functions, centered at each data point, and (3) Estimation, produced by fitting a set of coefficients c_m to the observations. Original figure published under the CC BY 4.0 license. Credit: Deringer *et al.*, Chem. Rev. 121, 10073 (2021). 21

2.9 Effects of the kernel length scale, σ_{length} on the GPR for different types of data. The effects of using a small (left) and large (right) σ_{length} are illustrated for learning from function values (a) and from derivatives of the function values (b). The lengths of the basis functions are represented by a solid green line in each case. The overfitting produced by the small σ_{length} are clear by the poor fitting results obtained on the left traces. Original figure published under the CC BY 4.0 license. Credit: Deringer *et al.*, Chem. Rev. 121, 10073 (2021). 22

2.10 Three main components for GAP: (1) training dataset, (2) a suitable mathematical representation of the atomic environments, and (3) the Gaussian process regression itself. By sampling the trajectories produced by running MD with the GAP and adding them to the training dataset together with their energies and forces, an iterative improvement on the accuracy of the potential is allowed. Original figure published under the CC BY 4.0 license. Credit: Deringer *et al.*, Chem. Rev. 121, 10073 (2021). 23

2.11 Schematic of a typical neural network used as a surrogate model of a PES. Red: input layer, green: hidden layer(s) (in this case, one), blue: output layer. The input to the neural network is a local atomic environment descriptor and the output is its local energy contribution. The information is processed in a feed-forward manner, as every neuron takes as inputs the outputs of the previous neurons and its output is a non-linear function of them. The connections between neurons are named *synapses* and their values are set during the training process. . 23

2.12 Illustration of a three-dimensional feed-forward neural network featuring dual hidden layers, each comprising four neurons. The connections between neurons are depicted by arrows, symbolizing the adjustable parameters (weights) crucial for the network's functioning. To enhance clarity, bias weights have been omitted from the visualization. 25

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

2.13 Architecture of the second generation neural network interatomic potential, applied to a system with N atoms. The output corresponds to the total energy (E), which is the aggregate of local atomic energy contributions (E_i). Each local atomic energy contribution is determined by the output of a dedicated neural network. The input for each atomic neural network consists of a vector describing the local chemical environment of atom i, computed from the coordinates of the atoms in the system (R_1, R_2, ..., R_N). 26

2.14 Schematic of a fourth generation neural network interatomic potential for an atomic system containing N_a and N_b atoms of two different elements. The total potential energy, E_total is the sum of the local atomic contributions, E_short and the long-range contributions, E_elec. Each local atomic contribution is computed by a neural network from a descriptor of the local atomic environments, G. A second neural network is used to compute the atomic electronegativities from the same descriptors, from which the atomic charges are computed using a charge equilibration method. The long-range contributions are calculated based on the resulting atomic charges. Original figure reproduced from Nature Communications vol. 12 no. 398 under Creative Commons Attribution 4.0 International License. Credits: Tsz Wai Ko *et al.* 28

2.15 Basic MPNN applied to an anthracene molecule, showcasing (a) the initial configuration with yellow hydrogen atoms and blue carbon atoms, (b) the first iteration, and (c) the second iteration. In this model, a straightforward update mechanism is employed, where the hue of each node is an average of its original hue, its current hue, and the hues of its neighboring nodes. The feature vectors for the edges (represented by straight black lines) are omitted for clarity. Red arrows highlight the propagation of information regarding the initial state of node 18 through the network. Nodes whose current state have been influenced by the initial state of node 18 are denoted by red circles. Reprinted from Journal of Chemical Physics 152, 050902 (2020), with the permission of AIP Publishing. Credits: Tim Mueller *et al.* 30

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

3.1 Schematic of the iterative re-training algorithm used to develop the ML interatomic potential for Si_3N_4 (left) and results obtained over time for subsequent generations (right). The ML interatomic potential is used to run MD; thereby generating a trajectory of atomic configurations. The last snapshots of this trajectory are collected, their energies, forces and/or stress tensors are recalculated with DFT and then added to the training dataset. The ML interatomic potential is thereafter re-trained and used to run a new MD, thereby repeating the process until the desired accuracy is reached. The right panel presents the improvement in the results of using the ML interatomic potential for successive re-training iterations. The earliest iterations yield unphysical heterogeneous distributions of Si and N; intermediate iterations show unphysical N-chains and $N - N$ bonds; the final potential overcomes all aforementioned issues and generates realistic and defect-free structures of Si_3N_4. 36

3.2 ML interatomic potential tested against DFT for forces (top) and energies (bottom) calculation of the atomic configurations in the testing dataset, together with their respective MAE. Atomic configurations were divided into liquid (blue) and amorphous (green). 38

3.3 Top: Total RDFs for structures generated with GAP (red), DFT (blue), experimental data (black), Tersoff (cyan), ReaxFF (yellow), Billeter without added terms (violet) and BHM (green). Bottom: Partial RDFs of Si-N (top), Si-Si (middle) and N-N (bottom) bonds for GAP (color) and DFT (shadow). For the total RDF (top), GAP presents the best agreement with the DFT reference and experimental data, closely followed by BMH. For the partial RDFs (bottom), GAP presents a good agreement with the DFT reference for the first and second peaks of the three relevant bonds. 40

3.4 Si-N bond-length distributions (left) and N-Si-N bond-angle distribution (right) for amorphous Si_3N_4 structures generated with GAP (red), DFT (blue) and BMH (green). 41

3.5 Comparison of the structure factor obtained from neutron scattering for the amorphous Si_3N_4 generated with GAP (red) and DFT (blue), together with experimental data found in the literature. 41

3.6 Coordination number statistics for Si (left) and N (right) in the atomic structures obtained with GAP (red) and DFT (blue). 42

3.7 Comparison of the VDOSs calculated with GAP and DFT. There is a good match over the entire range of relevant energy values. 42

3.8 Analysis of the computational cost of the proposed ML interatomic potential. Left: Running time vs. number of atoms in the system, for a simulation of 10,000 time-steps. Right: Performance vs. number of cores for a system size of 3,024 atoms. 43

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

3.9 Scaling of DFT computational cost versus system size (number of atoms) for a 10,000 time-steps MD at 300 K using 20 CPU cores. The fitting curve indicates a cubic scaling of the DFT computational cost with the system size. 43
3.10 Testing materials for exposition to neutron irradiation. (a) Empirical test: exposing the selected material to neutron irradiation and studying its effects on it. (b) Simulations: giving a high initial momentum to one of the atoms of the atomic structure of interest, running an MD simulation and studying the resulting changes in the atomic structure. 45
3.11 Creation of a dataset by running melt-and-quench MD simulations. The process is repeated for different concentrations of Ge. 46
3.12 Validation of the GAP against DFT for energy/forces calculations and structural properties of low Ge concentration structures. Color: GAP RDFs, grey: DFT RDFs for Si-Si and Si-Ge bonds. 46
3.13 Validating the GAP against DFT for energy/forces calculations and structural properties of high Ge concentration structures. Color: GAP RDFs, gray: DFT RDFs for Ge-Ge and Si-Ge bonds. 47
3.14 Examples of the results of running an MD simulation for neutron irradiation on SiGe for a low (top) and a high (bottom) energy incident particle. The stronger the color of an atom, the higher its velocity. The snaps are taken at 0 fs (left), 25 fs (center) and 50 fs (right) from the moment the incident neutron collides with the PKA. 48
3.15 Scaling of the system size in neutron irradiation MD simulations for SiGe allowed by GAP. 49
3.16 Schematic of the controlled dynamical sampling technique. The input of the plant is Δ, the number of time-steps of the MD simulation, while the output is ϵ, the errors between DFTB and GAP. An MD simulation is performed, starting from the current state of the atomic system of interest and for Δ time-steps, the DFTB and GAP energy/forces calculations for this snapshot are compared and their error, ϵ, is computed. By using a controller, the sampling frequency can be dynamically adapted depending on this error. Possible controllers are shown in the bottom-right panel, while a simplified schematic of this process is shown in the upper-right panel. 50
3.17 Oxidation process of a silicon slab exposed to O_2 in gas form. The MD simulation runs with GAP in LAMMPS through an iterative algorithm, in which the O_2 molecules are replaced periodically. 52
3.18 Zoom in to the oxidation process of a silicon slab exposed to O_2 gas. . 53
3.19 Results of simulating the oxidation process of a silicon nanowire, using the GAP developed to study the oxidation process of silicon surfaces exposed to O_2 gas. Credits: Franz Fehringer and Lukas Cvitkovich. . 53
3.20 (a) Schematic of a GAP. (b) Workflows for a traditional potential built with this ML model (left) and the proposed composite interatomic potential (right). (c) Schematic of the proposed composite potential, augmenting the main potential with an auxiliary potential. 54

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

3.21 Initial dataset for training and testing the *main* ML potential, created using the melt-and-quench technique and the ReaxFF force field. The MD begins with a crystalline SiO_2 structure, which is melted up to 5,000 K and thereafter quenched back to 300 K. Once the dataset is created, the energies are recalculated on a subset using DFT. Blue: Crystalline phase (C). Red: Liquid phase (L). Green: Amorphous phase (A). The objective of the initial dataset is to provide a sampling of the PES of interest as comprehensive as possible. 56

3.22 Testing the potential energy calculations accuracy for the traditional (red) and the composite (blue) ML potentials against DFT. The dashed black line represents the ideal results, in which all energy predictions would be equal to the DFT calculations. The closer the dots are to this line, the better the results are. This reference makes the X-axis immaterial. The results show a high accuracy for both potentials and virtually no difference between them. An example of the atomic structures used for testing is shown in the bottom-right sub-panel. 57

3.23 Panel comparing results of the proposed framework with those of a traditional ML interatomic potential, built as a single GAP. Both potentials were used to run the same MD simulation, as specified in (a). Two of the resulting structures obtained from the proposed approach are presented in (b). Resulting structures from using the traditional approach are presented in (c) and (d), together with examples for unphysical behavior found in them. The RDFs for one of the structures built with the new approach and one built with the traditional ML potential are presented in (e), together with the DFT reference. As can be seen, the composite potential performs much better when compared against the DFT reference. 59

3.24 Transferability tests for the proposed composite ML potential. (a) Computed VDOS in comparison to DFT results. The good match is remarkable, considering that the potential was not trained on forces. (b) Using the potential to create an a-SiO_2 structure 10 times the size of those used in the training dataset. No unphysical clustering is present in the model. 60

4.1 Workflow to predict a property from an atomic structure. The structure must be represented by a set of descriptors. Then, an ML model can be trained and subsequently used to predict the desired property in new atomic structures. 62

4.2 Example of two a-SiO_2 structures used in this work to train and test the ML models. Each structure contains 216 atoms and a total of 16 such structures are used in this work. 62

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

4.3 Hydrogen interaction with the a-SiO_2 matrix (left) can lead to the formation of a hydroxyl group, via the breakage of a strained Si-O bond, resulting in a hydroxyl E' center defect (right). The blue bubbles show the spin density associated with interstitial hydrogen and the hydroxyl E' center defect, respectively. 63

4.4 Formation energies of hydroxyl E' center defects in a-SiO_2, obtained from DFT calculations, together with their mean (μ) and standard deviation (σ), in eV. The broad distribution is due to the amorphous nature of the structures. 63

4.5 ACSF output vector: Initially, the output vector provides the values of the two-body symmetry functions G_1, G_2, and G_3 for each chemical species in the dataset. Subsequently, the vector includes the values of the three-body symmetry functions G_4 and G_5, listed for each unique combination of two chemical species. Reprinted from Computer Physics Communications 247, 106949 (2020). Original figure published under Creative Commons CC-BY 4.0 license. Credits: Lauri Himanen *et al.* 65

4.6 Correlation between formation energies and bond-lengths. The formation energy decreases for larger bond-lengths, so defects preferably form at strained Si-O bonds. The red line indicates the Si-O bond-length in alphaquartz. 66

4.7 Error distributions and MAE in the prediction of formation energy of hydroxyl E' center defects in a-SiO_2 for the different combinations of descriptors (SOAP, ACSF and BLBA) and ML models (neural network, KRR and DT). 68

4.8 Parity plots for the DFT calculation results and the ML-based predictions using SOAP (top), ACSF (middle) and BLBA (bottom) descriptors (in all cases, combined with a neural network). All descriptors perform similarly. However, SOAP shows a slightly better correlation in general, particularly for negative values. 69

4.9 Proposed workflow to predict the structure of defects in atomic 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 is used as an input for an 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). 71

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

4.10 (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 the 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 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. 72

4.11 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. 73

4.12 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. 73

5.1 Alternative computing architecture schematic, where the MD simulations run on a hybrid platform. The typical steps of an ML-based MD simulation are followed: The descriptors of the local atomic environment of the system are computed (I) and passed to the ML model (II) to compute the energy and forces (III) which are then used to update the atomic coordinates for the following time-step (IV). However, while the PES is evaluated within a neuromorphic accelerator (II-III), the remaining computation requirements are fulfilled by a standard computer (I-IV). 76

5.2 The four fundamental two-terminal circuit elements: resistor, capacitor, inductor and memristor. Reprinted from Nature 453, 80-83 (2008), with permission from the journal. Credits: Dmitri B. Strukov *et al.* . 78

5.3 SEM image of a novel germanium-based memristor fabricated at ETHZ (left), and zoom to the device (right). The device is composed of a vertical and a horizontal metallic electrode, and amorphous germanium in between them. In both figures, the triangle-like shapes are tapered to connect the small device to larger wires that then connect to electrical pads such that the devices can be contacted with probes, which can be measured. The dark lines that extend from the device area to the right are artifacts of the visualization technique. 78

5.4 SEM image of an array of 16 memristors fabricated at ETHZ. These arrays allow performing matrix multiplication in an analog manner. . 79

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

5.5 RV-curves of two novel germanium-based memristors fabricated at ETHZ. Note that for both devices the transition from the higher resistance to the lower resistance is gradual, meaning that the switching is not digital but analog. The device on the left presents a wider gap between the high- and low-resistant states, while the device on the right presents more cycles with less variance. These are highly desired and useful properties in memristors used in neuromorphic computing applications. 80
5.6 Zoom to the SEM image of a novel germanium-based memristor fabricated at ETHZ, together with the representative simulation cell used to perform MD simulations of the device. This is a cubic cell with a lattice parameter of 13.67 nm, which contains 110,592 atoms. The physical device is approximately 10 times larger than this cell. 80
5.7 Comparison of the RDF for the amorphous germanium generated with NVNMD (red) and DFT (black). 82
5.8 Comparison of the structure factor obtained from neutron scattering for the amorphous germanium generated with NVNMD (red) and DFT (black). 82
5.9 Comparison of the bond-length distributions in the amorphous germanium structures generated with NVNMD (red) and DFT (black). . . 83
5.10 Results of running a melt-and-quench MD simulation on a cell containing 110,592 atoms (left), and RDF validation against a DFT reference (right). Note that the main peak of the NVNMD structure is shorter than that of the DFT reference, this is because the latter is markedly smaller, with only 216 atoms. 84

6.1 Two different MEP waveforms (red and blue curves) with virtually identical latencies (black vertical line), defined as the time elapsed between the TMS pulse and the beginning of the MEP trace. The epoch starts at the time the TMS pulse is delivered. 86
6.2 Workflow for automated assessment of MEP latencies in a possible closed-loop TMS set-up. MEPs are measured with electrodes placed on the target muscle and stored in a 120-dimensional vector. The preprocessing is done by trimming, smoothing, centering, and normalizing the MEP. The resulting vector is used as an input to the neural network for the latency estimation. The dashed arrows show how the DELMEP algorithm could be applied to a closed-loop protocol (dashed box), in which the brain stimulation parameters are modified depending on the MEP responses. 88
6.3 (a) Pre-processing of the two MEPs shown in Figure 6.1, divided into: trimming (1), smoothing (2), centering (3) and normalizing (4). Panels (b) and (c) show the MEPs before and after the pre-processing, respectively. When comparing (b) and (c), note that the different MEPs look similar after the pre-processing, thus facilitating the training and later use of the neural network. 90

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

6.4 Automated MEP latency annotations with the proposed DELMEP algorithm in the training dataset (green) and testing dataset (orange). The results are compared to the manually assessed values. The MAE is presented for both datasets. 91
6.5 Illustrative MEPs from the testing dataset and their corresponding automated (dashed violet vertical line) and manually assessed (purple vertical line) latencies. These MEPs were not used to train the neural network. The similarity between both latencies indicates a good performance of the DELMEP algorithm. 92
6.6 Map of the DELMEP algorithm errors in estimating MEP latencies, as a function of the MEP V_{PP} and manually annotated latency. The upper-right panel shows the DELMEP estimation error versus the MEP V_{PP}, for MEPs with latency estimation error higher than 1 ms. 93
6.7 Relation between the average MAE obtained during the intra-subject CV tests and the dataset size of every subject (number of MEPs). Each data point represents a different subject in the dataset. 96
6.8 Latencies annotated with DELMEP and Bignoni's method vs. manually annotated latencies for the MEPs in Bigoni's dataset. The manually annotated latencies are the mean value of the annotations of three independent experts. 97

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

List of Tables

3.1 Parameters used to implement the Gaussian approximation potential (GAP) and its smooth overlap of atomic positions (SOAP), three-body and two-body descriptors. 34
3.2 Atomic structures added to the training dataset in every iteration. . . 35
3.3 Energy, force and stress convergence parameters for every structure type used in the training process 37
3.4 Validation of the GAP, by comparing relevant geometrical parameters of the resulting silicon oxidized structures to equivalent Density Functional based Tight Binding (DFTB) simulations. 52

5.1 Most relevant parameters used to train the quantized neural network (QNN), which performed the molecular dynamics (MD) simulations on the neuromorphic-based hardware accelerator. 81
5.2 Comparison of the performance of non-von Neumann molecular dynamics (NVNMD) to a traditional GAP interatomic potential, by using both methods to run the same MD simulation in LAMMPS. 83

6.1 Comparison of mean absolute error (MAE) obtained with DELMEP, Signal Hunter, absolute hard threshold estimation (AHTE), Squared Hard Threshold Estimator (SHTE), and Bigoni's method for the motor evoked potentials (MEPs) in the testing dataset. The MAE is reported for the entire dataset and for low (peak-to-peak amplitude (V_{PP}) < 100 µV) and high ($V_{PP} > 100$ µV) amplitude MEPs. MEPs discarded by Bigoni's method were also discarded on the other methods for a direct comparison. 94
6.2 CV of DELMEP across the entire dataset. The available data was divided into five different batches. In each test, one of the batches was used to test the algorithm and the remaining batches to train it. The MAE is reported for every batch (CV 1-5), as well as its average and standard deviation (SD) for all batches. 94
6.3 Intra-subject variability of DELMEP. A 5-fold CV was performed using data from one subject at a time, the process was repeated for all subjects. The table contains the MAE for every data batch of each subject, together with the average and SD for every subject. 95

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

6.4 Inter-subject variability of DELMEP. The data from one subject was used for testing and the data from the remaining subjects for training. This process was repeated for every subject. The MAE for every subject, together with its average and SD are reported. 95

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

Chapter 1

Introduction

Artificial intelligence (AI) is a relatively novel field. It is widely considered that it was born as a scientific discipline in 1956, during the Dartmouth Summer Research Project on Artificial Intelligence [1]. Therefore, compared to long-standing scientific disciplines like physics or chemistry, the development of AI has occurred within a minuscule timeframe, representing a fraction of the extensive history these fields have witnessed. Nonetheless, progress and development of AI have unfolded at an exponential pace. In 1996, the IBM supercomputer Deep Blue beat the world chess champion, Garry Kasparov, in a six-game match held in Philadelphia [2]. This milestone occurred only 40 years after the conceptualization of AI as a scientific field.

In the modern world, there is virtually no scientific discipline which has not been transformed by AI in general, and by the subfield of machine learning (ML) in particular. ML finds applications in a wide range of different domains, such as biology [3], physics [4], sociology [5] and neuroscience [6], where it focuses on the development of algorithms and models that allow information processing machines to learn and improve their performance without being explicitly programmed. The core concept of ML is to enable information processing machines to discover relationships and make accurate predictions or decisions based on the input data. The novelty of ML algorithms is that they learn from examples or experiences, adjusting their internal parameters to optimize their performance based on a provided dataset [7]. This is in contrast to traditional software paradigms, in which programmers would write explicit instructions for information processing machines to follow, defining every step and rule required to solve a problem.

Originally, ML was predominantly applied to simple tasks, with limited capabilities and narrow domains of application. Early ML algorithms were primarily focused on tasks such as pattern recognition [8] and classification [9], where the input data and desired outputs were relatively straightforward. These algorithms were capable of learning simple rules and making predictions based on the provided training data. However, as the field of ML advanced, researchers and practitioners began exploring its potential for more complex tasks.

At present time, ML algorithms can perform tasks as complex as processing natural language [10] and computer vision [11], with practical applications in medicine [12], finances [13], environmental sciences [14], etc. Among the scientific disciplines transformed by novel ML algorithms, we find the field of molecular dynamics (MD). MD is an umbrella term that encompasses a wide range of atomistic calculation techniques. What all of them have in common is that they are used to simulate the time evolution of the configuration of atomic systems, by considering their interatomic interactions. Even though MD is a well-established technique, researchers employing it face a long-standing dilemma: accuracy vs. efficiency [15][16]. On the one hand, *ab initio* methods, such as the density functional theory (DFT) evaluation of the potential energy surface (PES), offer high accuracy and require minimal information about the system at hand, other than its elemental composition, but at the expense of high computational costs which limit their use to only small atomic systems. On the other hand, empirical potentials, which are hand-crafted approximations to the PES, are computationally more efficient, but often not accurate enough for several applications, and they require expert knowledge to properly parametrize the potential for new materials.

In the last decade, ML interatomic potentials emerged as promising alternatives to *ab initio* methods and empirical potentials, aiming at combining the high accuracy of the former with the low computational cost of the latter. ML interatomic potentials work by training an ML model, commonly, but not always, a neural network, to act as a computationally affordable surrogate model of the PES [17][18].

The development of ML interatomic potentials has been nothing short of revolutionary in the field of MD. The following example should be enough to illustrate this. In 2018, Deringer *et al.* developed an ML interatomic potential for amorphous silicon [19]. The atomic structures generated with this potential contained as little as 2% defects, while reducing the computational times and costs by orders of magnitude when compared to *ab initio* methods. In fact, running the same melt-and-quench MD simulation to model amorphous silicon structures, required 16,000,000 core-hours with DFT, while only 40,000 with their ML interatomic potential. When considering the core-hour costs of the UK national supercomputer at the time, this translated in a cost reduction from USD 185,000 to USD 500.

The particularly good trade-off between accuracy and efficiency in ML algorithms aids researchers and engineers in performing simulations and data processing tasks which were previously out of range by orders of magnitude. This is particularly valuable in tasks in which massive amount of data must be processed, such as when simulating extensive atomic systems [20]; or when timing is crucial, as in the case of systems required to process data on-the-fly [21].

In general terms, commonly used ML models are agnostic to the specific application they were trained for. The same models, techniques and training algorithms can be used for seemingly unrelated applications in highly dissimilar fields, by providing the appropriate training data and making the necessary adjustments. This means that ML solutions developed for one specific application can potentially be of use to solve problems in a different scientific field. This concept is at the core of this thesis, since the solutions developed for interatomic potentials used in MD simulations were adapted to process biomedical data of neurological patients. The ML algorithms developed in this thesis serve to reduce computational times by orders of magnitude in the first case, and to accelerate the work of neuroscientist and medical doctors in the

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Wien Bibliothek Your knowledge hub

second. In particular, in this work the challenge of automating the analysis of motor evoked potentials (MEPs) generated by transcranial magnetic stimulation (TMS) was addressed. These are biomedical signals measured with electrodes on the target muscle of a patient. They are typically characterized by the time elapsed between the stimulation and the beginning of the muscle movement. This characterization provides medical doctors with information on the corticospinal excitability of the patient, which is needed, among a wide range of applications, to evaluate the progress of patients suffering neuromotor diseases, such as amyotrophic lateral sclerosis [22].

A preliminary assessment might create the perception that such ML algorithms must be unrelated and highly disparate. However, considering: (I) ML models as mathematical functions, (II) their training as an optimization problem, and (III) the data used to train them as abstract mathematical objects, the relationship between these seemingly unrelated problems becomes less blurry and more intuitive.

1.1 Research Goals

In general terms, the research goals of this thesis involve the development of ML tools and techniques aimed at reducing the computational times and costs required for running MD simulations with *ab initio* accuracy. Additionally, this thesis addresses the transdisciplinary application of these techniques in processing biomedical data.

In concrete terms, the research presented in this thesis has been conducted at the Institute for Microelectronics, TU Wien [23], in the context of the Modeling Unconventional Nanoscaled Device FABrication (MUNDFAB) research project [24]. It consists of the following research goals: (I) an ML interatomic potential for silicon-nitride, with special focus on modelling amorphous structures, and an active learning technique, in cooperation with Argonne National Laboratory in Lemont, USA [25]; (II) an ML interatomic potential for silicon-germanium, with application to studying the effects of neutron irradiation in this material in radiative environments, such as nuclear power plants and airspace, developed during my research stay at LAAS-CNRS in Toulouse, France [26]; (III) ML interatomic potentials to study the oxidation process of silicon (together with an active learning technique) and to model amorphous silicon-dioxide, with application to simulating this process which is crucial for the semiconductor industry; (IV) ML models to predict the formation energy and structure of hydrogen-based defects in amorphous silicon-dioxide; (V) an ML interatomic potential for germanium which runs on a neuromorphic hardware accelerated platform, with application to studying germanium-based memristors and the possibility of simulating systems composed of more than 100,000 atoms, developed during my research stay in the Computational Nanoelectronics Group at ETH Zurich in Switzerland [27]; and (VI) an ML-based technique to automate the analysis of MEPs signals generated by TMS, in cooperation with the Department of Neuroscience and Biomedical Engineering at Aalto University in Helsinki, Finland [28].

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

1.2 Thesis Outline

The structure of the thesis aims at first introducing the theoretical framework in Chapter 2, then presenting and discussing the results in the following chapters. Chapter 3 presents the results of developing ML interatomic potentials for silicon-nitride, silicon-germanium, the oxidation of silicon, and silicon-dioxide. These results are grouped into a single chapter as they are all based on using the Gaussian approximation potential (GAP) method. Chapter 4 discusses ML-based alternatives to interatomic potentials and presents the results of such developments to study the formation energy and structure of hydrogen-based defects in amorphous silicon-dioxide. In Chapter 5, the results of using a neuromorphic hardware accelerator for MD simulations to model amorphous germanium are presented. In Chapter 6, the transdisciplinary application of the previously considered ML developments to biomedical data processing is discussed. Finally, Chapter 7 presents the conclusions of the thesis.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Chapter 2

State-of-the-Art and Methods

The following chapter provides an overview of the current state-of-the-art in MD simulations in general, and ML-based interatomic potentials in particular. The main aim is to offer the essential theoretical groundwork for comprehending the results presented in the subsequent chapters. The Born-Oppenheimer approximation and the concept of PES are first discussed in Section 2.1, followed by an overview of MD simulation methods in Section 2.2. A discussion of relevant classical ineratomic potentials can be found in Section 2.3, while an introduction to ML interatomic potentials and chemical environment descriptors is presented in Section 2.4.

2.1 The Born-Oppenheimer Approximation

The Born-Oppenheimer approximation was first proposed by M. Born and R. Oppenheimer in 1927 [29]. The approximation assumes that the motion of the electrons is much faster than that of the nuclei, allowing the electronic motion to be treated as instantaneous with respect to the motion of the nuclei. This is because each proton or neutron in the nuclei has a mass several orders of magnitude larger than that of an electron [30].

This approximation allows the separation of the motion of atomic nuclei and electrons in a molecule or solid. In mathematical terms, this allows to approximate the total wavefunction of the atomic system as the product of an electronic wavefunction and a nuclear wavefunction:

$$\Psi_{\text{total}} = \psi_{\text{electrons}} \times \psi_{\text{nuclei}}, \tag{2.1}$$

this enables the separation of the Hamiltonian operator into electronic and nuclear terms, since electrons and nuclei are considered as decoupled systems under this approximation. Disentangling the motion of the electrons from that of the nuclei greatly reduces the mathematical complexity of the equations of motion, thereby also reducing the computational efforts needed to solve them.

In the Born-Oppenheimer approximation, the Schrödinger equation for the electrons is solved with the nuclear coordinates fixed. This implies considering the coor-

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

dinates of the electrons as functions of the coordinates of the nuclei. Therefore, given a system with N atoms, where the position of atom i is the 3-dimensional vector R_i, the potential energy of the system, E, can be written as:

$$E = E(R_1, R_2, \ldots, R_N). \tag{2.2}$$

This function is known as the PES, which relates the potential energy of the system, E, a scalar, with the positions of the atoms in it, a 3-N scalar field. Therefore, for every configuration of the system (i.e., for every set of atomic positions), the PES relates it to the potential energy of the system, E. An illustration of the PES for a hypothetical reaction system is shown in Figure 2.1. The x and y axis represent coordinates, while the z axis represents the potential energy of the system. The highest potential energy is indicated with red and the lowest with violet. The color scheme of the PES is also used in the 2-D contour map. In this example, 1 and 2 are saddle points. The reaction takes the system from the energy minimum to a local minima, passing through the saddle point number 2.

In essence, the objective of ML interatomic potentials is to act as computationally more efficient surrogate models of the PES of interest [17]. By *learning* from a relevant dataset of points on the PES of interest, ML interatomic potentials are thereafter able to interpolate and compute the potential energy and forces (*e.i.*, the height and curvature of the PES) for new atomic configurations.

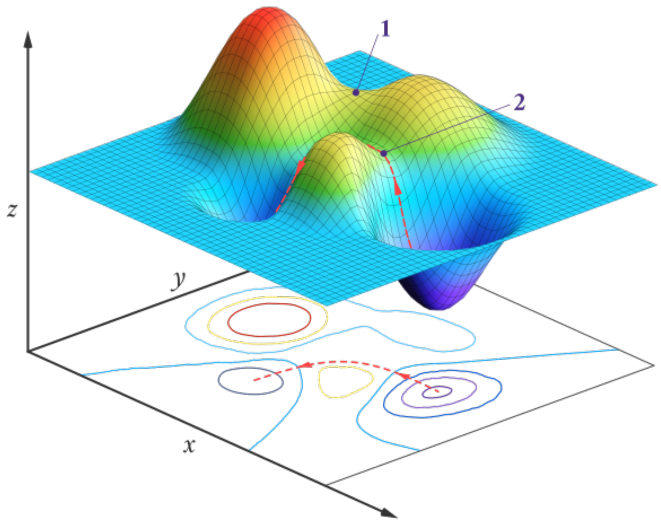

Figure 2.1: PES for a hypothetical reaction system with two coordinates. (CC BY-NC; Ümit Kaya via LibreTexts)

2.2 Molecular Dynamics Simulations

The technique of MD allows to simulate the behavior and evolution of atomic systems over time. MD is an essential atomic-scale research method, finding use in a wide range of fields, such as physics [31], biology [32] and nanotechnology [33]. In MD, the vast majority of simulation techniques assume the Born-Oppenheimer approximation, therefore treating the atomic nuclei independently of the electrons. This means that, in the classical limit, the nuclei obey Newton's equations of motion:

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = m_i \frac{\mathrm{d}^2 q_i}{\mathrm{d}t^2} = -\frac{\delta U(q)}{\delta q_i}, \tag{2.3}$$

where q are the system coordinates, q_i the coordinates of the atom i, U the total potential energy of the system, t the time, p_i the momentum of atom i, and m_i is the mass of atom i.

Since the Born-Oppenheimer approximation is assumed, the nuclei evolve on the PES. Therefore, the forces that govern their movements and interactions are computed as the negative of the derivative with respect to the atomic coordinates of the total potential energy of the system. Computing the exact PES of a complex atomic system is most frequently not possible in practical terms, and different levels of approximation are used. The theory level of approximation used in a specific MD simulation is referred to as the *potential*. The potential is the physical model of the PES. It can range from highly accurate *ab initio* methods, such as DFT, to simpler empirical equations in the form of classical potentials. Classical potentials will be discussed in Section 2.3, while modern ML-based potentials will be presented in Section 2.4.

Once the interatomic interactions are defined by the potential employed, the equations are discretized to enable a computational solution. This is done by defining a time-step, which is an empirically set parameter. The time-step must be short enough to allow access to the relevant physical mechanisms of interest. However, the shorter the time-step, the higher the total number of steps. Therefore, a reasonable trade-off is needed. A typical time-step, both for *ab initio* and classical methods, is 1 fs. The discretized equations are typically solved using a symplectic (i.e., approximately energy-conserving) integrator. An example of one such algorithm is the Verlet method [34],

$$\begin{aligned}
r_i(t+\delta t) &= r_i(t) + v_i(t)\delta t + \frac{f_i(t)}{2m_i}\delta t^2 \\
v_i(t+\delta t/2) &= v_i(t) + \frac{\delta t}{2}\frac{f_i(t)}{m_i} \\
f_i(t+\delta t) &= f_i(r_i(t+\delta t)) \\
v_i(t+\delta t) &= v_i(t+\delta t/2) + \frac{\delta t}{2}\frac{f_i(t+\delta t)}{m_i},
\end{aligned} \tag{2.4}$$

where t is the time, r_i is the coordinates vector, v_i is the velocity vector and i is the atom index. Given a potential, a reasonable time-step and a method to solve the

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

discretized equations, Newton's equations of movement can be solved for the isolated atomic system of interest. Often, however, it is desirable to study the interaction between the atomic system of interest and its environment, in the form of energy exchanges and changes in its volume. This can be done by modifying the equations of motion to simulate constant temperature or constant pressure conditions, by means of a thermostat or barostat, respectively. These elements will be explained next.

2.2.1 Thermostats

Thermostats are used to regulate the atomic system's temperature. They are modifications to the equations of motion that emulate the energy exchange between the atomic system and a theoretical heat reservoir, such as the environment. Given an atomic system, the instantaneous value of its temperature is related to the kinetic energy, via the equipartition theorem [35],

$$\begin{aligned} \langle K \rangle &= \frac{3}{2} N k_B T \\ K &= \frac{1}{2} \sum_i m_i v_i^2, \end{aligned} \tag{2.5}$$

where K is the kinetic energy, m_i is the mass of atom i, v_i is the velocity of atom i, k_B is the Boltzmann constant, T is the temperature of the system, and N is the number of atoms in the system. The operation of a thermostat on an atomic system can be viewed as a closed-loop control problem, where, using control theory jargon, the atomic system is the plant and the thermostat is the controller. In every step, the temperature of the atomic system is measured and compared to the target temperature. Based on this difference, the thermostat will modify the equations of motion to reach the target temperature. This modification must be done gradually, in order to avoid unphysical behaviors. The rate at which the modification is done is determined by either a single or multiple parameters, depending on the thermostat being used.

There is a wide range of thermostats used in MD simulations. An obvious and simple way to implement a thermostat is through *velocity scaling*. In this implementation, the velocities of all particles are scaled to reach the target temperature,

$$\begin{aligned} v_i^{\text{new}} &= v_i^{\text{old}} \lambda \\ \lambda &= \sqrt{T_0 / T(t)}, \end{aligned} \tag{2.6}$$

where T_0 is the target temperature, $T(t)$ is the current temperature, λ is the scaling factor, and $v_i^{\text{old}}/v_i^{\text{new}}$ is the previous/new velocity of atom i. However, one problem with this simple approach is that it does not allow for temperature fluctuations, which are present in the canonical ensemble, therefore making it a non-ideal candidate to model real-world energy exchanges.

An improvement from the velocity scaling method is the Berendsen thermostat [36], which works by simulating a weaker coupling between the atomic system of interest and a thermal bath,

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

$$\frac{\mathrm{d}T(t)}{\mathrm{d}t} = \frac{1}{\tau}(T_0 - T(t)), \tag{2.7}$$

where τ is the coupling parameter, which determines the strength of the heat exchange between the atomic system and the thermal bath. τ is given in units of time and plays an important role in the dynamics of the atomic system. The smaller the value of τ, the faster the dynamics will be, thereby taking less time to reach the target temperature, but producing larger temperature changes. On the contrary, larger values of τ result in a longer time to reach the target temperature, but smoother changes in the temperature values of the atomic system over time.

An alternative method, and probably the most common one, is the Nosé-Hoover thermostat, which adds an extra degree of freedom to the Hamiltonian of the atomic system,

$$H(P, R, p_s, s) = \sum_i \frac{p_i^2}{2ms^2} + \frac{1}{2}\sum_{ij,i\neq j} U(r_i - r_j) + \frac{p_s^2}{2Q} + gk_B T ln(s), \tag{2.8}$$

where g is the number of independent momentum degrees of freedom of the system, R and P represent all coordinates r_i and p_i, s is the additional degree of freedom, and Q represents a mass which should be chosen along with the system. The extra degree of freedom emulates the coupling between the atomic system and a heat bath, with which it can exchange energy [37][38]. In the Nosé-Hoover thermostat, the entire atomic system is coupled to a heat bath.

Another approach is the Langevin thermostat [39], in which each atom is coupled to a heat bath, emulated by a fluctuating force and a friction term,

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\delta U(q)}{\delta q_i} - \gamma_i m_i v_i + \eta_i \xi_i, \tag{2.9}$$

where γ_i is a positive real number, representing a friction term that dissipates energy from the atomic system; ξ_i is a noise term that imposes a stochastic force on atom i and η_i is a positive real number that determines the magnitude of this stochastic force on atom i. It is important to note that the Langevin thermostat is truly stochastic in comparison with the other methods, which are deterministic after initialization. Regardless of the selected method and its parametrization, the function of the thermostat is to regulate the atomic system temperature, in order to simulate the energy exchange with the environment.

2.2.2 Barostats

A barostat is analoguous to a thermostat but controls the pressure of the atomic system being simulated. In the same way in which the system temperature is related to the velocity of the particles in the atomic system, the pressure is related to the volume of the atomic system. As with thermostats, the time evolution of the free variable, i.e., the system volume, can be controlled by a parameter or set of parameters, depending on the selected barostat.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

In an analog way to the velocity scaling, the coordinates of the atoms in the system can be scaled to reach the target pressure,

$$\begin{aligned} V^{\mathrm{new}} &= V^{\mathrm{old}}\lambda \\ r_i^{\mathrm{new}} &= r_i^{\mathrm{old}}\lambda^{1/3}, \end{aligned} \tag{2.10}$$

where V^{new} is the volume of the atomic system in the new step, V^{old} is the volume of the atomic system in the previous step, $r_i^{\mathrm{new}}/r_i^{\mathrm{old}}$ are the atomic coordinates of atom i in the new/previous step, and λ is the scaling factor. As with the thermostat, the problem with this simple approach is that it does not allow the system pressure to oscillate, therefore making it not ideal to model real systems.

Analog to the Berendsen thermostat, the Berendsen barostat is an improvement over the direct scaling barostat, in which a weaker coupling with the pressure reservoir is implemented,

$$\frac{\mathrm{d}P(t)}{\mathrm{d}t} = \frac{1}{\tau}(P_0 - P(t)), \tag{2.11}$$

where $P(t)$ is the atomic system pressure over time, P_0 is the target pressure and τ is the coupling parameter. As with the thermostat, τ determines the dynamics of the pressure control.

In this same manner, a Nosé-Hoover barostat can be implemented, as an analog to the Nosé-Hoover thermostat. Moreover, a pure barostat can be used on an ensemble, or it can be combined with a thermostat to generate the isothermal-isobaric statistical ensemble.

2.3 Classical Potentials

When running MD simulations, the total energy of the atomic system and the forces acting on each of the atoms which compose the system must be computed for every time-step. Performing these calculations with an *ab initio* method is computationally challenging, even for simple or trivial atomic system, therefore severely limiting the system sizes that can be studied, and shortening the simulation times that can be accessed. The computational limits of *ab initio* methods are rapidly reached, since their computational costs typically scale cubically with the number of atoms in the system.

It is in this context that classical potentials, also known as *empirical potentials*, are developed. Classical potentials are computationally efficient approximations to the Born-Oppenheimer PES. Their analytical form is typically derived from fundamental physical arguments, while their parametrization is based on experimental data or more expensive *ab initio* calculations. The different physical interactions are typically tackled independently, in the form of separate terms. Moreover, intermolecular and intramolecular interactions are commonly modeled separately.

In the following sections, the Lennard-Jones [40], Tersoff [41] and ReaxFF [42] classical potentials will be discussed. Lennard-Jones is one of the earliest examples of a classical potential and its simplicity allows for it to serve as an illustrative example. Tersoff and ReaxFF are more complex and accurate classical potentials, which were used in several of the production MD simulations in this doctoral thesis.

2.3.1 Lennard-Jones

The Lennard-Jones classical potential was developed by John Lennard-Jones in an attempt to produce a physical model to explain the temperature-dependent variation of the viscosity in gases [40]. The analytical form of the Lennard-Jones classical potential considers only pairwise interactions of atoms, and it is composed of an attraction term and a repulsion term, both of which are functions of an inverse power of the interatomic distance,

$$U(r) = Ar^{-n} - Br^{m}, \tag{2.12}$$

where U is the potential energy, r is the interatomic distance, A and B define the magnitude of the attractive and repulsive terms, respectively, and n and m define the rate of change of these terms with respect to the interatomic distance. Therefore, the model is based on four parameters, all of which are positive numbers, that can be empirically set, based on experimental data.

In its most commonly used form, the Lennard-Jones potential adopts m and n to be equal to 12 and 6, respectively. Because of this reason, this classical potential typically receives the name *12-6 potential*. Since m and n are set, there are only two parameters left to be determined by the user, A and B. These parameters are typically re-arranged, to allow a simpler physical interpretation,

$$U(r) = 4\epsilon(-\frac{r}{\sigma}^{-6} + \frac{r}{\sigma}^{-12}), \tag{2.13}$$

where ϵ is the potential energy well depth, r is the interatomic distance and σ is the length scale. Many force-fields take this form as their basis, such as AMBER [43], OPLS [44] and AIREBO [45]. The simple form of these classical potentials makes them highly computationally efficient. However, it also limits their accuracy, as they can only account for 2-body interactions. When more accurate calculations are needed, it is necessary to account for multi-body interatomic interactions.

2.3.2 Tersoff

The previously discussed Lennard-Jones classical potential considers only two-atom interactions. This makes it computationally efficient but not accurate enough for several applications, as more complex interatomic interactions are missing. In a trade-off between computational efficiency and accuracy, the Tersoff potential considers not only two-atom interactions, but three-atoms interactions as well. This allows it to account for the role that bond-angles play in defining the potential energy of the atomic system. The Tersoff classical potential offers great flexibility and at a relatively low computational cost. It has the following analytical form [41],

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

$$E = \sum_{i=1} E_i = \frac{1}{2}\sum_{i\neq j} V_{ij}, \tag{2.14}$$

$$\begin{aligned}
V_{ij} &= f_c(r_{ij})[f_R(r_{ij}) + b_{ij} f_A(r_{ij})] \\
f_R(r_{ij}) &= A_{ij} e^{(-\lambda_{ij} r_{ij})} \\
f_A(r_{ij}) &= B_{ij} e^{(-\mu_{ij} r_{ij})} \\
f_c(r_{ij}) &= 1, \text{ if } r_{ij} < R \\
f_c(r_{ij}) &= \frac{1}{2} + \frac{1}{2}\cos(\pi\frac{r_{ij} - R_{ij}}{S_{ij} - R_{ij}}), \text{ if } R_{ij} < r_{ij} < S_{ij} \\
f_c(r_{ij}) &= 0, \text{ if } r_{ij} > R, \\
b_{ij} &= \chi_{ij}(1 + \beta_i^{\eta_i}\zeta_{ij}^{\eta_i})^{-1/2\eta_i} \\
\zeta_{ij} &= \sum_{k\neq i,j} f_c(r_{ik})\omega_{ik} g(\theta_{ijk}) \qquad (2.15) \\
g(\theta_{ijk}) &= 1 + c_i^2/d_i^2 - c_i^2/[d_i^2 + (h_i - \cos(\phi_{ijk}))^2] \\
\lambda_{ij} &= (\lambda_i + \lambda_j)/2 \\
\mu_{ij} &= (\mu_i + \mu_j)/2 \\
A_{ij} &= (A_i + A_j)^{1/2} \\
B_{ij} &= (B_i + B_j)^{1/2} \\
R_{ij} &= (R_i R_j^{1/2}) \\
S_{ij} &= (S_i S_j^{1/2}),
\end{aligned}$$

where i, j, and k are the labels of the atoms, r_{ij} is the bond-length between the atoms i and j, and θ_{ijk} is the bond-angle between the $i-j$ and $i-k$ bonds. The parameter χ_{ij} regulates the strength of the heteropolar bonds. λ_{ij}, μ_{ij}, A_{ij}, B_{ij}, R_{ij}, and S_{ij} are empirically defined parameters that depend on the atomic species being considered. The Tersoff potential has been succesfully parametrized for a wide variety of materials, such as silicon [46], silicon-dioxide [41], and carbon [47].

2.3.3 ReaxFF

The previously described Tersoff potential provides a higher degree of accuracy than the simpler Lennard-Jones potential. However, it is still insufficient for several applications. ReaxFF is a more complex potential, able to produce more accurate results, at the expense of a lower computational efficiency when compared to the simpler Tersoff and Lennard-Jones potentials. In the ReaxFF potential, the potential energy of the atomic system is divided in the following components,

$$E_{\text{system}} = E_{\text{bond}} + E_{\text{over}} + E_{\text{angle}} + E_{\text{tors}} + E_{\text{vdWaals}} + E_{\text{Coulomb}} + E_{\text{specific}}. \tag{2.16}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

A detailed explanation of each of these components can be found in [42]. In essence, E_{bond} is an empirical function of the bond formation energy. E_{over} is a penalty preventing over-coordination. The many-body effects are given in E_{angle} (three-body valence angle strain) and E_{tors} (four-body torsional angle strain). E_{vdWaals} represents the van der Waals interaction and E_{Coulomb} the non-bonding electrostatic energy. Finally, E_{specific} are system-specific inclusions.

In ReaxFF, the energy components can be either bond-order-dependent or bond-order-independent. The bond-order is defined as the number of chemical bonds between a pair of atoms and indicates the stability of a bond, and it is calculated from the interatomic distance,

$$\begin{aligned} BO_{ij} &= BO_{ij}^{\sigma} + BO_{ij}^{\pi} + BO_{ij}^{\pi\pi} \\ BO_{ij} &= e^{[p_{bo1}(\frac{r_{ij}}{r_0^{\sigma}})^{p_{bo2}}]} + e^{[p_{bo3}(\frac{r_{ij}}{r_0^{\pi}})^{p_{bo4}}]} + e^{[p_{bo5}(\frac{r_{ij}}{r_0^{\pi\pi}})^{p_{bo6}}]}, \end{aligned} \tag{2.17}$$

where BO_{ij} is the bond order between atoms i and j, r_{ij} is the interatomic distance, the r_0 terms are equilibrium bond lengths, and p_{bo} terms are empirical parameters, determined by experimental data for the material of interest. The main improvement of ReaxFF over Tersoff is that it explicitly acounts for the formation and breaking of bonds, which makes it well-suited for studying reactive materials. The ReaxFF potential has been succesfully used to study a wide variety of materials, systems and processes, such as hydrocarbon oxidation [48], hydrogen combustion [49] and modeling hafnia [50].

2.4 Machine Learning Interatomic Potentials

In the context of studying semiconductor devices and technologies, the main motivation to develop ML interatomic potentials is to overcome the limitations posed by the high computational costs of *ab initio* methods, such as DFT. These methods are typically limited to only relatively small systems ($< 10^3$ atoms) and short simulation times ($\sim 10^1$ ps). The final objective of ML interatomic potentials is to serve as a computationally efficient substitute to *ab initio* methods in the calculation of atomic configuration energies and forces, needed to run MD simulations. In general terms, an ML interatomic potential takes the atomic coordinate space as the input and transforms them into an output space of the system energy. Ideally, these calculations will be both accurate and computationally efficient, allowing to simulate system sizes and process times out of range for other solutions.

The energy of a given atomic configuration is a property of the configuration itself, and it is thereby independent of the coordinate system employed to represent it. Therefore, traditional cartesian coordinate systems are ill-suited as inputs to ML models, since they are dependent on an origin which is arbitrarily selected [36]. As a consequence of this, a descriptor must be used, to act as a buffer between the atomic structure coordinates and the ML model used to assess its energy and forces. A descriptor computes a mathematical representation (i.e., a vector or matrix) of a given atomic structure and serves as an intermediate between the atomic structure and the ML model. However, not any possible mathematical representation is efficient as a descriptor, as they must be invariant under translation and rotation of the atomic structures, as well as under permutation of identical atoms [51][52]. Such properties aid in training an ML model to compute a property from atomic structures, since, if atomic coordinates were used instead of a descriptor, the ML model would have to learn that any possible translation, rotation or atomic permutation in an atomic structure corresponds to the same atomic property. Namely, the ML model would have to learn that an infinite number of different inputs correspond to the same single output, therefore rendering the training process unfeasible. Descriptors can be either global (they represent the entire atomic system) or local (they represent the local or near environment of a given atom in the system), with the latter being the most commonly used for ML interatomic potentials. The ML model is trained to find a functional relationship between a descriptor (highly-dimensional) and the corresponding potential energy (scalar-value) of the atomic structure which it represents. Therefore, except for some novel methods, such as message-passing networks [53], traditional ML based interatomic potentials are composed of two main components: (A) descriptor and (B) ML model. A schematic of this framework is presented in Figure 2.2. Ideally, for any given material and application of interest, a combination of descriptor(s) and ML model(s) could be found to develop an interatomic potential that fulfills the accuracy and computational cost requirements.

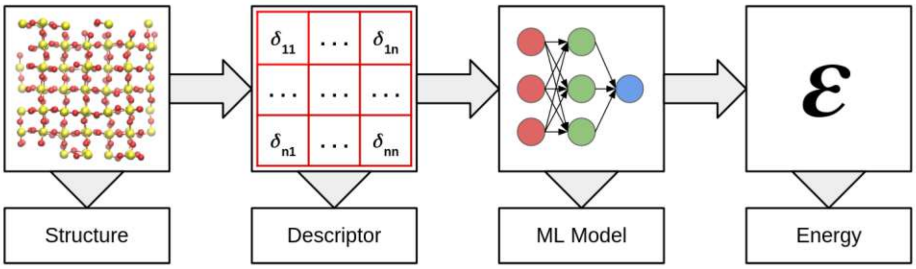

Figure 2.2: Workflow to compute the energy of an atomic configuration with an ML model. The input atomic configuration is first represented by means of a descriptor. This descriptor is subsequently used as input to the ML model, which produces the energy as its output.

2.4.1 Chemical Environment Descriptors

As previously mentioned, ML interatomic potentials generally need a descriptor to represent the atomic structures in a way which is compatible with ML algorithms. These descriptors can be as simple as a collection of geometrical values (e.g., bond-lengths, bond-angles, etc.) or take more complex mathematical forms that allow to represent the chemical environments with a higher degree of accuracy. This section will present the trivial geometrical descriptors, the more complex Behler-Parrinello [17] and the commonly used smooth overlap of atomic positions (SOAP) [54].

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

2.4.1.1 Geometrical Descriptors

Geometrical descriptors serve as fundamental tools for characterizing atomic structures within the context of ML interatomic potentials. These descriptors are essentially sets of meticulously ordered geometric parameters derived from the atomic structure under consideration, such as bond-lengths and bond-angles. The distinct advantage of employing geometrical descriptors lies in their computational efficiency; they are typically straightforward to calculate. Furthermore, these descriptors offer a high degree of interpretability, with each numerical value corresponding to a specific geometrical property of the atomic system.

Despite their merits, geometrical descriptors present certain limitations. Chief among them is their potential inadequacy in furnishing sufficient information for training highly accurate ML interatomic potentials. This shortfall arises due to the inherent simplicity of geometrical descriptors, which may not capture the intricacies of atomic interactions with the requisite depth. Additionally, when utilized as standalone descriptors on a global scale, they lack invariance concerning atom permutations. This lack of invariance is illustrated in Figure 2.3, where two distinct atomic structures can be generated using identical sets of six interatomic distances [54].

To enhance the effectiveness of geometrical descriptors in the context of ML interatomic potentials, complementary approaches that capture more nuanced aspects of interatomic interactions must be explored. This may involve incorporating additional descriptors that account for electronic structure, charge distribution, or other quantum mechanical properties. Striking a balance between the simplicity and interpretability of geometrical descriptors and the nuanced information provided by more complex descriptors is a key challenge in developing robust and accurate ML interatomic potentials for diverse atomic systems.

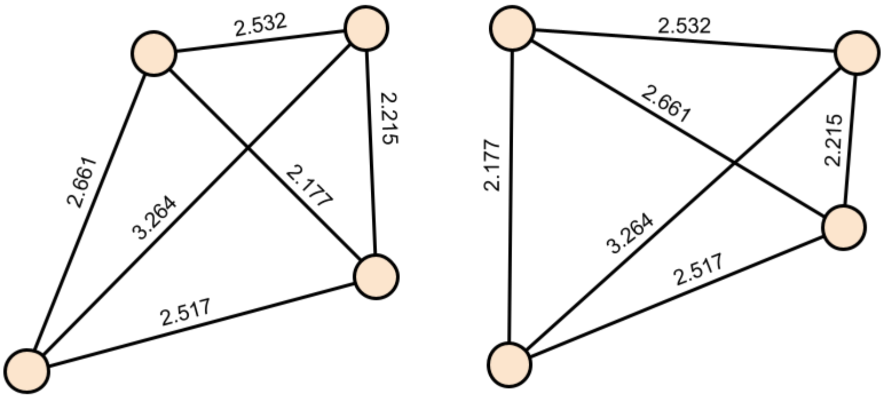

Figure 2.3: Two distinct atomic structures, constructed with the same set of six interatomic distances.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

2.4.1.2 Behler-Parrinello Descriptor

The Behler-Parrinello descriptor was initially introduced in 2007 in a foundational paper by Behler and Parrinello [17]. In this seminal work, the authors proposed a novel approach to fitting a neural network to DFT calculations of energies and forces for atomic systems of arbitrary size. The method was tested for bulk silicon and validated against experimental data. The Behler-Parrinello descriptor considers only interatomic interactions within a certain radial cutoff, defined as

$$f_{\mathrm{c}}(R_{ij}) = 0.5[\cos(\frac{\pi R_{ij}}{R_{\mathrm{c}}}) + 1], \tag{2.18}$$

for $R_{ij} \leq R_{\mathrm{c}}$, and zero for $R_{ij} > R_{\mathrm{c}}$. Radial symmetry functions are thereafter constructed as sums of Gaussians with the parameters η and R_{s},

$$G_i^1 = \sum_{j \neq i}^{all} e^{-\eta (R_{ij} - R_{\mathrm{s}})^2 f_{\mathrm{c}}(R_{ij})}. \tag{2.19}$$

On the other hand, angular terms are constructed for all triplets of atoms by summing the cosine values of the angles $\theta_{ijk} = \frac{R_{ij}.R_{ik}}{R_{ij}R_{ik}}$, centered at atom i, with $R_{ij} = R_i - R_j$,

$$G_i^2 = 2^{1-\zeta} \sum_{j,k \neq i}^{all} (1 + \lambda \cos(\theta_{ijk}))^{\zeta} e^{-\eta (R_{ij}^2 R_{ik}^2 R_{jk}^2)} f_{\mathrm{c}}(R_{ij}) f_{\mathrm{c}}(R_{ik}) f_{\mathrm{c}}(R_{jk}), \tag{2.20}$$

where the parameter λ can take the value 1 or -1 and different values of the parameters η, R_s and ζ can be used to generate an arbitrary number of invariants. This framework allows for the capture of intricate interatomic relationships, making the Behler-Parrinello descriptor a potent tool in the development of ML interatomic potentials.

2.4.1.3 Smooth Overlap of Atomic Positions (SOAP)

The SOAP descriptor was originally introduced in 2013 by Bartók *et al.* where it was tested by fitting models to the PES of small silicon cluster and the bulk crystal [54]. In the realm of ML interatomic potentials, the SOAP descriptor stands as a crucial tool for capturing the structural information of materials at the atomic level. SOAP is a representation scheme designed to encode the local environment around each atom in a material system. This descriptor plays a pivotal role in developing accurate and transferable interatomic potentials, facilitating the exploration of material properties and behaviors.

The SOAP descriptor achieves its effectiveness by quantifying the spatial distribution of atoms within a given cutoff distance, employing a local expansion of a Gaussian smeared atomic density with orthonormal functions based on spherical harmonics and radial basis functions. This descriptor starts from defining the neighborhood density of a given atom, a, as

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

$$\rho_a(\boldsymbol{r}) = \sum_b e^{-\frac{(\boldsymbol{r}-\boldsymbol{r}_{ab})^2}{2\sigma_{\text{at}}^2}} f_{\text{cut}}(r_{ab}), \tag{2.21}$$

where the sum is over all atoms within the f_{cut} cutoff, which goes smoothly to zero at r_{cut}. An example of a cutoff function can be seen in Figure 2.4. $\boldsymbol{r}$ is the coordinate vector of atom a, $\boldsymbol{r}_{ab}$ is the coordinate vector from atom a to atom b, and σ_{at} is the parameter that controls the smoothness of the potential. This local neighborhood density is thereafter expanded on a basis set of spherical harmonics Y_{lm} and orthogonal radial basis functions g_n,

$$\rho_a(\boldsymbol{r}) = \sum_{nlm} c^{(a)}_{nlm} g_n(\mathrm{a}) Y_{lm}(\hat{\boldsymbol{r}}), \tag{2.22}$$

where n is the index for the radial basis functions up to n_{max}, l is the angular degree of the spherical harmonics up to l_{max}, and m goes from $-l_{max}$ to l_{max}. These expansion coefficients are then used to form the spherical power spectrum,

$$P^{(a)}_{nn'l} = \sqrt{\frac{8\pi^2}{2l+1}} \sum_m (c^{(a)}_{nlm}) * c^{(a)}_{n'lm}. \tag{2.23}$$

A graphical representation of the process of building the SOAP descriptor for an atomic system can be seen in Figure 2.4, while the mathematical derivations are available in Ref. [54]. The SOAP descriptor emerged as a cornerstone in the development

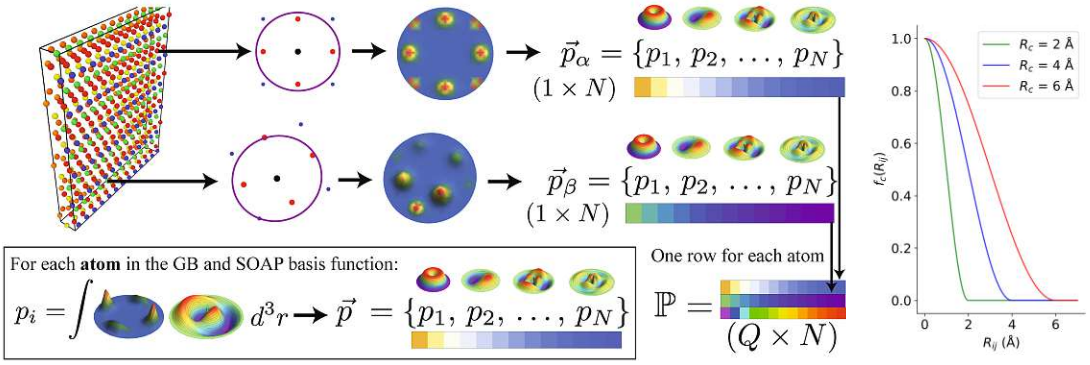
$\vec{p}_\alpha = \{p_1, p_2, \ldots, p_N\}$
$(1 \times N)$
$\vec{p}_\beta = \{p_1, p_2, \ldots, p_N\}$
$(1 \times N)$
For each **atom** in the GB and SOAP basis function:
$p_i = \int d^3r \rightarrow \vec{p} = \{p_1, p_2, \ldots, p_N\}$
One row for each atom
$\mathbb{P} =$
$(Q \times N)$
1.0
0.8
0.6
0.4
0.2
0.0
$f_c(R_{ij})$
0
2
4
6
R_{ij} (Å)
R_c = 2 Å
R_c = 4 Å
R_c = 6 Å

Figure 2.4: Left: Illustration of the process of extracting a SOAP matrix P for a single grain boundary (GB). Given a single atom in the GB, a Gaussian particle density function is placed at the location of each atom within a local environment sphere around the atom. Next, the total density function produced by neighbors is projected into a spherical basis consisting of radial basis functions and the spherical harmonics, as shown in the *boxed* region. Each basis function produces a single coefficient p_i in the SOAP vector $\vec{p}$ for the atom, the magnitude of which is represented in the figure by the colors of the arrays. Once a SOAP vector is available for all Q atoms in the GB, they are collected into a single matrix P that represents the GB. Figure reproduced from npj Computational Materials (2017) 3, 29, with permission from the authors. Right: Example of a trigonometric f_{cut} function for different cutoff parameters.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

of ML interatomic potentials, providing a robust framework for representing atomic structures and fostering a deeper understanding of material behaviors. Moreover, it is an integral part of the GAP method [55], which will be presented in the next section.

2.4.2 The Gaussian Approximation Potential Method

The purpose of this section is to serve as an introduction the GAP method, on which most of the results on interatomic potential developments of this thesis are based. Detailed explanation on GAP theory can be found in [55][54][56].

The GAP method was originally introduced by Bartók *et al.* in 2010 [55]. The objective of GAP is to model the Born-Oppenheimer PES of an atomic system by running a Gaussian process regression (GPR). As with other commonly used ML interatomic potentials, the GAP method consists of a set of descriptors and an ML model. An overview of the GAP method can be seen in Figure 2.5.

The most commonly used descriptors in the GAP method are 2-body (interatomic distances) and SOAP, which are sometimes combined with 3-body (distances between atoms and angles in triplets). The combination of simple geometrical descriptors and more complex mathematical representations of the atomic environments allow for high levels of accuracy, while reducing overfitting. A representation of the aforementioned descriptors can be found in Figure 2.6. Similarly to most ML interatomic potentials, GAP employs local atomic descriptors. Therefore, it is based on the hypothesis that, when computing the total potential energy, E_{total}, the long-range interatomic interactions, E_{long}, are negligible when compared to the short-range interatomic interactions, E_{short}, since $E_{\text{short}} \gg E_{\text{long}}$. In other words,

$$E_{\text{total}} = E_{\text{short}} + E_{\text{long}} \approx E_{\text{short}}. \tag{2.24}$$

The threshold distance that divides interatomic interactions between short and long is known as the cutoff range, r_{cut}. This is a system-dependent parameter which is empirically set. Therefore, the local atomic environment of a given atom i is defined

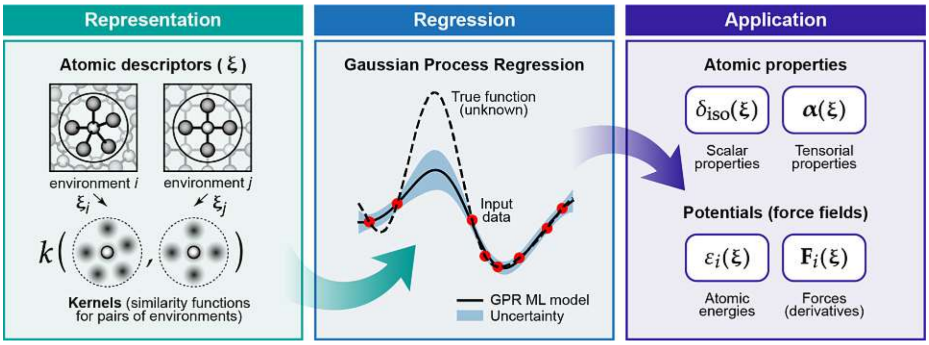

Figure 2.5: The three basic steps in GAP modeling: Representing the atomic structures by means of atomic descriptors, GPR and application. Original figure published under the CC BY 4.0 license. Credit: Deringer *et al.*, Chem. Rev. 121, 10073 (2021).

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

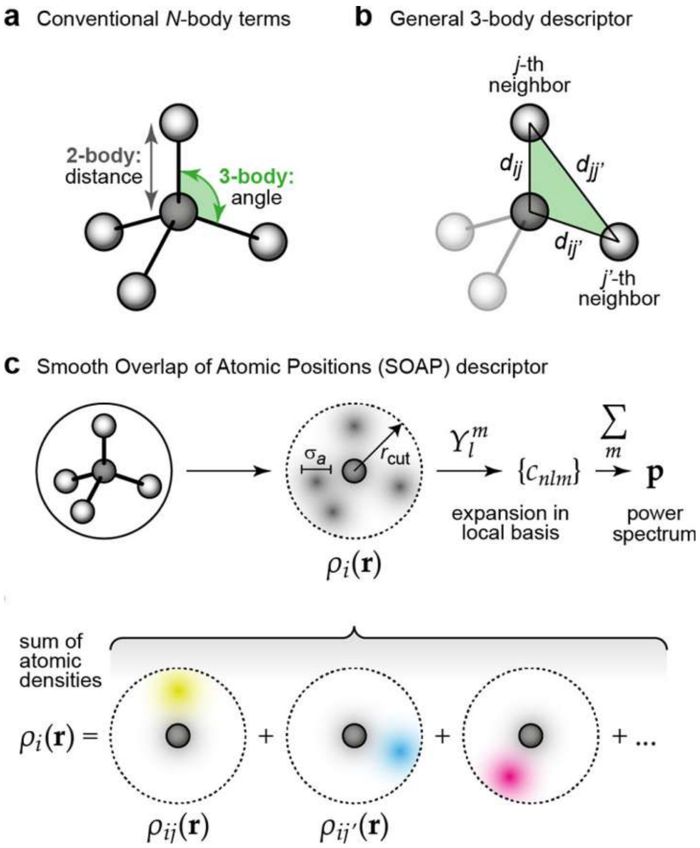

Figure 2.6: Descriptors for atomic structures commonly used in the GAP method: (a) conventional N-body terms (e.g., 2-body: interatomic distance), (b) General 3-body descriptor, and (c) SOAP. Modified from original figure published under the CC BY 4.0 license. Credit: Deringer *et al.*, Chem. Rev. 121, 10073 (2021).

as the sphere of radius r_{cut} around the atom i. Only the interaction with atoms inside this sphere are considered, while those outside are neglected.

Neglecting the long-range interatomic interactions means that only the local environment of each atom is relevant to compute the total potential energy [57],

$$E_{total} = \sum_{i=1}^{N} E_i(d_i), \tag{2.25}$$

where E_i is the local energy contribution to the total potential energy of atom i, and d_i is the local atomic environment descriptor of atom i. This hypothesis has been used since the first developments of ML interatomic potentials [17], and it is the base of most of the current methods found in the literature [58][59][60]. However, unlike other commonly used ML interatomic potentials, such as neural networks, GAPs do not have a fixed functional form. Instead, they estimate the local contribution of a given atom i, ϵ_i, as:

$$\epsilon_i = \delta^2 \sum_{s=1}^{N_s} \alpha_i k(d_i, d_s), \tag{2.26}$$

where d_i is the descriptor of the local environment of atom i, d_s (for s = 1, 2, ..., N_s) are the descriptors of the local atomic environments of the atoms in the training dataset, δ is a parameter defining the energy scale, $k()$ is a kernel function that measures the similarity between the atomic environments d_i and d_s, and α_i is the fitting coefficient. This means that GAPs are capable of modeling complex PESs without any previous knowledge or assumption about its functional form [55], since they are based on GPR. A schematic of a GAP is shown in Figure 2.7.

As presented in Figure 2.8, a GPR is composed of a set of observations on an unknown function. In the case of the GAP method, these observations are points on the PES of interest. Basis functions, characterized by their length, σ_{length}, are placed on each point, and an estimation to the unknown function is produced by fitting a set of coefficients to the observations. The benefit of this approach is that the potential energy, a non-linear function of the atomic positions, is re-written as a linear function in the kernels. This allows one to use linear algebra to obtain the fitting coefficients during the training stage [61]. However, the GAP method is particularly sensible to the selected kernel length, σ_{length}, as shown in Figure 2.9. If the kernel length is too small, the training will result in overfitting; therefore, σ_{length} must be properly set for the PES of interest. Training a GAP means finding the α values which minimize the error for the energies and forces of the atomic configurations in the training dataset, when compared against the theory level used to compute them (e.g., DFT).

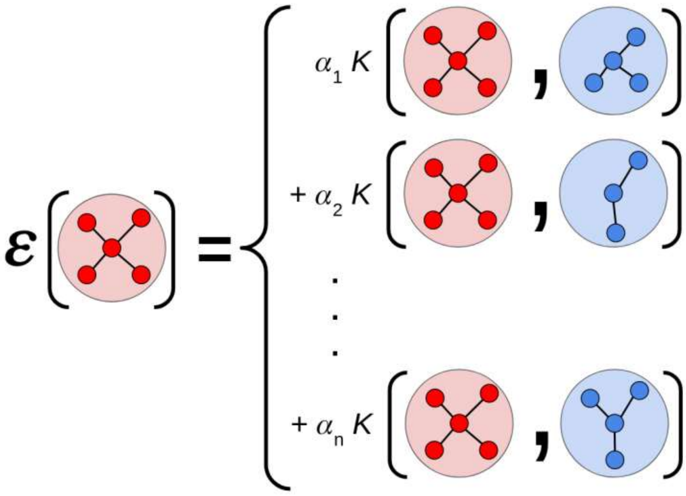

Figure 2.7: Schematic of a GAP used as a surrogate model to a PES. The input atomic configuration (red) is compared to every atomic configuration available in the training data-set (blue) by means of a similarity measure or kernel function (K). The energy of the input atomic configuration is then computed as the sum of these similarity measures, weighed by the vector α. The values of this vector are learned in the training process.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
Bibliothek Your knowledge hub
TU WIEN

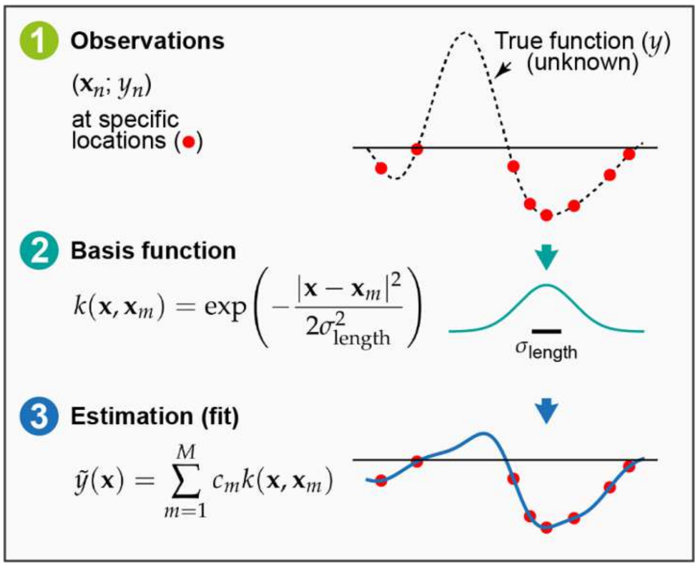
① Observations
$(\mathbf{x}_n; y_n)$
at specific locations (●)
True function (y) (unknown)
② Basis function
$k(\mathbf{x}, \mathbf{x}_m) = \exp\left(-\frac{|\mathbf{x}-\mathbf{x}_m|^2}{2\sigma^2_{\text{length}}}\right)$
σ_{length}
③ Estimation (fit)
$\tilde{y}(\mathbf{x}) = \sum_{m=1}^{M} c_m k(\mathbf{x}, \mathbf{x}_m)$

Figure 2.8: Basic elements of the GPR: (1) Observations, in the form of a training dataset composed of multiple data points, (2) Basis functions, centered at each data point, and (3) Estimation, produced by fitting a set of coefficients c_m to the observations. Original figure published under the CC BY 4.0 license. Credit: Deringer *et al.*, Chem. Rev. 121, 10073 (2021).

The training process can be seen as an optimization problem, where the set of variables to optimize are the α-values. Since there is only one α-value for each atomic configuration in the training dataset, training a GAP is relatively inexpensive and requires substantially less training data that other methods, such as neural networks. The combination of these two advantages allows for relatively rapid development times. However, since the kernel function must be computed for every atomic configuration in the training dataset, the computational cost of a GAP grows with the size of the training dataset. This is a disadvantage which is particularly challenging when active-learning schemes are used, since the number of atomic configurations in the training dataset can grow significantly.

At the present time, the GAP method has been successfully used to develop ML interatomic potentials for a wide range of atomic systems, such as amorphous carbon [62], gold [63], silicon [64], hafnium-dioxide [65], silicon:hydrogen [66], platinum [67] and iron [68]. A comprehensive list of the currently open-access available GAPs can be found in [69]. The versatility of this method allows building highly accurate yet computationally efficient ML interatomic potentials for any PES of interest. Moreover, it can be combined with active-learning algorithms to allow for iterative improvement, as schematized in Figure 2.10.

The results of this thesis show that a GAP can be combined with an efficient algorithm, resulting in a more efficient use of computational resources. In Chapter 3, the results of employing this approach to develop GAP interatomic potentials for materials for which this method has not yet been used are presented: silicon-nitride, silicon-germanium and silicon-dioxide.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

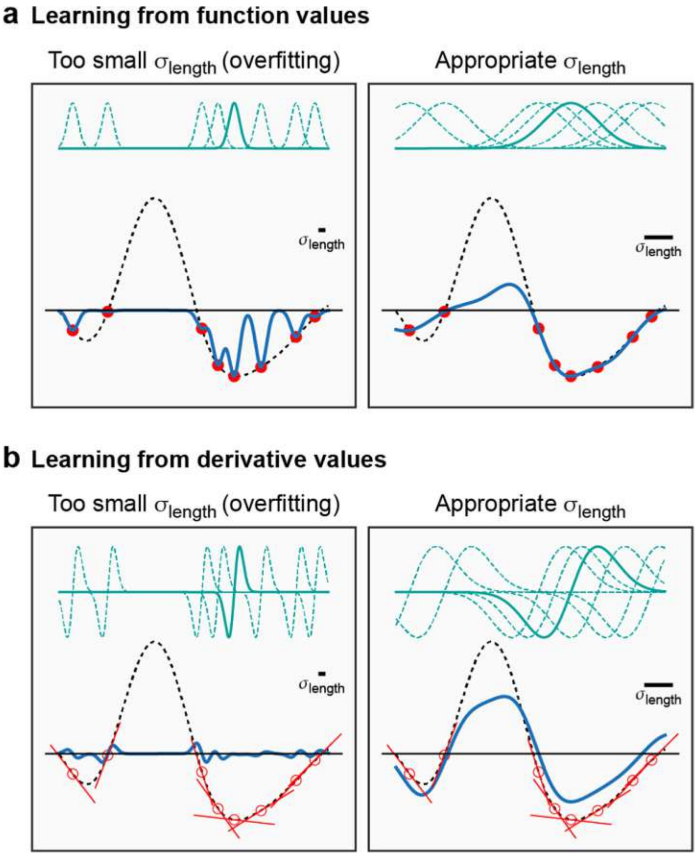

Figure 2.9: Effects of the kernel length scale, σ_{length} on the GPR for different types of data. The effects of using a small (left) and large (right) σ_{length} are illustrated for learning from function values (a) and from derivatives of the function values (b). The lengths of the basis functions are represented by a solid green line in each case. The overfitting produced by the small σ_{length} are clear by the poor fitting results obtained on the left traces. Original figure published under the CC BY 4.0 license. Credit: Deringer *et al.*, Chem. Rev. 121, 10073 (2021).

This section introduced and discussed the GAP method. This ML algorithm has interesting advantages, such as reduced training dataset size requirements and relatively short training times, which made it a good candidate for the developments presented in Chapter 3. However, GAP is not the only method available to develop accurate and efficient ML interatomic potentials, and an alternative algorithm might be better suited for other materials or applications. Moreover, this is a relatively young and highly dynamic field, in which new methods are constantly being developed. This section will discuss other alternative methods to GAP: artificial neural networks and message-passing networks, with the objective of briefly analyzing other options which were also considered before selecting GAP. More detailed reviews of established and emerging methods can be found in the literature [70][36].

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

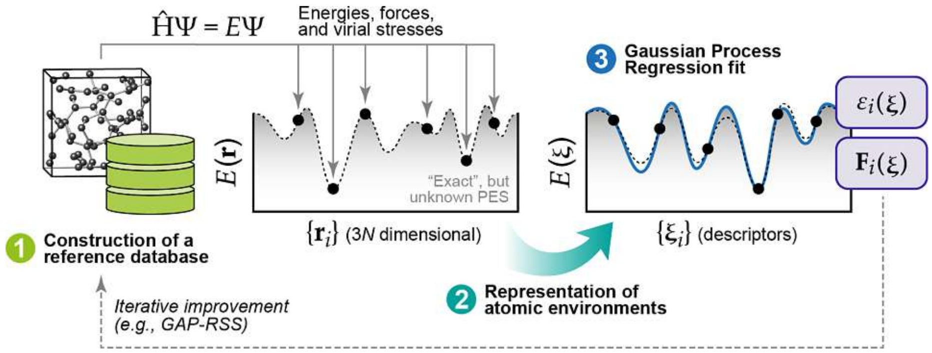
$\hat{H}\Psi = E\Psi$
Energies, forces, and virial stresses
3 Gaussian Process Regression fit
$\varepsilon_i(\xi)$
$\mathbf{F}_i(\xi)$
$E(\mathbf{r})$
"Exact", but unknown PES
$E(\xi)$
1 Construction of a reference database
$\{\mathbf{r}_i\}$ (3N dimensional)
$\{\xi_i\}$ (descriptors)
2 Representation of atomic environments
Iterative improvement (e.g., GAP-RSS)

Figure 2.10: Three main components for GAP: (1) training dataset, (2) a suitable mathematical representation of the atomic environments, and (3) the Gaussian process regression itself. By sampling the trajectories produced by running MD with the GAP and adding them to the training dataset together with their energies and forces, an iterative improvement on the accuracy of the potential is allowed. Original figure published under the CC BY 4.0 license. Credit: Deringer *et al.*, Chem. Rev. 121, 10073 (2021).

2.4.3 Artificial Neural Networks

An artificial neural network is an ML model, inspired in the way in which the human brain processes information. This method consists of two main components: neurons and synapses. Neurons can be seen as independent nodes in which the actual information processing takes places, while synapses are the connections between these nodes. The group of all neurons at the same level is known as a *layer*. Layers are classified, depending on their position in the neural network, into: *input*, *hidden* and *output* layers, as depicted in Figure 2.11.

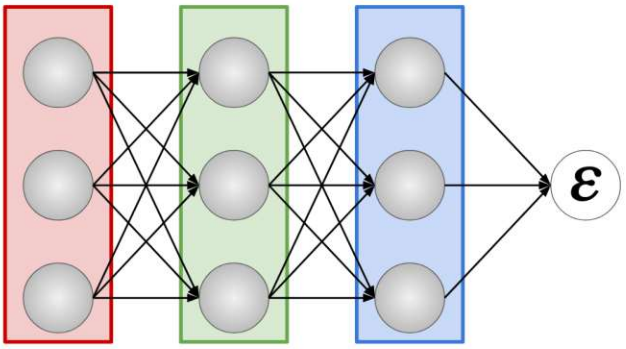

Figure 2.11: Schematic of a typical neural network used as a surrogate model of a PES. Red: input layer, green: hidden layer(s) (in this case, one), blue: output layer. The input to the neural network is a local atomic environment descriptor and the output is its local energy contribution. The information is processed in a feed-forward manner, as every neuron takes as inputs the outputs of the previous neurons and its output is a non-linear function of them. The connections between neurons are named *synapses* and their values are set during the training process.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Neurons can present a single or multiple inputs. However, they always present a single output, which is a non-linear function of its input(s), known as the *activation function*. There is a wide range of activation functions (e.g., linear, rectifier, sigmoid, etc.) and a single neural network can combine different types of them in different layers. On the other hand, synapses connect the output of the neurons in a layer to the inputs of the neurons in the following layer, and they are defined by a single value, known as *weight*. The higher the weight, the stronger is the connection between the neurons.

The classification of neural networks is done based on the amount of layers they present, the number of neurons in each layer and the way in which these layers are interconnected by synapses. The most common artificial neural network topology used in interatomic potentials is the multi-layer fully-connected perceptron. In this model, there is at least one hidden layer and the output of every neuron feeds the input of every neuron in the following layer. In these, and most commonly used artificial neural networks, the information is propagated from the inputs to the output, as every neuron receives its inputs from some or all neurons in the previous layer, and provides its output to some or all neurons in the following layer, in what is known as *feed-forward* information propagation.

Training a neural network is done by optimizing the weights of its synapses, so that the function which associates the inputs and outputs (in this case, atomic configurations and their energies/forces) can be better approximated. This is done by using a training algorithm (e.g., Adam [71]) to find the optimal values for the synapses.

Neural network based interatomic potentials (and ML interatomic potentials in general) can be classified into four different generations [72]. The first generation started in 1995, with the work of Blank *et al.* [73]. These neural network interatomic potentials were only applicable to low-dimensional PESs, that means, for systems containing only a few atoms. An example of these first neural networks can be seen in Figure 2.12. In these neural networks, the input is given by the coordinates of a 3-dimensional PES, $\mathbf{G} = \{G_i\}$. The outputs of the neurons in the first layer are given by applying the nonlinear activation function to the sum of each input, multiplied by its corresponding weight,

$$y_i^1 = f_i^1(b_i^1 + \sum_{j=1}^{3} a_{ji}^{01} G_j), \tag{2.27}$$

were y_i^1 is the output of the i^{th} neuron in the first layer, f_i^1 is the activation function of that neuron and b_i^1 is its bias, G_j is the j^{th} input to the neural network and a_{ji}^{01} is the weight of the synapse that connects G_j to the i^{th} neuron. Similarly, the outputs of the neurons in the second hidden layer are given by applying the nonlinear activation function to the sum of the outputs of the first hidden layers, multiplied by its corresponding weight. Combining these equations, the resulting energy as a function of the input PES coordinates is given by

$$E = f_1^3\{b_1^3 + \sum_{l=1}^{4} a_{l1}^{23} f_l^2[b_l^2 + \sum_{k=1}^{4} a_{kl}^{12} f_k^1(b_k^1 + \sum_{j=1}^{3} a_{jk}^{01} G_j)]\}. \tag{2.28}$$

This first generation of neural networks has been successfully used to model low-dimensional PESs of simple atomic systems, particularly molecules, as several examples in the literature show [74][75][76][77][78][79][80]. However, these initial neural networks suffered from a series of issues when applied to more complex atomic systems. In first place, once a neural network has been build, the number of atoms in the system cannot be changed, since adding or removing an atom would change the dimensions of the PES, which would change the dimension of the input vector, rendering it incompatible with the number of synapses connecting it to the input layer of the neural network. This drastically reduces the versatility of the neural network, as it could only be used to compute the energies and forces on atomic systems with the same number of atoms as those used to train it. In second place, increasing the dimensions of the PES would inevitably increase the size of the neural network, therefore making it computationally less efficient. Finally, these neural networks require that the order in which the coordinates of the PES presented at its input are always respected, since it is not invariant with respect to the permutation of identical atoms. These disadvantages rendered the first generation of neural network interatomic potentials impractical to model mode complex PESs, therefore fostering the development of the second generation of neural network interatomic potentials.

The second generation was introduced in 2007 by Behler and Parrinello [17], in which the concept of nearsightedness was introduced, i.e., considering the total potential energy of the system as a sum of local contributions from each atom, each of which depends only on a local atomic environment. These were the first neural network interatomic potentials that could be used to model highly-dimensional PESs. The second generation is contemporary with the previously discussed GAP method.

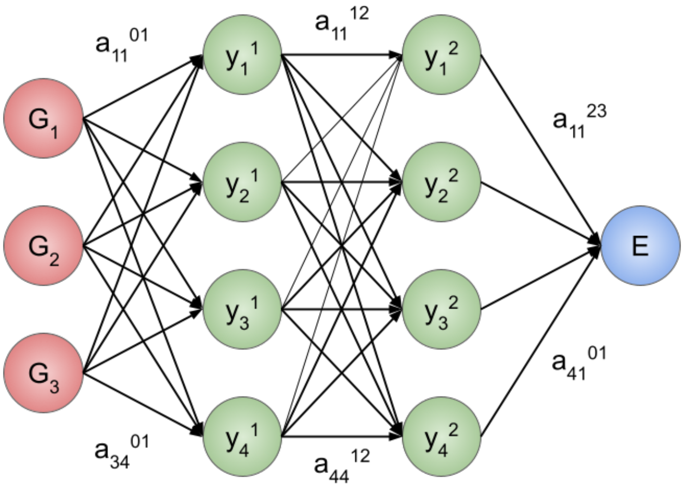

Figure 2.12: Illustration of a three-dimensional feed-forward neural network featuring dual hidden layers, each comprising four neurons. The connections between neurons are depicted by arrows, symbolizing the adjustable parameters (weights) crucial for the network's functioning. To enhance clarity, bias weights have been omitted from the visualization.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

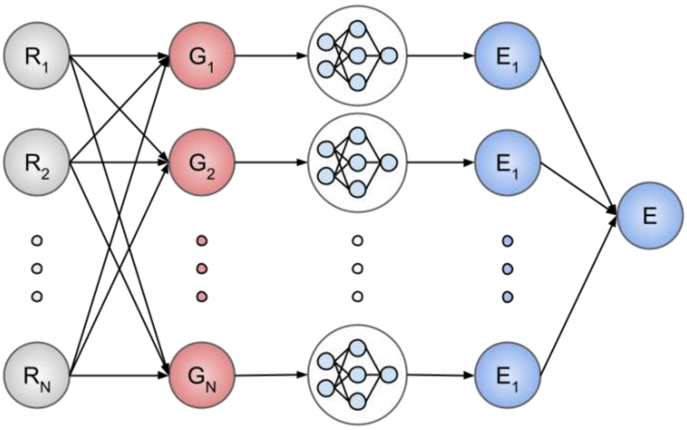

Figure 2.13: Architecture of the second generation neural network interatomic potential, applied to a system with N atoms. The output corresponds to the total energy (E), which is the aggregate of local atomic energy contributions (E_i). Each local atomic energy contribution is determined by the output of a dedicated neural network. The input for each atomic neural network consists of a vector describing the local chemical environment of atom i, computed from the coordinates of the atoms in the system (R_1, R_2, ..., R_N).

A schematic of a second generation neural network can be seen in Figure 2.13. The novelty of the second generation neural networks is that the total potential energy of the system, E, is computed as a sum of local energy contributions, E_i, one from each atom,

$$E = \sum_{i=1}^{N} E_i. \tag{2.29}$$

The local atomic energy contributions are computed by a neural network, from the local descriptor of each atom in the system, G_i,

$$E = \sum_{i=1}^{N} E_i(G_i). \tag{2.30}$$

These descriptors are built in turn based on the coordinates of the atoms in the system, $\mathbf{R}$,

$$E = \sum_{i=1}^{N} E_i(G_i(\mathbf{R})). \tag{2.31}$$

The advantage of this approach is that the neural network interatomic potential can be used for any number of atoms in the system. In the first generation of neural networks, if an atom is removed from the system, or added to it, the neural network cannot be used anymore. In the second generation, adding or removing an atom from the system simply translates in computing one more or less descriptor and evaluating its local energy contribution with the neural network. The second generation

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

of neural networks have been successfully used to develop interatomic potentials for sodium [81], carbon [82] and silicon [83][84], among other materials.

The main limitation of the second generation of neural network interatomic potentials is that they neglect long-range interactions, *e.i.*, beyond the cutoff radius. This issue was addressed in the third generation, by constructing environment-dependent atomic charges. The third generation of neural network interatomic potentials started in 2011, with the work of Artrith *et al.* [85] and Morawietz *et al.* [86]. At the present time, only few examples can be found of third generation neural network interatomic potentials in which the Coulomb interactions are computed without truncation. Instead, the most common way to account for the long-term interactions is to train a second neural network which computes the atomic charges based on the same local atomic environment descriptors used to compute the local energy contributions [87],

$$\begin{aligned} E_{\text{total}} &= E_{\text{short}} + E_{\text{elec}}, \\ E_{\text{total}} &= \sum_{i=1}^{N} E_i(G_i) + \sum_{i\,j}^{N} \frac{q_i(G_i)q_j(G_j)}{R_{ij}}, \end{aligned} \tag{2.32}$$

where E_{total} is the total potential energy, E_{short} is the sum of all local energy contributions, E_{elec} is the long-range energy contribution, G_i/G_j is the descriptor of the local atomic environment of atom i/j, q_i/q_j is the atomic charge of atom i/j, and R_{ij} is the interatomic distance between atoms i and j.

The main limitation of the third generation of neural network interatomic potentials is their inability to describe long-range charge transfer and different charge states of a system. This is because the atomic partial charges are expressed as a function of the local chemical environment only. The inclusion of non-local charge transfer in the models has given rise to the fourth generation, which has been successfully used to develop interatomic potentials for several system types, from organic molecules to ionic solids [72]. Based on the third generation, the fourth generation computes the total potential energy of the system as the sum of a short-range and a long-range component. They operate by deriving a descriptor from the coordinates of the atoms in the system, which is then utilized by a neural network to assess atomic electronegativities. Employing a charge equilibration method, the atomic charges are subsequently calculated based on these electronegativities. The long-range energy component is then determined from the atomic charges, typically through the application of Coulomb's equation. The short-range energy component is computed as the sum of local atomic energies, with one computed for each atom in the system. Similar to previous generations, individual local atomic energies are computed by neural networks, utilizing descriptors of the local atomic environment. However, in the fourth generation, these neural networks also incorporate atomic charges as part of their inputs. A schematic of these neural networks is depicted in Figure 2.14.

All neural networks discussed up to this point have the use of descriptors with a static functional form in common. These descriptors typically contain a set of hyperparameters which allow the user to optimize them for the intended application. However, these hyperparameters are normally tuned before training the neural network interatomic potential and remain fixed during this process. In contrast to

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

this, beyond the fourth generation, the field has seen the development of learnable descriptors. This idea was first introduced by Duvenaud *et al.* [88] in 2015. Taking inspiration from extended-connectivity fingerprintss (ECFPs) [89], they approached the representation of molecules by conceptualizing them as graph networks. To process these graphical representations and predict their final feature vectors (descriptors), a convolutional layer was employed. While they did not utilize these methods explicitly

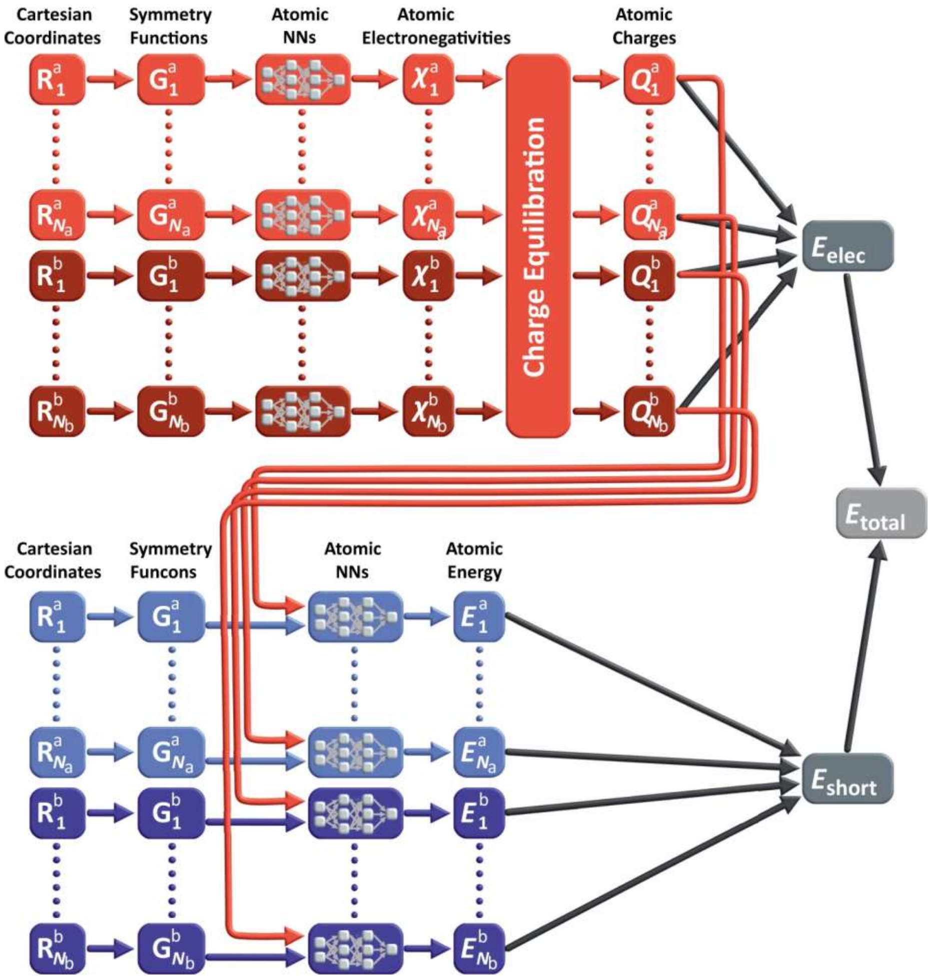

Figure 2.14: Schematic of a fourth generation neural network interatomic potential for an atomic system containing N_{a} and N_{b} atoms of two different elements. The total potential energy, E_{total} is the sum of the local atomic contributions, E_{short} and the long-range contributions, E_{elec}. Each local atomic contribution is computed by a neural network from a descriptor of the local atomic environments, G. A second neural network is used to compute the atomic electronegativities from the same descriptors, from which the atomic charges are computed using a charge equilibration method. The long-range contributions are calculated based on the resulting atomic charges. Original figure reproduced from Nature Communications vol. 12 no. 398 under Creative Commons Attribution 4.0 International License. Credits: Tsz Wai Ko *et al.*

TU Bibliothek, WIEN Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

for modelling PESs, their efforts laid the foundation for a novel category of neural network interatomic potentials. In 2017, Gilmer *et al.* introduced the term *message passing neural networkss (MPNNs)* [90] to characterize this class. MPNN methods share a common trait wherein predefined descriptors are substituted with automatically derived descriptors learned from the geometric structure.

In contrast to previous generations, the input to a MPNN is not an atomic environment descriptor but a graph, in which its nodes represent the atoms and its edges the bonds in the atomic structure [53]. A typical MPNN applied to compute the potential energy of an atomic system works by following these steps:

(I) A graph network is built based on the input atomic structure, in which nodes represent atoms and edges represent bonds.

(II) The graph network is initialized with information extracted from the input atomic structure. Nodes are commonly initialized with their corresponding atomic number, covalent radius, number of valence electrons, etc. Edges are commonly initialized with their corresponding bond length.

(III) Information is propagated along the graph in messages. Each node receives messages from neighboring nodes with information on their internal feature vectors. In the first iteration, nodes receive information only from the neighbor nodes. In the second iteration, information of their neighbors and the neighbor of their neighbors, etc. This process is iterated N times.

After N iterations, the feature vector for each atom has been updated based on information from its N^{th}-nearest neighbors [70]. This type of neural network interatomic potential is represented in Figure 2.15. One of the main advantages of MPNNs is their natural ability to account for long-distance interatomic interactions. This is in contrast with traditional methods to develop ML interatomic potentials, in which the only way to account for long-distance interatomic interactions is to increase the cutoff radius of the local atomic environment when computing the descriptor of every atom, therefore increasing the computational power requirements.

Artificial neural networks remain one of the most commonly used methods to develop ML interatomic potentials, due to their high accuracy and versatility. The size of the artificial neural network (i.e., the number of layers and neurons in each layer) can be easily modified to reach the best possible cost-to-accuracy ratio. The popularity of this method in developing ML interatomic potentials is further aided by the wide availability of different software packages which can be used to build a surrogate model of the PES of interest. Some available software packages are DeepMD [58], PANNA [59], AisNet [91], AMP [60] and SIMPLE-NN [92]. The topic of artificial neural networks for MD simulations will be addressed again in Chapter 5, where this method will be used in the development of an ML interatomic potential for amorphous germanium, due to its compatibility with neuromorphic hardware accelerators for MD simulations.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

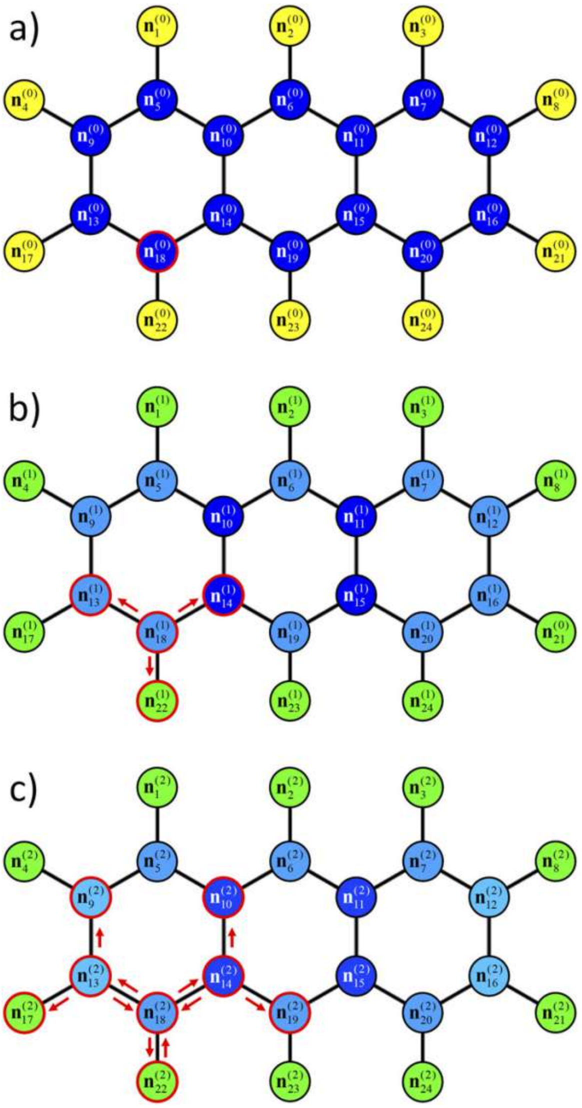

Figure 2.15: Basic MPNN applied to an anthracene molecule, showcasing (a) the initial configuration with yellow hydrogen atoms and blue carbon atoms, (b) the first iteration, and (c) the second iteration. In this model, a straightforward update mechanism is employed, where the hue of each node is an average of its original hue, its current hue, and the hues of its neighboring nodes. The feature vectors for the edges (represented by straight black lines) are omitted for clarity. Red arrows highlight the propagation of information regarding the initial state of node 18 through the network. Nodes whose current state have been influenced by the initial state of node 18 are denoted by red circles. Reprinted from Journal of Chemical Physics 152, 050902 (2020), with the permission of AIP Publishing. Credits: Tim Mueller *et al.*

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

2.4.4 Moment Tensor Potentials

The starting hypothesis of most ML based interatomic potentials is that there is a function that correlates local atomic environments with their local potential energies. This function acts as a computationally more efficient approximation to the PES and can be used to compute the total potential energy of a given atomic configuration, by summing all the local potential energies, contributed by every atom in the configuration.

The key concept behind moment tensor potentials, introduced by Shapeev in 2016 [93], is that this aforementioned function can be built as a linear combination of polynomial basis functions which represent one-, two- and three-body interactions between atoms in the given configuration,

$$V(R, Z) = \sum_{i=1}^{N} V_{\text{local}}(R - r_i, Z, z_i), \tag{2.33}$$

where V is the total potential energy of the atomic configuration, V_{local} is the function that computes local potential energy contributions from local atomic environments, r_i are the coordinates of the i^{th} atom, R is the set of coordinates of all atoms in the configuration, z_i is the atomic number of the i^{th} atom and Z is the set of all atomic numbers in the configuration. This expression is expanded in Eq. (2.34),

$$V_{\text{local}}(R - r_i, z, z_i) = \sum_{i=1}^{N} V_\alpha B_\alpha(R - r_i, Z, z_i), \tag{2.34}$$

where V_α is the linear coefficient and B_α is the basis function. The training of a moment tensor potential is an optimization process, in which the coefficients that multiply every basis function in the weighed sum are learned. In a manner similar to the learning process of kernel-based methods, these weights are optimized to reduce the prediction errors. Although moment tensor potentials are a relevant part of the state-of-the-art of ML interatomic potentials, they fall outside the scope of this thesis. The mathematical derivations of this method can be found in the original paper by Shapeev [93], and a discussion of their impact in the field can be found in the review by Müller *et al.* [70].

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Chapter 3

Gaussian Approximation Potential Results

The previous chapter presented the theory of the GAP method. In the present chapter, the results of employing this method to develop interatomic potentials for silicon nitride (Si_3N_4), silicon-germanium (SiGe), the oxidation of Si and silicon dioxide (SiO_2) are presented. The high accuracy and low computational costs found for these different atomic systems are a strong indication of the potential of this method.

3.1 GAP for Silicon-Nitride (Si_3N_4)

The results and figures presented in this section have been partially published in the Journal of Chemical Physics [94] and Nanomaterials [95].

Silicon nitride (Si_3N_4) is a material with a high melting-point (1900 °C), which finds applications as storage layer in non-volatile charge trap flash devices [96], in the automotive industry, where it is used in engine fabrications [97], in the renewable energy industry, as an anti-reflective coating for solar cells [98], in the engines of the NASA Space Shuttle, as armorial-bearing [99], and in microelectronics, as a dielectric in nanoscale metal oxide semiconductor field-effect transistors (MOSFETs) [100]. In this last application, it is particularly valued for its use in the reduction of leakage currents and the prevention of boron diffusion. Moreover, Si_3N_4 is more stable than other Si-nitrogen compounds [101][102], and it can crystallize in different phases, such as the trigonal α-Si_3N_4, the hexagonal β-Si_3N_4, and the cubic γ-Si_3N_4.

In order to study properties of this material at the atomistic scale, MD simulations are frequently employed. Examples of previously designed empirical potentials for Si_3N_4 include Tersoff [103], Billeter [104], Born-Mayer-Huggins (BMH) [105] and ReaxFF [106]. These empirical potentials are orders of magnitude faster than DFT and accurate enough for the specific applications for which they were developed, such as studying diffusion [103], the interactions with H atoms [103] and interfaces with other materials [104]. However, their accuracy is insufficient for several other important applications, including modeling amorphous Si_3N_4, which is relevant in the semiconductor industry, since most thin films deposited for use in transistors are amorphous. In addition, empirical potentials sometimes present systematic errors in describing important structural features of amorphous Si_3N_4, such as the coordination number. The modeling of amorphous Si_3N_4 is typically done using the melt-and-quench technique, which is the direct motivation for the development of this universally applicable ML-based potential.

In order to create realistic amorphous structures, *ab initio* accuracy combined with comparatively slow quenching rates of 5 K/ps or below are required [107][65]. This directly translates to high numbers of time-steps ($>$ 1,000,000), which are out of range for DFT calculations due to its high computational costs. On the other hand, empirical potentials are a very useful tool to run these extensive MD simulations due to their low computational costs; however, their accuracy is not always near that of *ab initio* methods.

Here, a general purpose ML interatomic potential for Si_3N_4 is presented, with a special focus on modelling its amorphous phase. The employed approach combines an established ML model with a simple active-learning algorithm in a bootstrap process. This allowed to semi-automatically develop an accurate ML interatomic potential, starting from an inexpensive empirical potential, poorly optimized for amorphous Si_3N_4. The proposed ML interatomic potential is validated by comparing structural properties of the resulting amorphous Si_3N_4 models to those of DFT calculations, empirical potentials and experimental data found in the literature. This is the first time an ML interatomic potential is developed for Si_3N_4.

In the development of this interatomic potential, the GAP method was paired with a set of three descriptors: SOAP [54], *two-body* and *three-body*. The two-body descriptor is the interatomic distance between atoms within the cutoff distance, while the three-body descriptor is a triangle-like representation of the three sides of a triangle of three atoms. The implementation of GAP and its descriptors was done using the software package QUIP [108]; its most relevant parameters are summarized in Table 3.1.

Table 3.1: Parameters used to implement the GAP and its SOAP, three-body and two-body descriptors.

Parameter	SOAP	Three-body	Two-body
δ (eV)	0.4	1	4
r_{cut} (Å)	4	3	4
r_Δ (Å)	1	-	-
n_{max}	8	-	-
l_{max}	4	-	-
ζ	4	-	-

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Among the most relevant parameters in the training process, δ is the scaling of the kernel, r_{cut} is the cutoff of the descriptor, r_{Δ} is the region in which the cutoff function goes smoothly to zero, n_{max} and l_{max} are the number of angular and radial basis functions of the SOAP descriptor, and ζ is the power the kernel is raised to.

To make use of the proposed GAP, the reader is referred to the open-access repository, where the ML interatomic potential and the training dataset are publicly available [109]. The proposed active-learning algorithm is based on an iterative re-training approach, derived from the original algorithm of Deringer *et al.* [57]. The algorithm allows starting with a simple and inexpensive empirical potential for an initial sample of the PES of interest, to eventually build up an ML interatomic potential with near *ab initio* accuracy. This approach aims at saving computational time, since it avoids running an initial MD simulation with DFT. Moreover, due to its iterative nature, this active-learning algorithm automatically runs and improves the ML interatomic potential until the desired accuracy is reached.

For the generation of the first training dataset, an MD simulation was run with the inexpensive empirical potential Billeter [104], in QuantumATK [110]. During this initial run, an N-N repulsive term was included to suppress the formation of N-N bonds, with the aim of reducing the number of under-coordinated atoms in the resulting amorphous structures. This initial MD started from a defect-free α-Si_3N_4 crystal containing 224 atoms at 300 K, which was thereafter heated up above its melting point to 5,000 K and subsequently quenched back to 300 K in 1,000,000 time-steps, using a step-size of 0.5 fs, resulting in a quench rate of 4.7 K/fs. The resulting trajectory was subsequently sub-sampled, leading to 565 atomic configurations. Furthermore, single-atom configurations for Si and N (needed for the GAP to learn the correct free energies) were added to the training dataset, together with 267 dimers (Si-Si, Si-N and N-N).

The energies and forces of these atomic structures were re-calculated with DFT in the software package CP2K [111], using the Perdew-Burke-Ernzerhof (PBE) func-

Table 3.2: Atomic structures added to the training dataset in every iteration.

Iteration Nº	Structure type: number of configurations
1	Single atoms: 2; dimers: 267; initial MD: 565
2	Non-homogeneous amorphous Si_3N_4: 39
3	Amorphous Si_3N_4 with N-chains: 82
4	Amorphous Si_3N_4 with N-clusters: 30
5	Crystal α-Si_3N_4: 99
6	Bulk Si: 87
7	Amorphous Si_3N_4: 101
8	Geometry optimized Si_3N_4: 18
9	Amorphous Si_3N_4 with under-coordinated Si: 60
10	Amorphous Si_3N_4 with over-coordinated Si: 49
11	Vibration of hexagonal β-Si_3N_4: 82
12	Amorphous Si_3N_4 with density 2.71 g cm^{-3}: 141

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

tional [112], in conjunction with the Goedecker-Teter-Hutter (GTH) pseudopotentials [113]. Once the calculations for the initial training dataset were finished, a GAP was trained on it. This GAP was used to run a series of sequential melt-and-quench MD simulations. After each MD, the last snapshots from the trajectory were saved and their energies, forces and/or stress tensors were re-calculated using DFT and added to the training dataset. The GAP was re-trained on the extended training dataset before running the next MD and the process was repeated until the desired accuracy was reached. This technique was applied to α-Si_3N_4 and β-Si_3N_4 with different densities, as well as to bulk silicon, in order to provide a thorough sampling of the PES of interest. Table 3.2 shows the structure types added in each iteration.

The selection of the atomic configurations from which to start running the MD is manually done by the user, based on the application of interest. Thereafter, the sampling of the PES is automatically done by the algorithm, therefore constituting a semi-automatic active-learning method.

The results presented in the right panel of Figure 3.1 show how GAPs from earlier iterations of the training process were prone to create non-homogeneous atomic structures when used to model amorphous Si_3N_4. Furthermore, GAPs from intermediate iterations still yielded chains and clusters of Si and N. These unphysical features were not present in the quenched structures after the final iteration of the retraining cycle. Table 3.3 shows the energy, force and stress σ parameters, which determine the

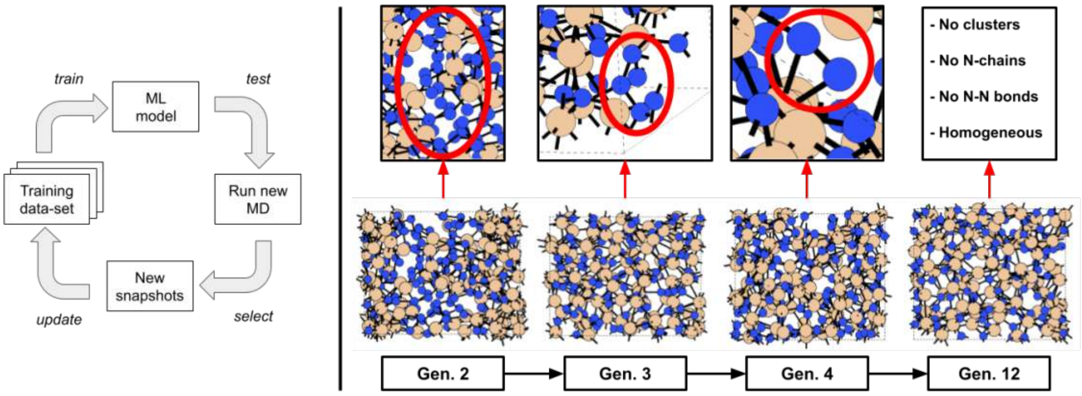

Figure 3.1: Schematic of the iterative re-training algorithm used to develop the ML interatomic potential for Si_3N_4 (left) and results obtained over time for subsequent generations (right). The ML interatomic potential is used to run MD; thereby generating a trajectory of atomic configurations. The last snapshots of this trajectory are collected, their energies, forces and/or stress tensors are recalculated with DFT and then added to the training dataset. The ML interatomic potential is thereafter re-trained and used to run a new MD, thereby repeating the process until the desired accuracy is reached. The right panel presents the improvement in the results of using the ML interatomic potential for successive re-training iterations. The earliest iterations yield unphysical heterogeneous distributions of Si and N; intermediate iterations show unphysical N-chains and $N-N$ bonds; the final potential overcomes all aforementioned issues and generates realistic and defect-free structures of Si_3N_4.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 3.3: Energy, force and stress convergence parameters for every structure type used in the training process

Structure type	σ_{energy}	σ_{force}	σ_{stress}
Dimers	0.01	0.1	-
Amorphous Si_3N_4 with N-clusters	0.01	0.1	-
Non-homogeneous amorphous Si_3N_4	0.01	0.1	-
Crystalline	0.002	0.02	-
Amorphous Si_3N_4 with N-chains	0.002	0.02	-
Bulk Si	0.002	0.02	-
Initial MD	0.002	0.02	-
Hexagonal β-Si_3N_4	0.002	0.02	0.02
Amorphous Si_3N_4 ($2.71\,\mathrm{g\ cm^{-3}}$)	0.002	0.02	0.02
Amorphous Si_3N_4 with over-coordinated Si	0.0005	0.005	0.005
Amorphous Si_3N_4 with under-coordinated Si	0.0005	0.005	0.005
Final amorphous Si_3N_4	0.0005	0.005	0.005
Geometry optimized Si_3N_4	0.0005	0.005	-
Single atoms	0.0001	0.001	-

convergence criteria in the GAP training process used for each structure type. Note that the DFT stress tensors were only calculated for the last five iterations.

The training process of the GAP consists in an optimization problem: finding the set of α values which produce the lowest errors for the atomic configurations in the training dataset. However, if not specified, the convergence parameters will be the same for all atomic configurations, in which case the GAP would overfit to the atomic configurations with the highest energies and/or forces, since the optimization algorithm would be biased towards reducing the nominal error of these atomic configurations. Therefore, the data is divided into 4 blocks, each with distinct σ parameters, as a measure to avoid overfitting to configurations with higher energies and/or forces.

The resulting GAP was validated against energies and forces from DFT calculations for the atomic structures in a testing dataset; the results are shown in Figure 3.2. The testing dataset contains 1,000 liquid and amorphous atomic configurations, randomly selected from the initial MD simulation, none of which were used in the training process. The deviations are quantified by computing the mean absolute error (MAE) between both calculation methods. The energies and forces calculations yielded a MAE of 8 meV/atom and 0.26 eV/Å, respectively. Moreover, the clear linear correlation between GAP and DFT indicates no systematic error.

The GAP was used to run an MD simulation, starting from a crystalline Si_3N_4 system composed of 224 atoms. The initial velocities were drawn from a Maxwell-Boltzmann distribution at 300 K. Then, the system was heated to 5,000 K. An equilibration phase continued at 5,000 K for 10,000 time-steps, followed by a quenching phase over 1,000,000 time-steps, driving the system back to a temperature of 300 K. A time-step of 0.5 fs was used, resulting in a quenching rate of 4.7 K/fs.

TU Bibliothek WIEN Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

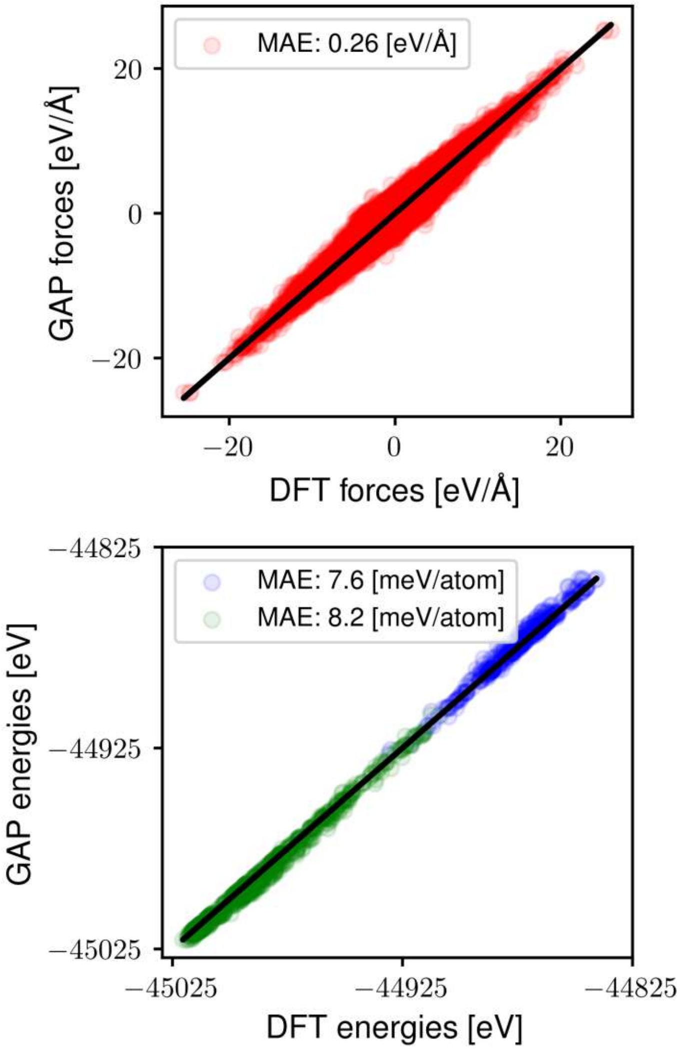

Figure 3.2: ML interatomic potential tested against DFT for forces (top) and energies (bottom) calculation of the atomic configurations in the testing dataset, together with their respective MAE. Atomic configurations were divided into liquid (blue) and amorphous (green).

A DFT reference atomic structure was obtained by running a geometry optimization of the final snapshot of the melt-and-quench MD performed with the modified Billeter empirical potential [104]. The geometry optimization was done using the PBE [112] functional, in CP2K [111]. This technique to build a reference DFT structure was preferred over running a DFT MD, due to the high computational costs of the latter. As will be shown next, this reference structure correlates well with experimental data.

Both the GAP and DFT atomic structures were first compared by computing their total and partial radial distribution functions (RDFs); the results are presented in Figure 3.3. The top panel compares the total RDF (i.e., all atomic species were con-

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

sidered) for GAP, DFT and the empirical potentials Tersoff [103], ReaxFF [106] and Billeter [104]. The amorphous Si_3N_4 structures for each of these empirical potentials were obtained by running melt-and-quench MD simulations in QuantumATK [110], with a NoseHoover NVT thermostat, starting from exactly the same initial Si_3N_4 atomic configuration used for the GAP MD simulation. The same recipe was used as for the GAP MD simulation, with identical time-step and quench-rate. Moreover, an amorphous Si_3N_4 structure generated with the BMH empirical potential obtained from [105] was added to the comparison. This BMH structure contains 112 atoms, a density of 2.9 g/cm^3, and it was generated using a quench rate of 0.1 K/fs. In addition, the results were validated by comparing them with experimental data found in the literature, measured with a technique based on x-ray diffraction of the structure of chemically vapor-deposited amorphous Si_3N_4 [114].

The amorphous Si_3N_4 structures created with the ReaxFF and Billeter empirical potentials yield high peaks for short interatomic distances, which do not correlate with experimental data. The Tersoff and BMH empirical potentials yield a good agreement to the DFT reference, as well as the experimental data, over the entire range of relevant interatomic distances. The GAP correlation to these references is slightly better than that of the empirical potentials, particularly if the height of the first and secondary peaks is considered.

Figure 3.4 compares the distributions of Si-N bond-lengths and N-Si-N bond-angles of these atomic structures. The distributions in the structure generated with the BMH empirical potential are added for reference purposes. The bond-lengths distribution is virtually identical for GAP and DFT, both share minimum and maximum values at 1.6 A and 2.4 A, respectively, as well as a single peak at 1.79 A for DFT and 1.78 A for GAP. On the other hand, for the bond-angle distributions, the minimum and maximum values for DFT/GAP are 72.0°/75.2° and 176.3°/176.8°, respectively. The mean angle for GAP is 109.0°, while for DFT it is 109.2°. Overall, the GAP yields a better agreement to the DFT reference than the BMH empirical potential, which slightly over-estimate the angles and under-estimate the bond-lengths.

A further validation test was performed by computing the neutron scattering factor for the GAP and DFT amorphous Si_3N_4 structures. The results are shown in Figure 3.5, where experimental data for an amorphous Si_3N_4 sample with a density of 2.60 gcm^{-3} [114] was added for reference purposes. Note that the density of the GAP amorphous Si_3N_4 was 2.92 gcm^{-3}. The similarity between the GAP results and the experimental data is remarkable for low values of q. For higher values, the discrepancy is more noticeable; however, the GAP and DFT results remain in good agreement, therefore hinting at limitations at the theory level and not in the training of the ML interatomic potential. If more accurate results were needed for this, or any other specific application, the same framework could be used with a different theory level [65].

The main application of the ML interatomic potential is to be employed in melt-and-quench MD simulations to model amorphous Si_3N_4 with vanishing defect densities, starting from different initial configurations, which vary in crystal type and density. Therefore, to analyze the defects in the structure generated with the potential, their coordination numbers were analyzed and compared to that of the DFT

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

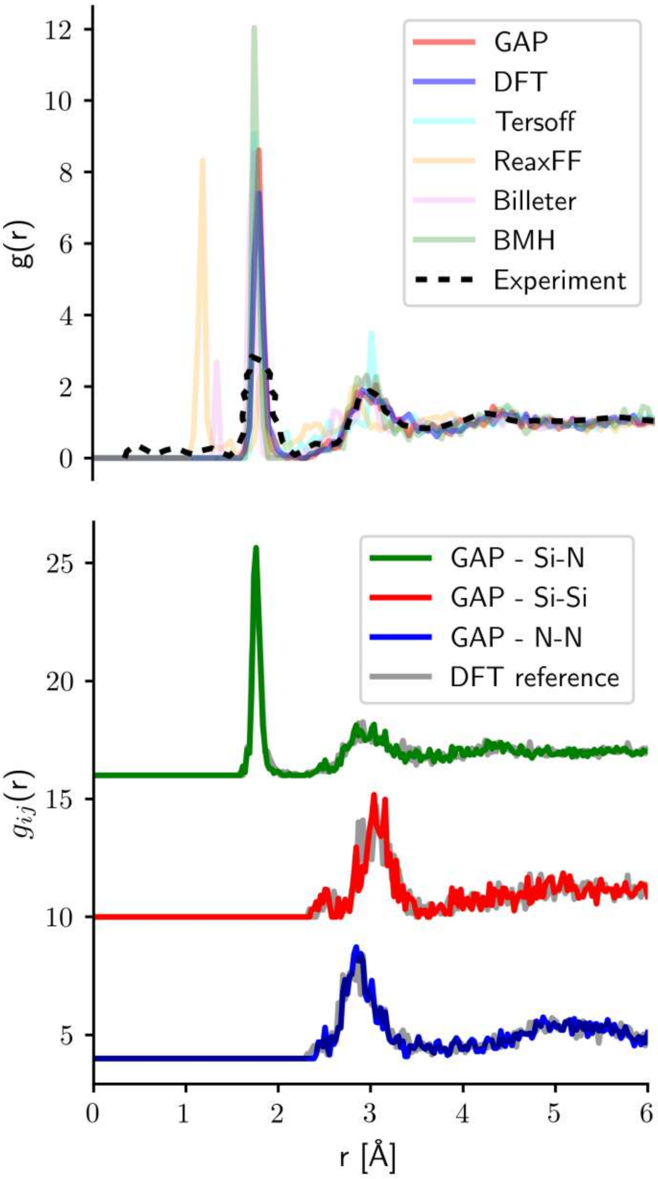

Figure 3.3: Top: Total RDFs for structures generated with GAP (red), DFT (blue), experimental data (black), Tersoff (cyan), ReaxFF (yellow), Billeter without added terms (violet) and BHM (green). Bottom: Partial RDFs of Si-N (top), Si-Si (middle) and N-N (bottom) bonds for GAP (color) and DFT (shadow). For the total RDF (top), GAP presents the best agreement with the DFT reference and experimental data, closely followed by BMH. For the partial RDFs (bottom), GAP presents a good agreement with the DFT reference for the first and second peaks of the three relevant bonds.

reference structure. Figure 3.6 shows the coordination number of silicon atoms (left) and of nitrogen atoms (right). The qualitative similarity between both coordination number distributions is striking.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

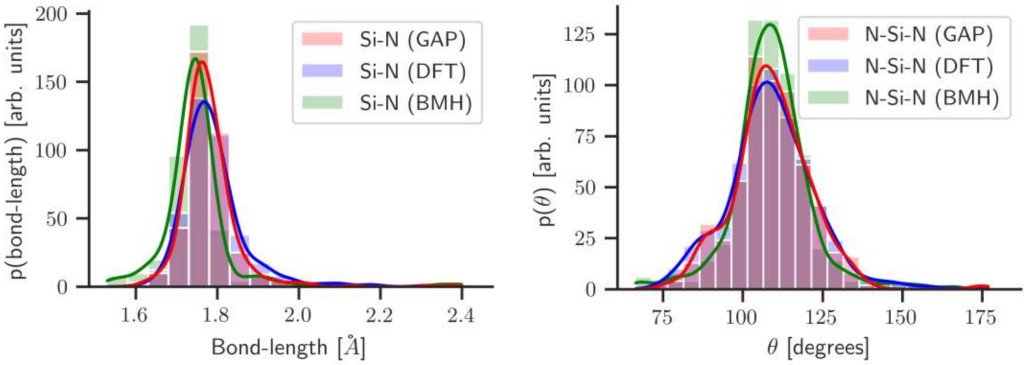

Figure 3.4: Si-N bond-length distributions (left) and N-Si-N bond-angle distribution (right) for amorphous Si_3N_4 structures generated with GAP (red), DFT (blue) and BMH (green).

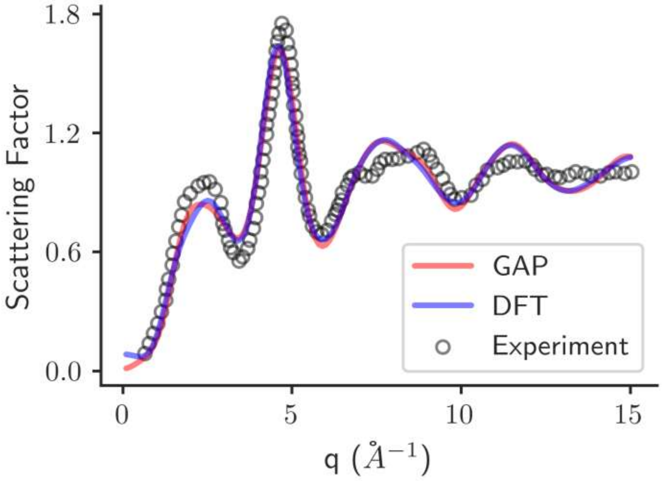

Figure 3.5: Comparison of the structure factor obtained from neutron scattering for the amorphous Si_3N_4 generated with GAP (red) and DFT (blue), together with experimental data found in the literature.

In order to further test the proposed ML interatomic potential, the vibrational density of states (VDOS) of the amorphous Si_3N_4 structure generated using the melt-and-quench technique was computed. The finite difference calculations were carried out by displacing a single atom by 0.01 Å along each coordinate axis. The resulting dynamical matrix was then diagonalized in mass-weighted coordinates to obtain the system's normal modes and the corresponding phonon energies. The results of these calculations are presented in Figure 3.7. The VDOS spectra obtained by DFT calculations and the ML interatomic potential match over the relevant energy range.

The good match is partially due to including forces in the training process, as this provides the ML interatomic potential with the explicit value of the first derivative of the PES with respect to the atomic coordinates and a much more accurate representation of the PES curvature, reducing the error when computing the VDOS.

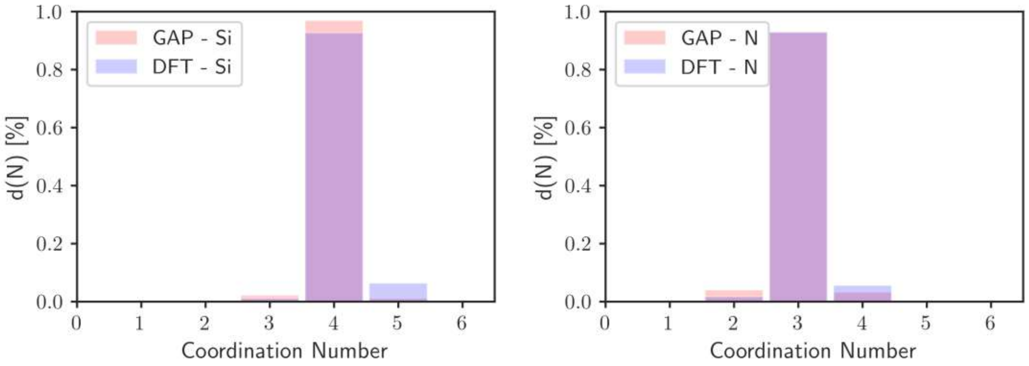

Figure 3.6: Coordination number statistics for Si (left) and N (right) in the atomic structures obtained with GAP (red) and DFT (blue).

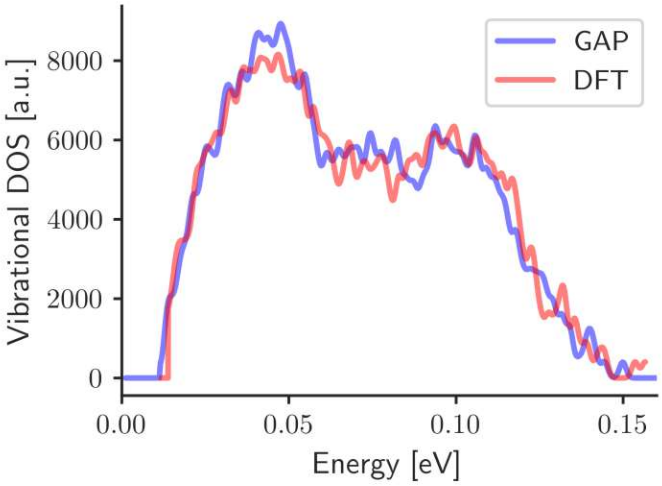

Figure 3.7: Comparison of the VDOSs calculated with GAP and DFT. There is a good match over the entire range of relevant energy values.

As mentioned above, the main advantage of ML interatomic potentials over *ab initio* methods when running MD is their more attractive accuracy-to-cost ratio. Maintaining a level of accuracy comparable to that of DFT, while being orders of magnitude computationally cheaper, allows for ML interatomic potentials to be used in MD, resulting in access to simulation times and system sizes previously out of range. In order to analyze the computational costs of the proposed ML interatomic potential, two different tests were conducted. In the first test, it was analyzed how the time needed to run an MD simulation scales with the number of atoms in the system. Therefore, a Si_3N_4 structure was considered and an MD simulation was re-run with 112, 378, 896, 1,750 and 3,024 atoms. The MD simulation was performed in LAMMPS using 20 cores, and it consisted of 10,000 time-steps at 1,000 K with a time-step of 1 fs. The results can be seen in the right panel of Figure 3.8, where a nearly linear scale of the simulation time and the number of atoms can be observed.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

The second test aimed at analyzing how parallelization can be used to reduce the computational times needed to run MD simulations with the proposed ML interatomic potential. Therefore, LAMMPS was used to re-run an MD simulation, which consisted of a Si_3N_4 system with 3,024 atoms, evolving over 10,000 time-steps at 1,000 K, with a time-step of 1 fs. The simulation was repeated with 2, 4, 8, 16 and 32 cores and the results are presented in the left panel of Figure 3.8, where a nearly linear correlation between performance and number of cores can be seen. By running an MD simulation with DFT for the same system using 4 cores, the performance was estimated to be about 0.0011 [K steps/hour]. The GAP yields a performance of 1.28 [K steps/hour] under the same circumstances, therefore making it about 3-4 orders of magnitude faster than DFT for this particular scenario. However, the computational cost of DFT scales cubically with the number of atoms in the system, as shown in

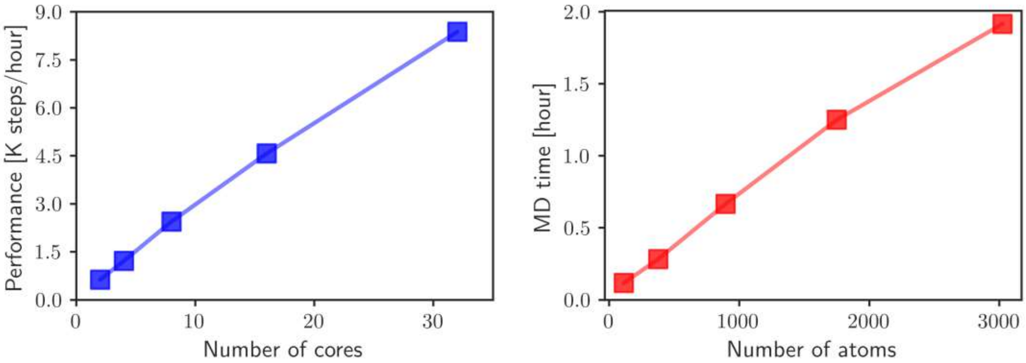

Figure 3.8: Analysis of the computational cost of the proposed ML interatomic potential. Left: Running time vs. number of atoms in the system, for a simulation of 10,000 time-steps. Right: Performance vs. number of cores for a system size of 3,024 atoms.

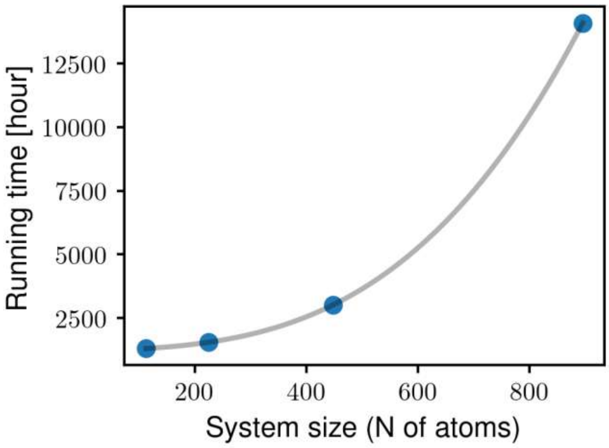

Figure 3.9: Scaling of DFT computational cost versus system size (number of atoms) for a 10,000 time-steps MD at 300 K using 20 CPU cores. The fitting curve indicates a cubic scaling of the DFT computational cost with the system size.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Figure 3.9. Therefore, the performance gap between DFT and the ML interatomic potential grows with the system size.

In summary, an ML potential for Si_3N_4 was presented, developed using the GAP method. The training was done using an active-learning algorithm, consisting of iterative re-trainings. Following this approach, a GAP was used to run melt-and-quench MD simulations. The resulting atomic trajectory was sub-sampled and the energies, forces and/or stress tensors of the final atomic configurations re-calculated using DFT. These data were added to the training dataset, the GAP was re-trained and the process was repeated until the desired accuracy was reached. By applying this method to a variety of different atomic systems, the PES of interest was thoroughly sampled. Compared to energies from DFT calculations, the resulting ML potential yields an MAE of approximately 8 meV/atom. Moreover, when used to model amorphous Si_3N_4, the GAP yields a remarkable agreement to DFT results and experimental data, while being approximately 3-4 orders of magnitude faster than DFT. This combination of highly accurate results and remarkably low computational costs enable for much larger and more realistic structures, allowing MD simulations over much larger atomic system sizes and time scales.

3.2 GAP for Silicon-Germanium (SiGe)

The results and figures presented in this section have been partially presented at the TouCAM 2022 conference [115] and the EMRS 2023 conference [116]. This work is the result of a collaboration between Technische Universität Wien (TUW) and Laboratoire d'analyse et d'architecture des systèmes (LAAS)-Centre national de la recherche scientifique (CNRS) [26], and a product of my research stay in Toulouse from September to November 2022. The main motivation for this work was to develop an accurate and efficient interatomic potential to study the effects of neutron irradiation in SiGe by running MD simulations.

Microelectronic components can be subjected to displacement damage (DD) effects in radiative environments, such as nuclear plants or space [117]. DD can be described as a disruption induced in an atomic structure by its interaction with an incident high energy particle. When an energetic particle (e.g., a neutron) collides with an atom, the former transfers part of its kinetic energy to the latter, causing a displacement. The higher this energy transfer is, the further the atom will be displaced. An atom displaced by an incident particle is known as a primary knock-on atom (PKA). If the energy transfer is sufficiently high, the PKA will hit other atoms in the structure, starting a collision cascade, which can cause permanent damage.

Empirical tests are performed to test the effects of radiative environments on microelectronic components, consisting in exposing the component to a controlled source of radiation. However, gaining insight into atomistic processes via this path is difficult experimentally and MD simulations are often used as a complementary tool [118], as shown in Figure 3.10.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek, WIEN Your knowledge hub

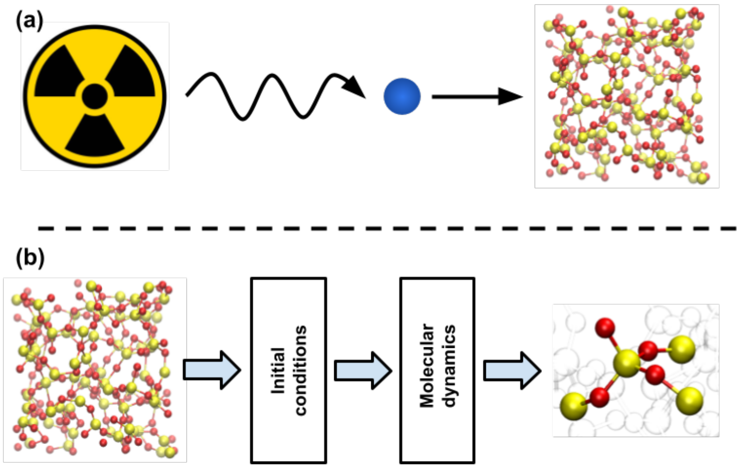

Figure 3.10: Testing materials for exposition to neutron irradiation. (a) Empirical test: exposing the selected material to neutron irradiation and studying its effects on it. (b) Simulations: giving a high initial momentum to one of the atoms of the atomic structure of interest, running an MD simulation and studying the resulting changes in the atomic structure.

These MD simulations are performed starting from the atomic system being studied, at room temperature. For neutron irradiation, the collision of an energetic neutron is simulated by transferring momentum to one of the atoms, causing the beginning of a collision-cascade. For the material of interest, SiGe, these simulations can be performed using *ab initio* methods, such as DFT, or using empirical potentials, such as the Stillinger-Weber force field [119]. However, the first option is computationally expensive, limiting its application to small atomic system sizes and simulation times and, therefore, rendering it impracticable for most cases. The second option is computationally more efficient, but its accuracy is in question, since important physical processes are not properly captured, such as the electronic dissipation effects during the collisions between the incident particles and the atoms in the structure. It is in this context that an ML interatomic potential becomes a suitable alternative.

Since the motivation of this work was to simulate and study the effects of neutron irradiation not only on crystal, but also in amorphous SiGe, the first step consisted in training and testing the GAP for crystal, liquid and amorphous SiGe structures. In order to create an initial dataset of relevant atomic configurations, the melt-and-quench technique was used [120] with an MD algorithm as implemented in QuantumATK [110]. The computationally inexpensive Stillinger-Weber empirical potential [119] was used to melt an initially defect-free crystal SiGe structure at a temperature of 5,000 K, therefore forming a liquid, which was thereafter quenched back to room temperature, creating an amorphous SiGe structure at the end of the process. Since SiGe can contain different ratios of Si and Ge, this process was repeated for atomic systems with 0%, 10%, 25%, 50%, 75%, 90% and 100% Ge. The results of these melt-and-quench MD simulations are shown in Figure 3.11.

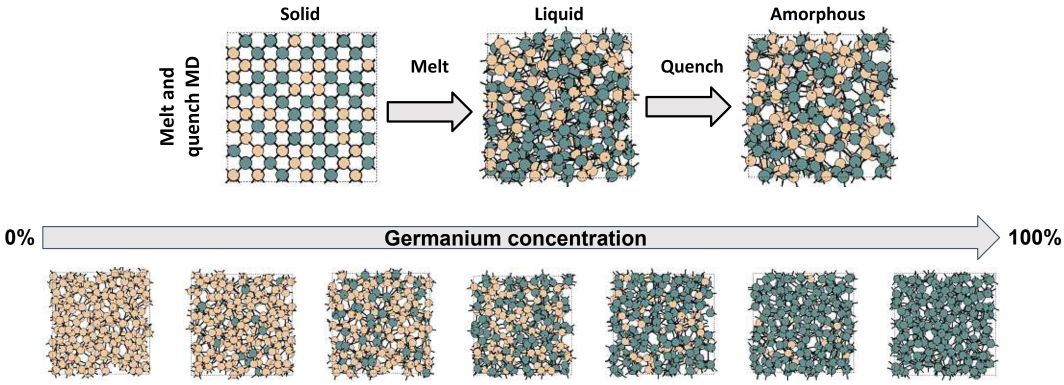

Figure 3.11: Creation of a dataset by running melt-and-quench MD simulations. The process is repeated for different concentrations of Ge.

The resulting atomic trajectory from each MD simulation was sub-sampled to 1,000 atomic configurations, for which the energies and forces were re-calculated with DFT, using the PBE functional [112] in the CP2K software package [111]. The atomic configurations for which the DFT calculations did not converge (about 5%) were discarded. The remaining atomic configurations, together with their DFT energies and forces were collected into the initial dataset. A considerable amount of computational time was saved, since there was no need to run *ab initio* MD simulations.

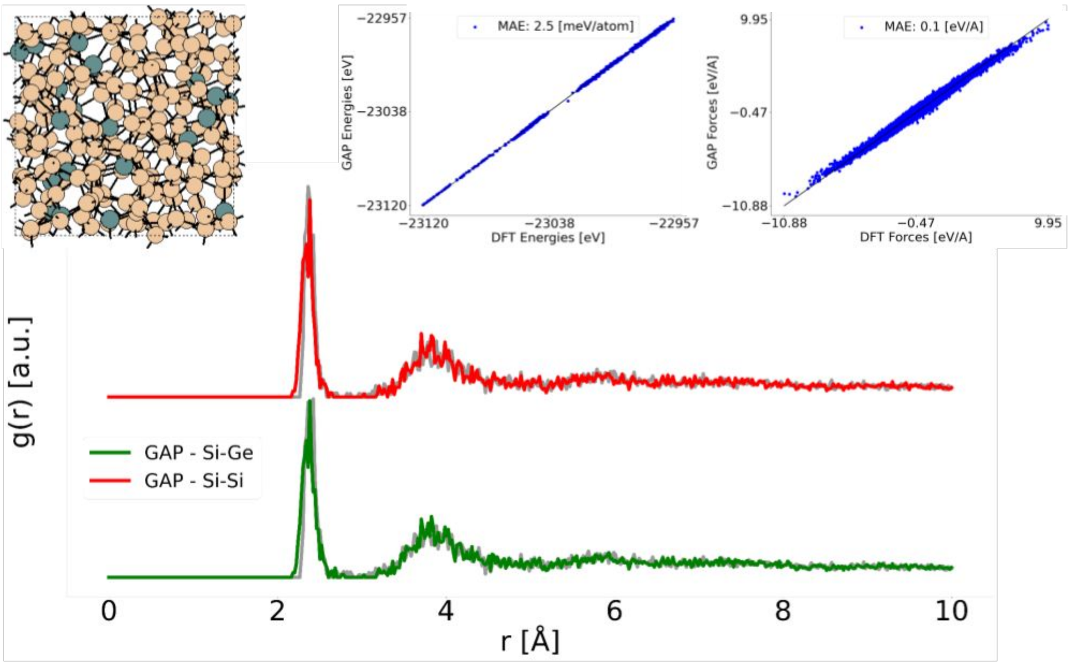

Figure 3.12: Validation of the GAP against DFT for energy/forces calculations and structural properties of low Ge concentration structures. Color: GAP RDFs, grey: DFT RDFs for Si-Si and Si-Ge bonds.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

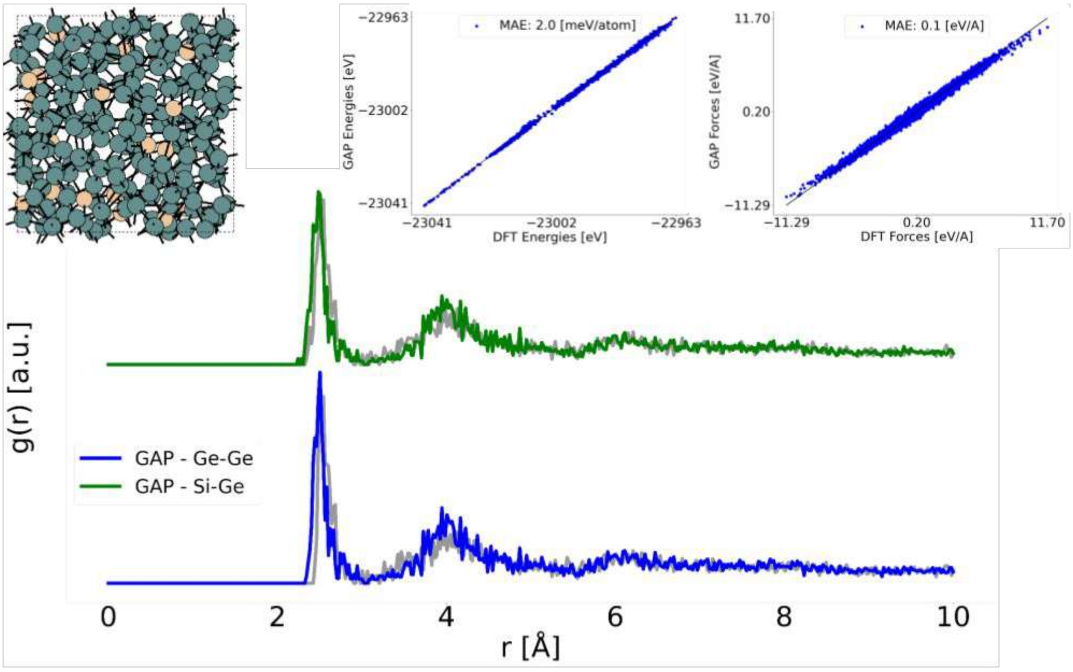

Figure 3.13: Validating the GAP against DFT for energy/forces calculations and structural properties of high Ge concentration structures. Color: GAP RDFs, gray: DFT RDFs for Ge-Ge and Si-Ge bonds.

Once the initial dataset was built, 200 evenly-spaced atomic configurations from each trajectory corresponding to different concentrations of Ge were used to build the training dataset. In addition, Si-Si, Si-Ge and Ge-Ge dimers, as well as Si and Ge single atom configurations were added to the training dataset. On the other hand, 1,000 of the remaining atomic configurations were randomly selected to build the testing dataset. No atomic configuration in the training dataset was used in the testing dataset.

After the training and testing datasets were built, a GAP was trained on the training dataset and thereafter tested by comparing it to DFT for the energy and forces calculations of the atomic configurations in the testing dataset. The top panels of Figure 3.12 and Figure 3.13 show the results for the atomic configurations with 10% and 90% Ge, respectively.

The next test consisted in using the GAP to run melt-and-quench MD simulations in LAMMPS for each of the different Ge concentrations. The partial RDF of the resulting amorphous SiGe configurations were compared to DFT references. The results for 10% and 90% Ge are shown in the bottom panel of Figure 3.12 and Figure 3.13, respectively.

The forces computed using the GAP yielded a consistent MAE of 0.1 eV/Å for all Ge concentrations, for the energies a MAE between 2 meV/atom and 2.5 meV/atom. On the other hand, the RDFs of the amorphous SiGe structures created with the GAP correlate well with the RDFs of DFT-relaxed reference structures. The last step of the testing process was to use the GAP to run neutron irradiation MD simulations

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

for different energy levels and system sizes. The higher the energy of the incident particle, the larger the atomic system must be to accurately simulate the cascade effect. Figure 3.14 shows examples of neutron irradiation simulations ran with GAP for low and high incident particle energies. It can be seen from these results how, for this given system size, the effects of low energy incident neutrons can be simulated. In contrast, for high energy incident neutrons, some atoms leave the system, thereby indicating that a larger cell is needed for this simulation.

The calculations carried out with the resulting GAP were between 3 and 4 orders of magnitude faster than DFT and with a linear scaling in the computational costs, as compared to the cubic scaling of DFT. This allowed simulating atomic systems out of reach for DFT, as in the example shown in Figure 3.15.

The results of this work include a fully functional general-purpose ML interatomic potential for SiGe, trained and tested for crystal, liquid and amorphous phases, as well as for a wide variety of different Ge concentrations. The results showed a remarkably high level of accuracy when compared against DFT for energy and forces calculations. Moreover, the high correlation to the DFT calculations indicated no systematic errors. When used to model amorphous SiGe by using the melt-and-quench technique, the RDFs of the resulting atomic structures yielded a perfect correlation to those of the DFT references. Finally, the preliminary results on using this ML interatomic potential to run neutron irradiation MD simulations indicated a drastic reduction in the computational costs when compared to DFT, which thereby allows to simulate system sizes previously out of reach by orders of magnitude. As well as that, a high

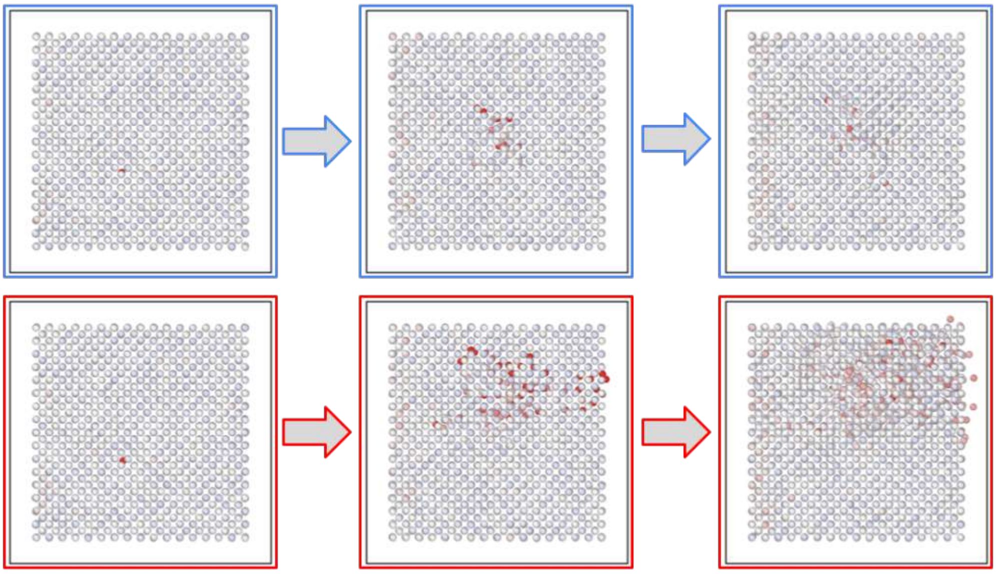

Figure 3.14: Examples of the results of running an MD simulation for neutron irradiation on SiGe for a low (top) and a high (bottom) energy incident particle. The stronger the color of an atom, the higher its velocity. The snaps are taken at 0 fs (left), 25 fs (center) and 50 fs (right) from the moment the incident neutron collides with the PKA.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

accuracy and no trace of clustering nor other unphysical phenomena was found. An alternative quantitative validation of the resulting SiGe structures is currently a work in progress. Their sheer size make them impossible to geometry-relax with DFT and an alternative method of validation must be found.

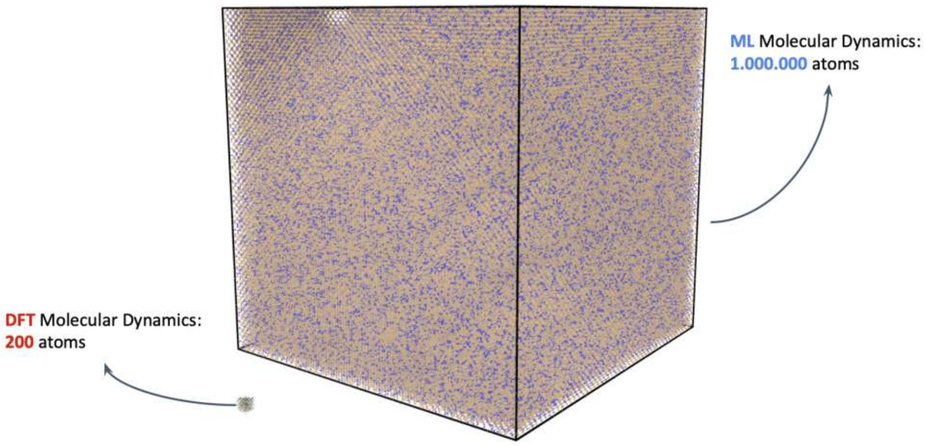

Figure 3.15: Scaling of the system size in neutron irradiation MD simulations for SiGe allowed by GAP.

3.3 GAP for the Oxidation of Silicon (Si)

The results and figures presented in this section have been partially presented at SiO_2*-2023 conference [121].*

Studying the oxidation of silicon is of utmost importance due to its significant implications in various fields, ranging from semiconductor technology to nanoelectronics and beyond. The understanding and exploration of silicon oxidation provides valuable insights and pave the way for advancements in numerous areas.

Silicon is the backbone of the semiconductor industry, serving as the primary material for fabricating electronic devices such as integrated circuits. The process of silicon oxidation plays a vital role in the formation of a thin layer of silicon-dioxide, commonly known as the oxide layer [122]. This layer acts as an insulator, protecting the underlying silicon and enables the desired functionality of the electronic components. By studying the oxidation of silicon, researchers can optimize the process, enhancing device performance, and enabling the production of more advanced and efficient semiconductor devices.

Furthermore, the oxidation of silicon significantly impacts the reliability and long-term stability of semiconductor devices. The formation of the oxide layer is critical for isolating different components, preventing leakage currents, and ensuring proper functioning. However, defects and imperfections in the oxide layer can affect device characteristics, leading to performance degradation and reduced reliability [123].

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

The oxidation process of silicon can be studied by running MD simulations. However, developing an ML interatomic potential for this particular application is far from trivial. The main challenge lies on the wide variety of interatomic interactions for which the potential must be trained:

- Interactions in the molecular oxygen (O_2) gas (O-O).
- Initial interactions between the O_2 gas and the silicon surface (Si-O).
- Interactions between the oxidized silicon surface and the O_2 gas (SiO_2-O_2).
- Interactions between the oxidized silicon layers on top and the non-oxidized silicon layers at the bottom (SiO_2-Si).
- Interactions between the bottom silicon layer of the slab and the H atoms used to passivate it (Si-H).

To tackle these prolems, an active-learning solution was developed, schematized in Figure 3.16. Active learning is a pivotal paradigm in the realm of developing ML interatomic potentials, playing a crucial role in efficiently harnessing computational resources to enhance model accuracy and generalization. In this context, active learning emerges as a strategic approach to intelligently select informative data points for model training, thereby optimizing the predictive performance of the potential.

Traditional approaches in ML often involve training models on datasets build by randomly selecting datapoints, leading to resource-intensive processes that may demand substantial computational power and time. In contrast, active learning dynamically adapts the training set by iteratively selecting instances that are deemed

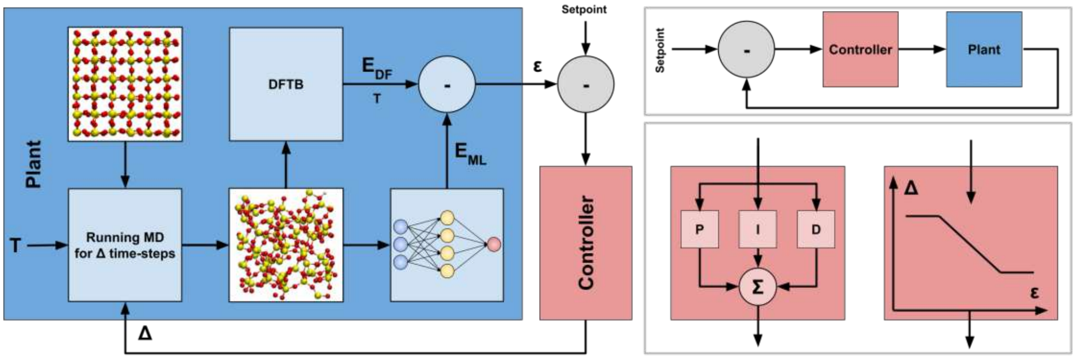

Figure 3.16: Schematic of the controlled dynamical sampling technique. The input of the plant is Δ, the number of time-steps of the MD simulation, while the output is ϵ, the errors between DFTB and GAP. An MD simulation is performed, starting from the current state of the atomic system of interest and for Δ time-steps, the DFTB and GAP energy/forces calculations for this snapshot are compared and their error, ϵ, is computed. By using a controller, the sampling frequency can be dynamically adapted depending on this error. Possible controllers are shown in the bottom-right panel, while a simplified schematic of this process is shown in the upper-right panel.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

most valuable for refining the model. This adaptability is particularly advantageous when dealing with interatomic potentials, where the complexity of atomic interactions demands precise calibration for accurate predictions.

In the pursuit of constructing robust interatomic potentials, active learning strategically identifies data points that pose the greatest uncertainty or contribute the most to model error. By iteratively selecting these influential instances, the model learns to refine its predictions in regions of the input space where it initially struggled.

The particular active-learning technique developed in this work, which will be referred to as *controlled dynamical sampling* is based on the active learning technique presented in Section 3.1 and the work done by Deringer *et al.* [57]. Controlled dynamical sampling relies on combining control theory [124] and MD simulations to dynamically sample the PES. Its operation is based on the following concept: An MD simulation starting from an initial atomic configuration X_i, and running for Δ time-steps at a temperature T, ends in a final atomic configuration X_f. If T is sufficiently high, the higher Δ is, the more different X_f will be from X_i and, therefore, the further away they will be in the PES. This relationship between Δ and displacement on the PES is the core of the controlled dynamical sampling technique.

A system or, using control theory jargon, a *plant*, can be defined with a single input (Δ) and a single output (ϵ). This system is defined by the following process: running an MD simulation at a fixed temperature T for Δ time-steps from the current state of the atomic system, X_A, and to produce a new atomic configuration, X_B. Thereafter, the energy/forces of X_B are computed with the selected theory level, e.g., Density Functional based Tight Binding (DFTB), and with the ML interatomic potential being developed. The output of the system, ϵ, is then defined as the difference between both. Naturally, ϵ represents the accuracy of the ML interatomic potential for that point on the PES.

Intuitively, a high ϵ means that the accuracy for that point in the PES is low and, therefore, the ML interatomic potential is unreliable and needs more training data in that area. On the contrary, a low value of ϵ indicates a successful training of the ML interatomic potential for that particular area of the PES and that no extra data is needed.

As shown schematically in Figure 3.16, the proposed technique works by using a controller to determine Δ for the next iteration, based on ϵ from the last iteration. The higher ϵ, the less accurate the ML interatomic potential is and the more data in that area of the PES is needed, therefore, the lower the next Δ will be. If the energy/forces ϵ is higher than an empirically defined threshold (e.g., an energy error equal or higher than 10 meV/atom), the given atomic configuration is added to the training dataset. The ML interatomic potential is re-trained after a certain number of new atomic configurations is collected in the training dataset (e.g., 50).

Any controller available in the control theory literature can be used in this technique, such as the popular proportional-integral-derivative (PID) [124], or a simple and intuitive proportional (P) controller. The results shown next are based on using a P controller, which regulated the value of Δ from 100 to 10,000 time-steps.

The final training dataset for the GAP developed in this work consisted of a total of 992 atomic configurations, including single atom configurations, dimers, bulk

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

silicon, bulk amorphous silicon-dioxide and silicon surfaces with different levels of oxidation. The descriptors used were two-body (cutoff = 4 Å) and SOAP ($n_{\max} = 8$, $l_{\max} = 8$, cutoff = 4 Å). The validation of the GAP was done by running MD simulating the oxidation of silicon slab surfaces exposed to O_2 gas. The silicon slabs were passivated and kept in place by a bottom layer of H atoms, for which the coordinates were fixed throughout the simulation. The surface of the silicon slab was oxidized by interaction with the O_2 gas above it and the simulation was re-started periodically in order to replace the O_2 molecules which bonded with the silicon slab. A depiction of the oxidation process produced by running an MD simulation with the proposed GAP in LAMMPS can be seen in Figure 3.17 and Figure 3.18. The quantitative validation of the resulting structures was done by comparing relevant geometrical properties of the resulting atomic structure to an equivalent structure generated with DFTB. The geometrical parameters used were: (1) mean Si-O bond-length, (2) mean O-Si-O bond-angle and (3) mean Si-O-Si bond-angle. The results are presented in Table 3.4.

The extension of this potential to simulating the oxidation of silicon nanowires is work-in-progress and out of the scope of the present thesis. Future work includes using the previously described controlled dynamical sampling technique to actively re-train

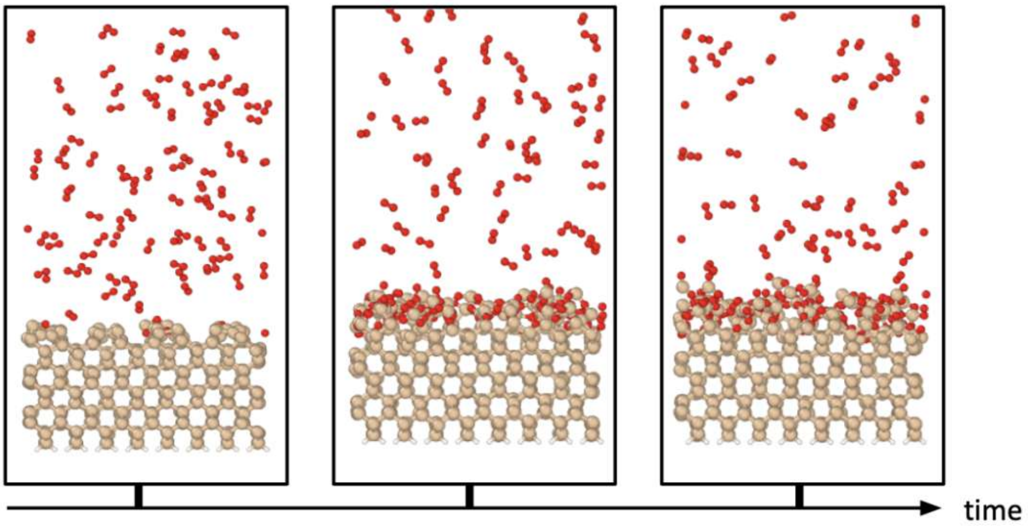

Figure 3.17: Oxidation process of a silicon slab exposed to O_2 in gas form. The MD simulation runs with GAP in LAMMPS through an iterative algorithm, in which the O_2 molecules are replaced periodically.

Table 3.4: Validation of the GAP, by comparing relevant geometrical parameters of the resulting silicon oxidized structures to equivalent DFTB simulations.

Parameter	GAP	DFTB
Si-O mean bond-length	1.66Å	1.66Å
mean O-Si-O bond-angle	108.9°	108.0°
mean Si-O-Si bond-angle	132.0°	132.5°

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

the GAP. Preliminary results on using the current version of the GAP to simulate the oxidation of silicon nanowires can be seen in Figure 3.19. The final objective would be developing an accurate and effective interatomic potential to model this process, motivated by the important role that these devices play in modern microelectronics. Silicon nanowires have emerged as a fascinating area of research and development in the field of nanotechnology. These structures, which consist of Si atoms arranged in a nanoscale wire-like configuration, possess unique properties that make them highly promising for a wide range of applications. Silicon nanowires would offer excellent compatibility with existing technologies, making them an attractive candidate for integration into electronic devices and systems.

The fabrication of silicon nanowires have seen significant advancements in recent years, allowing for precise control over their size, shape, and properties [125]. Their high aspect ratio, large surface-to-volume ratio, and tunable electronic properties have sparked interest in various scientific and technological domains. Silicon nanowires exhibit exceptional electrical, thermal, and mechanical characteristics, paving the way for potential breakthroughs in energy storage [126], sensing [127], and biomedical devices [128]. Therefore, their accurate simulation could potentially be of use across multiple industries.

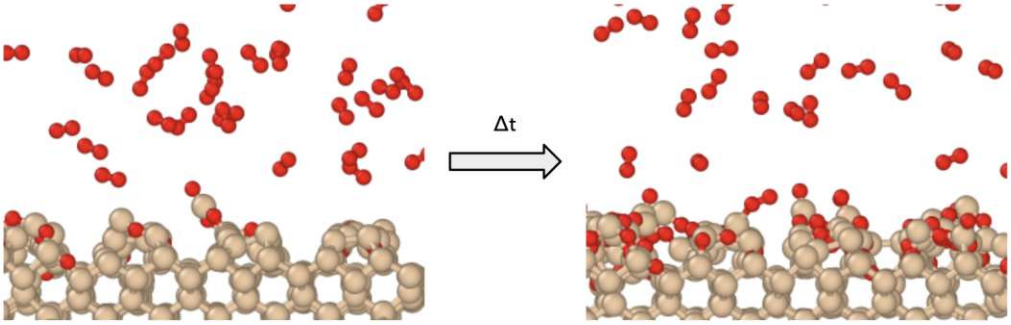

Figure 3.18: Zoom in to the oxidation process of a silicon slab exposed to O_2 gas.

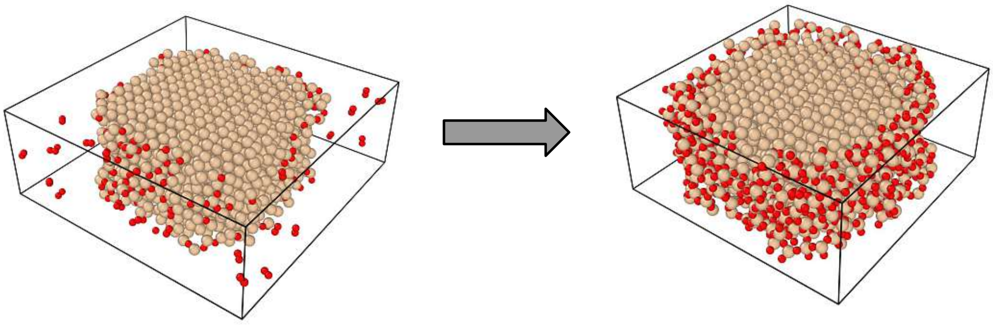

Figure 3.19: Results of simulating the oxidation process of a silicon nanowire, using the GAP developed to study the oxidation process of silicon surfaces exposed to O_2 gas. Credits: Franz Fehringer and Lukas Cvitkovich.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

3.4 GAP for Silicon-Dioxide (SiO_2)

The results and figures presented in this section have been partially published in Solid State Electronics [129].

The development of general-purpose ML interatomic potentials would be useful for a wide range of industries and scientific fields. However, the development of transferable ML interatomic potentials remains a topic of active research. In this context, *transferability* refers to the ability of an ML potential to retain its accuracy when employed in tasks it was not trained for (e.g., to accurately reproduce forces, despite only trained on energies). One of the main obstacles which weakens the transferability of ML potentials is overfitting, an undesired statistical phenomena in which ML models present a high accuracy on training data, but perform poorly in previously unseen testing data and, therefore, in real applications.

The interatomic potentials deployed in this work are implemented as instances of the well-established GAP [55]. The proposed workflow and a comparison to a traditional ML interatomic potential are depicted in Figure 3.20.

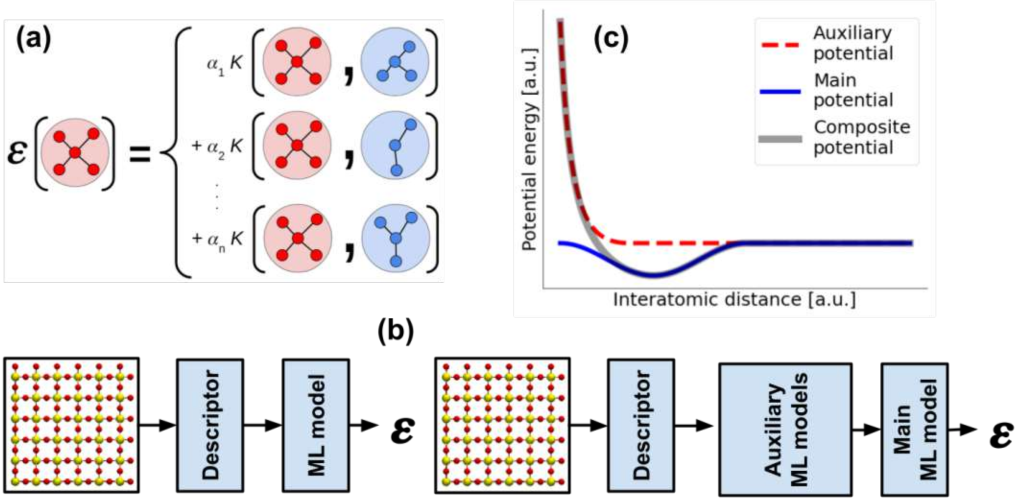

Figure 3.20: (a) Schematic of a GAP. (b) Workflows for a traditional potential built with this ML model (left) and the proposed composite interatomic potential (right). (c) Schematic of the proposed composite potential, augmenting the main potential with an auxiliary potential.

In this work, a systematic solution to mitigate the effects of overfitting was proposed, by augmenting a GAP with a set of independent repulsive-only potentials, in contrast to previous solutions found in the literature [65][130]. The proposed solution is demonstrated by building a potential for amorphous silicon dioxide (a-SiO_2). The potential was validated against DFT calculations and compared to the results obtained with a traditional ML potential.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

In a traditional ML interatomic potential, the total potential energy of a given atomic system is approximated by a sum of local energy contributions from every atom, computed by a single ML model [17],

$$E_{\text{total}} = \sum_{i}^{\text{Atoms}} E_i(\boldsymbol{d}_i), \tag{3.1}$$

where d_i is a local descriptor for the environment of the i^{th} atom and E_i is its local contribution to the total potential energy, computed by the ML model from the local descriptor. The proposal in this work was to build a potential which computes these local energies as a composite set of multiple components, i.e.,

$$E_i = \sum_{j}^{\text{Models}} E_i^j(\boldsymbol{d}_i^j), \tag{3.2}$$

where j runs over all components of the final potential. Each component can either be a non-parametric potential, or an ML model paired with different local descriptors, allowing for varying degrees of complexity.

In this work, an interatomic potential for a-SiO$_2$ was built composed of a complex main model responsible for giving accurate results for the atomic systems of interest and a simpler auxiliary potential, which represents the basic physics through pairwise short-range interactions, as shown in Figure 3.20(c),

$$E_i = E_i^{\text{main}}(\boldsymbol{d}_i^{\text{main}}) + E_i^{\text{aux}}(\boldsymbol{d}_i^{\text{aux}}). \tag{3.3}$$

The main ML potential is responsible for the intricate details of the PES of the atomic systems of interest. On the other hand, the auxiliary potential is repulsive-only, meaning that it only produces positive energies. The closer two atoms are to one another, the higher the computed energy will be. This potential matches the positive part of the pairwise dissociation curve, as obtained with DFT, by using a simple two-body descriptor with a short and smooth cut-off, as implemented in QUIP [108][131]. The sum of both energies results in the energy computed by the proposed composite ML interatomic potential.

The main potential employs the more sophisticated and highly-dimensional descriptor SOAP [54]. The training dataset for the main GAP was created by running MD, according to the melt-and-quench technique [120] within the LAMMPS engine [132]. A defect-free 216-atoms SiO$_2$ system was melted at 5,000 K and subsequently quenched to 300 K, using the classical force-field ReaxFF [133] with a time-step of 0.25 fs. The process is depicted in Figure 3.21. The resulting trajectory was sequentially sub-sampled to a training dataset of only 1,500 atomic configurations, for which the energies were calculated with DFT, using the PBE functional [112] in the CP2K software package [111].

Ideally, the training dataset must be comprehensive, while containing only atomic configurations relevant to the intended application for which the ML interatomic potential will be used. In other words, the atomic environments contained in it must represent a faithful sampling of the PES subset relevant for a particular application. An

TU Bibliothek WIEN Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

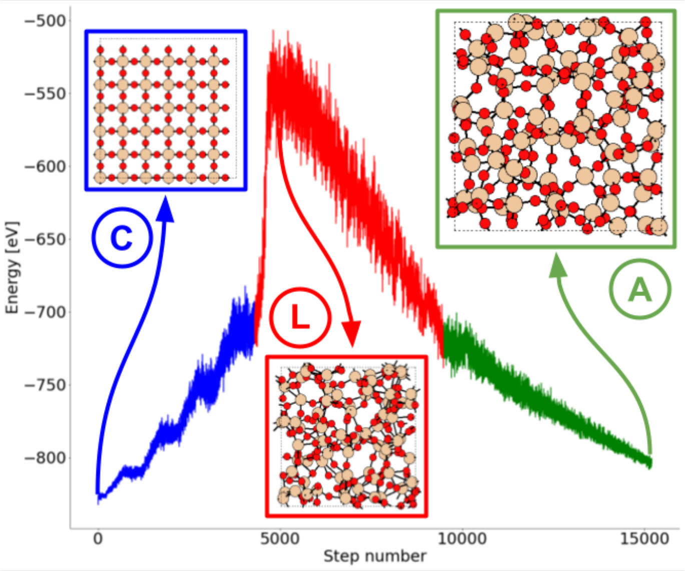

Figure 3.21: Initial dataset for training and testing the *main* ML potential, created using the melt-and-quench technique and the ReaxFF force field. The MD begins with a crystalline SiO_2 structure, which is melted up to 5,000 K and thereafter quenched back to 300 K. Once the dataset is created, the energies are recalculated on a subset using DFT. Blue: Crystalline phase (C). Red: Liquid phase (L). Green: Amorphous phase (A). The objective of the initial dataset is to provide a sampling of the PES of interest as comprehensive as possible.

under-sampling of the possible atomic environment space would undoubtedly result in poor ML prediction accuracy, while over-sampling results in an excessive computational cost. Furthermore, the sampling should not be biased towards any particular area of the PES, in order not to over-sample a sub-set of the possible atomic configurations space, while under-sampling the rest. It is for this reason that the training dataset is composed of a sub-set of the atomic configurations obtained by performing an MD simulation. The testing dataset is composed of 1,000 randomly selected atomic configurations from the same melt-and-quench MD trajectory. It is important to notice that no atomic configuration belonging to the training dataset was used in the testing dataset.

The main potential was trained on the residual between the DFT energies and the predictions of the auxiliary potential for the atomic configurations in the training dataset. The training was performed using the software package QUIP. The descriptor for the auxiliary potential was defined using smooth cut-offs: $r_{\text{cut}}^{\text{SiSi}} = 1.60$ Å, $r_{\text{cut}}^{\text{SiO}} = 1.10$ Å and $r_{\text{cut}}^{\text{OO}} = 0.80$ Å. The SOAP descriptor for the main GAP was constructed with $n_{\max} = 6$, $l_{\max} = 6$ and $r_{\text{cut}}^{\text{SOAP}} = 4.0$ Å. Therefore, the main potential acts on the entire atomic neighborhood, while the auxiliary potential acts only if two atoms are close to one another.

Following the traditional approach, a single GAP was trained on the dataset used to train the main ML model of the proposed composite potential. This potential was thereafter used to evaluate the performance of the proposed composite ML in-

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

teratomic potential. In the following tests, both potentials were used to perform a number of tasks and their results were compared to one another, taking DFT calculations as reference.

The first test for the composite and traditional ML potentials was to compute the potential energies of the 1,000 atomic configurations in the testing dataset. The results of each ML potential were compared to DFT calculations and shown in Figure 3.22. The proposed composite ML potential yielded an MAE of 4.7 meV/atom, therefore rendering it slightly more accurate than the traditional potential, which yielded a MAE of 4.8 meV/atom. In practical terms, this negligible difference in accuracy between both potentials means that they should produce near identical results when used in real applications. However, the following tests will demonstrate that this first impression is wrong and that the simple auxiliary potential plays an important role in the overall ML potential.

The second test employed both ML potentials to run the exact same melt-and-quench MD simulation, begining with an initially defect-free crystalline structure, which is melted at 5,000 K for a period of 30,000 time-steps (step-size = 1 fs). Once melted, the atomic system remains at a temperature of 5,000 K for another 20,000 time-steps, and it is then quenched back to 300 K within a period of 100,000 time-steps. The final result of this melt-and-quench MD simulation is an a-SiO_2 structure, as shown in Figure 3.23. The structures produced by the traditional model suffer from unphysical atomic cluster formations, as is apparent from the RDFs. In contrast, the composite potential shows excellent agreement with the DFT reference structure. The reason for this is that training datasets built by running MD rarely contain short-range interactions, as they are high in energy and therefore unlikely to be present in the MD trajectory. This prevents the traditional ML potential from

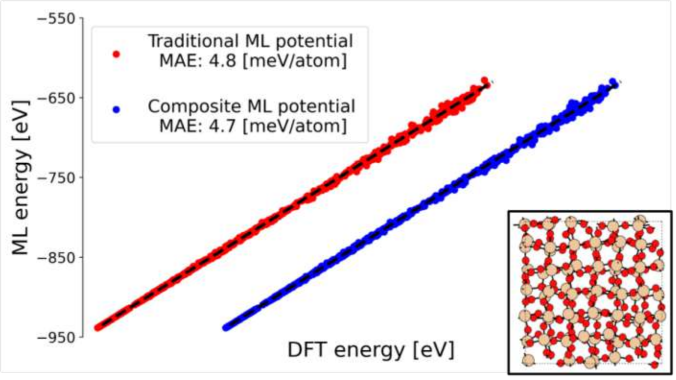

Figure 3.22: Testing the potential energy calculations accuracy for the traditional (red) and the composite (blue) ML potentials against DFT. The dashed black line represents the ideal results, in which all energy predictions would be equal to the DFT calculations. The closer the dots are to this line, the better the results are. This reference makes the X-axis immaterial. The results show a high accuracy for both potentials and virtually no difference between them. An example of the atomic structures used for testing is shown in the bottom-right sub-panel.

learning that short interatomic distances correspond to high energies, resulting in the unphysical clustering seen in the results. Since this information is explicitly included in the auxiliary potential, the proposed approach is much more robust while also retaining a high accuracy, with only little computational overhead (roughly 5%).

In order for an ML interatomic potential to be useful in practical applications, it must be scalable. This means that it must retain its accuracy for atomic systems noticeably larger than those used to train it. Ideally, the ML interatomic potential would also be able to yield accurate results when applied to tasks for which it was not specifically trained on, thus indicating a successful and comprehensive training process. As previously explained, accurate results for testing situations similar to those found in the training, paired with inaccurate results for different and new situations, would be a strong sign of overfitting.

To validate the transferability of the composite ML interatomic potential, the VDOS for one of the resulting a-SiO_2 structures was computed and compared against DFT results, as shown in Figure 3.24(a). The VDOS was determined by employing a finite-difference scheme in order to expand the system's PES to second order around the minimum configuration [134]. An a-SiO_2 system was chosen resulting from an MD run, and it was further relaxed using DFT, as well as the composite potential. The finite difference calculations were carried out by displacing a single atom by 0.01 Å along each coordinate axis. The resulting force constant was then diagonalized to obtain the system's normal modes and the corresponding phonon energies. As can be seen, there is excellent agreement between DFT and the proposed ML potential in predicting the VDOS for low-energy phonons. The fact that the proposed ML potential is able to produce accurate results for a task for which it was not trained is a strong indication of a successful training process. The VDOSs differ for higher energies. However, if higher accuracy was needed in that energy range, this could be improved by including forces in the training process. Including forces in the training process would provide the ML potential with the explicit value of the first derivative of the PES with respect to the atomic coordinates and a much more accurate representation of the PES curvature, reducing the errors when computing the VDOSs.

The training and testing datasets for this application were composed of atomic configurations extracted from the same MD run, meaning that they all share certain distinctive properties, such as the number of atoms and the size of the cell in which they are contained. Moreover, as all frames belong to the same trajectory, a certain degree of correlation or similarity is therefore expected between the atomic environments of the structures in both datasets. It is for these reasons that, to validate the scalability of the composite ML potential, a new test was needed. The proposed composite ML potential was employed to perform MD simulations on an atomic system 10-times the size of the atomic structures found in the training dataset. The results are shown in Figure 3.24(b), where no signs of unphysical clustering were found, therefore hinting at a successful training process and indicating the scalability of the proposed ML potential. These melt-and-quench MDs were performed using a relatively high quench rate, which allowed them to quench in barely 100,000 time-steps, a feat out of reach for several empirical potentials. This, paired with the small dataset needed for training, makes the proposed ML potential highly efficient.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

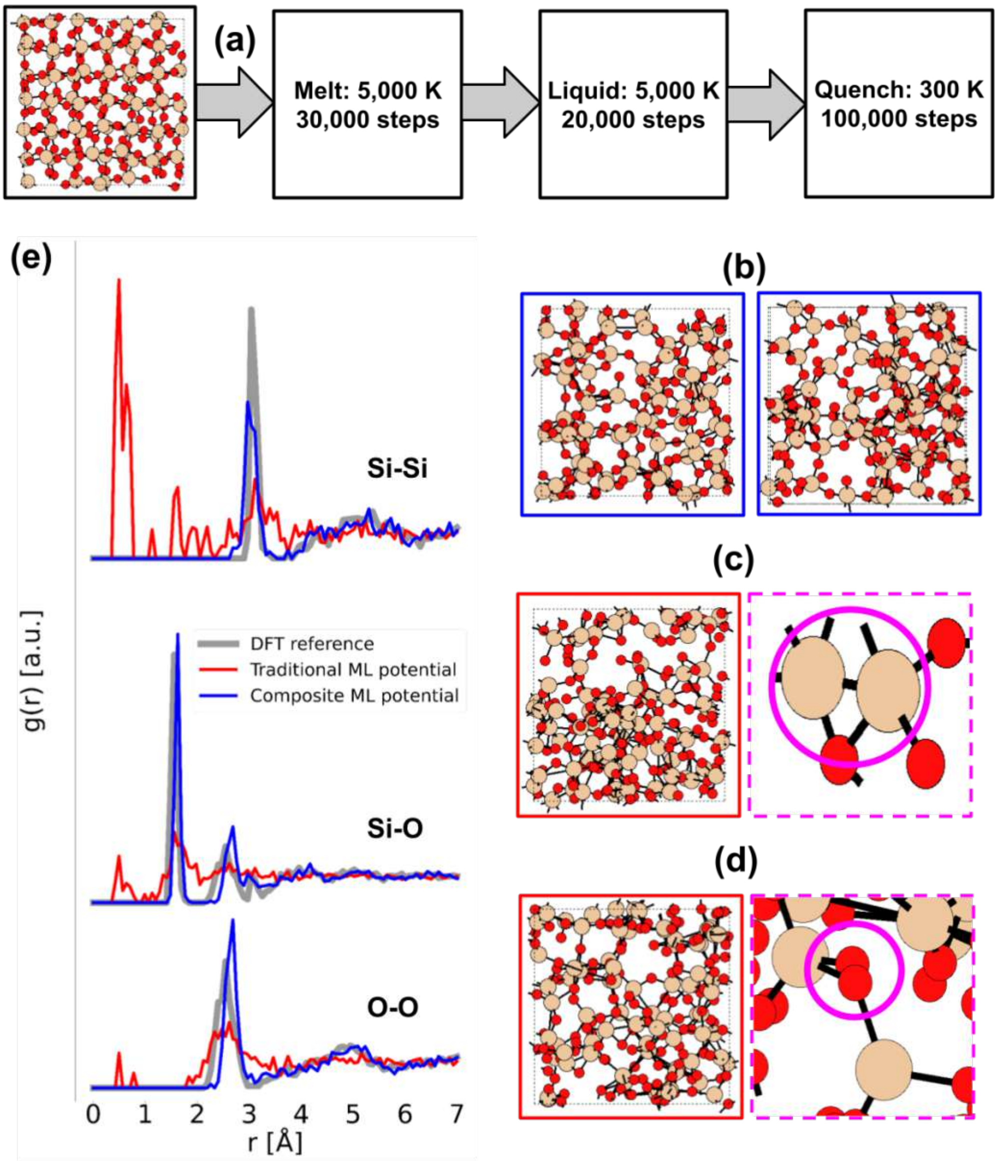

Figure 3.23: Panel comparing results of the proposed framework with those of a traditional ML interatomic potential, built as a single GAP. Both potentials were used to run the same MD simulation, as specified in (a). Two of the resulting structures obtained from the proposed approach are presented in (b). Resulting structures from using the traditional approach are presented in (c) and (d), together with examples for unphysical behavior found in them. The RDFs for one of the structures built with the new approach and one built with the traditional ML potential are presented in (e), together with the DFT reference. As can be seen, the composite potential performs much better when compared against the DFT reference.

ML models are a powerful tool for developing highly accurate and computationally efficient interatomic potentials. However, when the underlying physical mechanisms are not explicitly taken into account, their transferability is in question. In

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

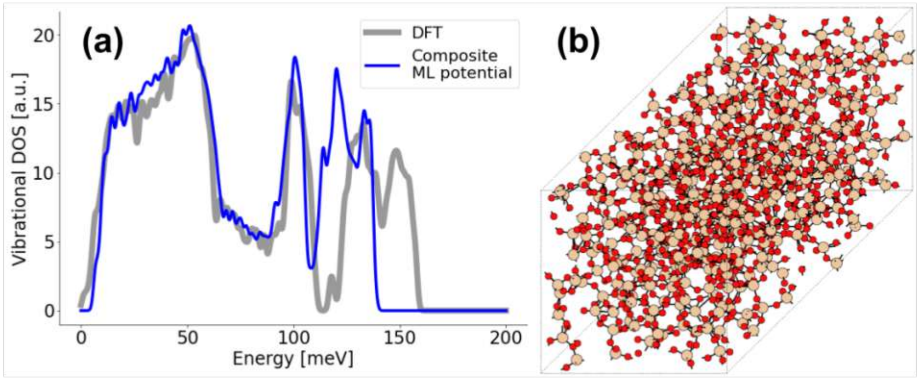

Figure 3.24: Transferability tests for the proposed composite ML potential. (a) Computed VDOS in comparison to DFT results. The good match is remarkable, considering that the potential was not trained on forces. (b) Using the potential to create an a-SiO_2 structure 10 times the size of those used in the training dataset. No unphysical clustering is present in the model.

the proposed approach, an ML potential is augmented by a set of simpler potentials. Results show that this significantly reduces the unphysical behavior when facing unexplored atomic environments, at a small increase in the computational costs. Future improvements to this work would include the forces to the training process, as opposed to the use of only energies, and the use of a wider range and variety of ML models combined with different descriptors.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

Chapter 4

Machine Learning Alternatives to Interatomic Potentials

One of the most fertile applications of ML techniques and algorithms in MD simulations and atomic calculations is to build surrogate models of the PES of interest, therefore constituting an ML interatomic potential. As shown in the previous sections, in general terms, this has the advantage of being computationally more efficient than *ab initio* methods and more accurate than empirical potentials. Nonetheless, an alternative solution to this framework is the employment of ML models to directly compute the property of interest from a given atomic system. In this chapter, the development of two of such alternative ML solutions will be presented. In Section 4.1, the direct prediction of defect formation energies will be discussed, while Section 4.2 focuses on the direct prediction of defect atomic structures.

4.1 Prediction of Defect-Formation Energies

The results and figures presented in this section have been partially published at the SISPAD 2020 conference [135].

In this work, the possibility of using ML models to calculate the formation energies of hydroxyl E' center defects in a-SiO_2 structures [136] was studied. Studying the behavior of these defects, especially their formation during device processing, is of great importance for the development of modern microelectronics, as they are suspected to be responsible for bias temperature instability (BTI) and random telegraph noise (RTN) in MOS transistors [137][138][139]. Since the formation of hydroxyl E' centers depends on the availability of hydrogen, the concentration of these defects is not directly accessible in DFT. However, it could be derived from a kinetic Monte Carlo (KMC) process model [140], coupled with ML-based on-the-fly prediction of formation energies as a function of the local environment. The conceptual workflow proposed in this work, combining a descriptor and an ML algorithm, is shown in Fig. 4.1.

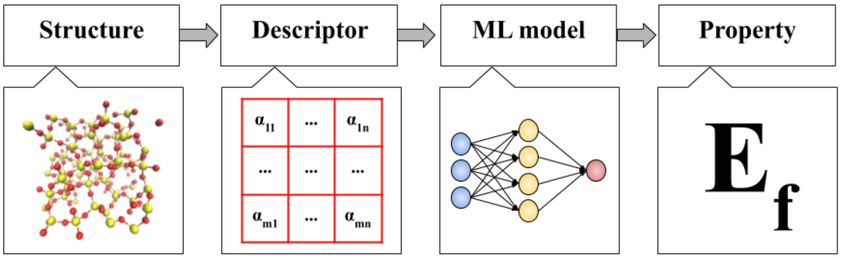
Structure
Descriptor
ML model
Property
α_{11} ... α_{1n}
... ...
α_{m1} ... α_{mn}
E_f

Figure 4.1: Workflow to predict a property from an atomic structure. The structure must be represented by a set of descriptors. Then, an ML model can be trained and subsequently used to predict the desired property in new atomic structures.

In this study, 16 different a-SiO_2 structures were prepared, each with a total of 216 atoms. These structures were created with LAMMPS [132] by using the ReaxFF [133] force field and the melt-and-quench technique, as described in [120]. Two examples of these structures are shown in Fig. 4.2. Within these structures, 1271 hydroxyl E' centers were created, as shown in Fig. 4.3. Their formation energies were extracted using DFT; the calculations were performed using the PBE functional [112] in the CP2K software package [111]. The resulting distribution of formation energies can be seen in Fig. 4.4.

Three popular ML models were used to predict the formation energy of the hydroxyl E' centers: A multi-layer perceptron, which is a type of neural network, kernel ridge regression (KRR) and decision tree (DT), as implemented in the scikit-learn package [141]. These models were combined with two sets of local descriptors common in the literature: SOAP [54] and atom-centered symmetry functions (ACSF) [142], implemented in the Python package DScribe [143].

The KRR is an ML technique that extends traditional linear regression to handle non-linear relationships in data. It falls under the umbrella of kernel methods, which utilize a mathematical construct known as a kernel function to implicitly map input data into higher-dimensional feature spaces. The primary motivation behind KRR is to capture complex patterns that may exist in the data.

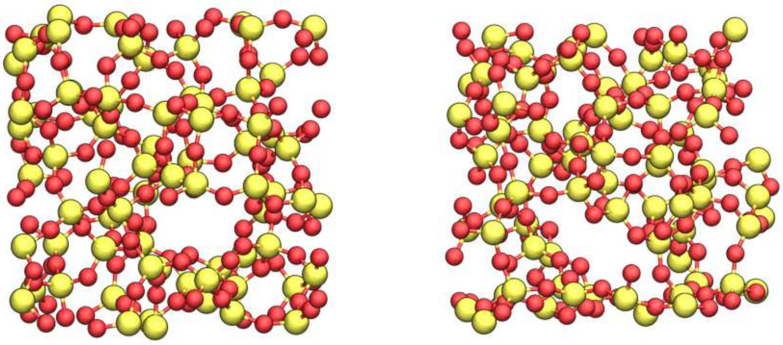

Figure 4.2: Example of two a-SiO_2 structures used in this work to train and test the ML models. Each structure contains 216 atoms and a total of 16 such structures are used in this work.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

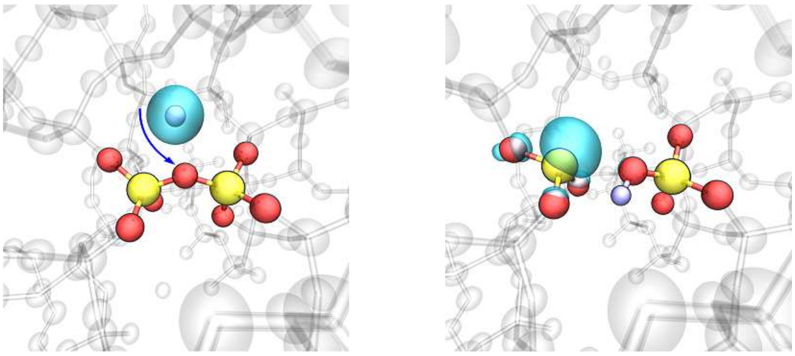

Figure 4.3: Hydrogen interaction with the a-SiO$_2$ matrix (left) can lead to the formation of a hydroxyl group, via the breakage of a strained Si-O bond, resulting in a hydroxyl E' center defect (right). The blue bubbles show the spin density associated with interstitial hydrogen and the hydroxyl E' center defect, respectively.

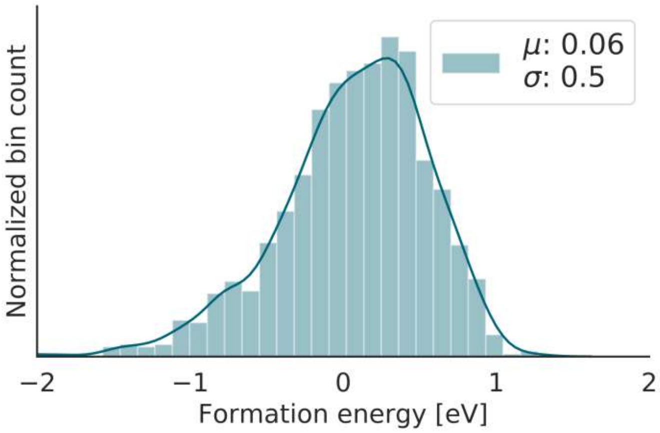

Figure 4.4: Formation energies of hydroxyl E' center defects in a-SiO$_2$, obtained from DFT calculations, together with their mean (μ) and standard deviation (σ), in eV. The broad distribution is due to the amorphous nature of the structures.

The KRR combines elements of ridge regression, which introduces a regularization term to prevent overfitting, with the expressive capabilities of kernel functions. The regularization term helps to stabilize the model and prevent it from fitting the noise in the data excessively, promoting better generalization to unseen examples. The kernel function, on the other hand, allows the KRR to handle non-linear relationships by implicitly mapping the input data into a higher-dimensional space, where linear relationships may become apparent. There are different types of kernels that can be used in this method, with the polynomial kernel being one of the most common,

$$K(x_1, x_2) = (\gamma \cdot x_1^T \cdot x_2 + c)^d, \tag{4.1}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

where γ is the kernel coefficient, c is the independent term and d is the degree of the polynomial. Training a KRR model is equivalent to finding the set of γ values that produce the best match between inputs and outputs in the given training dataset.

In contrast to KRR, a DT employs a tree-like structure. The nodes in the tree represent decision points based on specific features, and the leaves correspond to the predicted numerical values. During the training process, the algorithm recursively splits the dataset based on the features that result in the most effective reduction of the variance in the target variable. The decision at each node is typically determined by minimizing the mean squared error (MSE). The goal is to create partitions that result in homogenous subsets with similar target values. DTs in regression are beneficial for capturing complex non-linear relationships in the data and are particularly interpretable due to their tree structure. However, they are susceptible to overfitting, especially when the tree becomes too deep and starts fitting the noise in the training data.

The SOAP descriptor has been presented and discussed in Section 2.4.1.3. On the other hand, the ACSF descriptor builds a fingerprint representation of the local environment of a given atom composed of the output of multiple two- and three-body functions [142][143]. Three different two-body symmetry functions are used, G_i^{1,Z_1}, G_i^{2,Z_1} and G_i^{3,Z_1}, which are defined as

$$
\begin{aligned}
G_i^{1,Z1} &= \sum_{j}^{|Z_1|} f_c(R_{ij}) \\
G_i^{2,Z1} &= \sum_{j}^{|Z_1|} e^{-\eta(R_{ij}-R_\mathrm{s})^2} f_c(r_{ij}) \\
G_i^{3,Z1} &= \sum_{j}^{|Z_1|} \cos(\kappa R_{ij}) f_c(r_{ij}),
\end{aligned}
\tag{4.2}
$$

where the summation for j runs over all atoms with atomic number Z_1, η, R_s and κ are user-defined parameters, $R_{ij} = |R_i - R_j|$ and f_c is a smooth cutoff function, defined as

$$
f_\mathrm{c}(r) = \frac{1}{2}[\cos(\pi \frac{r}{r_\mathrm{cut}}) + 1], \tag{4.3}
$$

where r_cut is the cutoff radius. Additionally, the DScribe package implements the following three-body functions,

$$
\begin{aligned}
G_i^{4,Z_1,Z_2} &= 2^{1-\zeta} \sum_{j\neq i}^{|Z_1|} \sum_{k\neq i}^{|Z_2|} (1 + \lambda\cos(\theta))^\zeta e^{-\eta(R_{ij}^2+R_{ik}^2+R_{jk}^2)} f_\mathrm{c}(R_{ij}) f_\mathrm{c}(R_{ik}) f_\mathrm{c}(R_{jk}) \\
G_i^{5,Z_1,Z_2} &= 2^{1-\zeta} \sum_{j\neq i}^{|Z_1|} \sum_{k\neq i}^{|Z_2|} (1 + \lambda\cos(\theta))^\zeta e^{-\eta(R_{ij}^2+R_{ik}^2)} f_\mathrm{c}(R_{ij}) f_\mathrm{c}(R_{ik}),
\end{aligned}
\tag{4.4}
$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

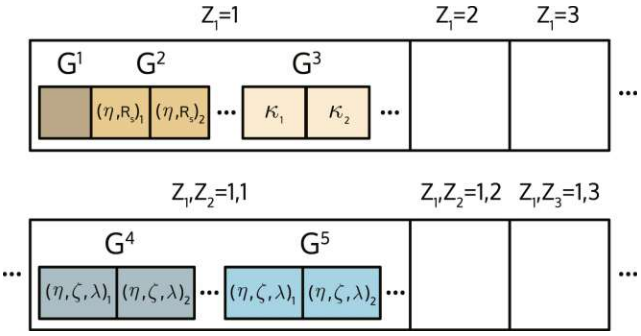

Figure 4.5: ACSF output vector: Initially, the output vector provides the values of the two-body symmetry functions G_1, G_2, and G_3 for each chemical species in the dataset. Subsequently, the vector includes the values of the three-body symmetry functions G_4 and G_5, listed for each unique combination of two chemical species. Reprinted from Computer Physics Communications 247, 106949 (2020). Original figure published under Creative Commons CC-BY 4.0 license. Credits: Lauri Himanen *et al.*

where the summation of j and k runs over all atoms with atomic numbers Z_1 or Z_2, respectively, η , λ and η are user-defined parameters, and θ is the angle between the three atoms (with the i^{th} atom in the center). A schematic of the ACSF output vector can be seen in Figure 4.5.

It should be clear from Figure 4.6 that there is an inverse correlation between the formation energy of a hydroxyl E' center defect and the length of the Si-O bond which has to be broken in order to form the defect. In other words, the formation energy is influenced by directly accessible geometric quantities (e.g., bond-lengths and bond-angles). Based on this observation, a simple geometry-based descriptor is proposed: bond-lengths and bond-angles (BLBA). This descriptor is built by using the bond-lengths and bond-angles of the atoms in the immediate vicinity of the possible defect site and arranging them in a vector in decreasing magnitude; then, this process is iterated for atoms further away from the possible defect site until a desired number of iterations is reached (3 iterations are considered in this implementation).

Although the BLBA descriptor cannot describe the local environment in the same level of detail of the SOAP and ACSF descriptors, it displays very valuable properties: since it contains only physical values which are relative to the possible defect site, the descriptor is invariant to spatial translations and rotations of the coordinate system. Moreover, since its components are ordered, it is also invariant with respect to permutation of indexes. Finally, it is highly compact, i.e. it contains sufficient information to be used for the prediction of the formation energies of hydroxyl E' center defects, while keeping its size and complexity to a minimum.

The performance of the BLBA descriptor will also be considered, as a demonstration of the potential of simple geometry-based descriptors. The 1271 hydroxyl E' center defects were randomly divided into a training and a testing dataset, with a ratio of 4:1. Every permutation of descriptor and ML model was trained with the training dataset and used to predict the values for the testing dataset. In every case, the MAE was calculated between the predictions and the targets.

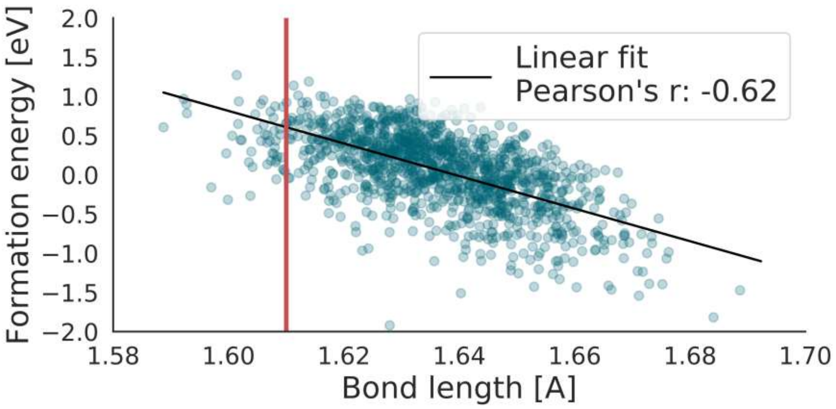

Figure 4.6: Correlation between formation energies and bond-lengths. The formation energy decreases for larger bond-lengths, so defects preferably form at strained Si-O bonds. The red line indicates the Si-O bond-length in alphaquartz.

The results of this study are summarized in Figure 4.7, where the prediction errors for the formation energies in the testing dataset are presented, together with its MAE, for every combination of descriptor (SOAP, ACSF and BLBA) and ML model (neural network, KRR and DT) studied in this work. Ultimately, the worst results were obtained with the ACSF descriptor in conjunction with a DT model, which presented a MAE of 0.39 eV. On the other hand, the best results were obtained with the SOAP descriptor in conjunction with a neural network model, which presented a MAE of 0.26 eV. This was expected, since SOAP describes the local environment in a higher level of detail. In this study, SOAP required 108 parameters to describe the local environment of the atom of interest, while ACSF required 36 and BLBA 17. However, it is important to stress that these numbers depend on the hyperparameters chosen for the SOAP and ACSF descriptors. A logical explanation for the results found is that the information contained in the extra parameters allows the SOAP descriptor to achieve more accurate predictions of formation energies of defects in the atomic structures.

Overall, BLBA performed similarly to SOAP and ACSF. However, it is important to note the following differences:

(I) For the specific configurations chosen in this study, BLBA was more compact than SOAP and ACSF, since it is able to properly describe the relevant local environment with a significantly lower amount of information.

(II) SOAP and ACSF have several additional parameters particularly designed to represent the chemical surroundings. This makes them more accurate, but it also means that the user requires a deeper understanding of the underlying atomic nature. Moreover, such necessity to adjust parameters which depend on the specific atomic structure makes them less accurate when this information is not available. On the contrary, BLBA does not require the adjustment of any

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

parameter (except for N, the number of iterations, but this adjustment does not require any previous knowledge of the atomic structure).

(III) Spherical harmonics decomposition is a familiar technique in all corners of atomistic calculations, producing relatively easy to interpret representations. However, every element of BLBA represents a physical property of the atomic structure (since each element of the descriptor represents a particular bond-length or bond-angle). Arguably, this makes it simpler than SOAP and ACSF and gives BLBA a higher interpretability.

(IV) The BLBA descriptor performed slightly better than the SOAP and ACSF descriptors when combined with a KRR or DT model. This could be explained given the fact that the BLBA descriptor is able to describe the local atomic environments with a lower amount of parameters. Therefore, it performs better with reduced datasets, by avoiding overfitting. However, more complex ML models, such as the neural network model, might make use of the extra information provided by the SOAP descriptor to increase the formation energy predictions accuracy.

The aim was to use a combination of a descriptor and an ML model to predict the formation energies of defects in a-SiO_2. Such predictions will be combined with a KMC method in future work, in order to assess the likelihood of defect formation under specific conditions. In other words, this work can be considered as a first step towards a more sophisticated ultimate goal: to predict reaction barriers by using a combination of a KMC method and an ML model.

It is also important to remember that the accuracy of the predictions is highly dependent on the atomic structures one considers as well as the electronic property one wants to extract or predict from them. Therefore, the results obtained in this work do not necessarily imply that the prediction accuracy of these descriptors would be unchanged if one applies them to different materials or different electronic properties from the one considered here. In case of applying a descriptor-based ML solution to predict a certain electronic property, it is crucial to consider different combinations of descriptors and ML models and to analyze the results obtained from them before choosing the best combination.

In order to allow a clear comparison between the BLBA descriptor and the well-established SOAP and ACSF descriptors, a correlation between the DFT results and the predictions made by using each of these descriptors is shown in Figure 4.8. In all cases, the descriptors were combined with a neural network model, since it was the ML model which showed the best results in the previous tests. The fact that all distributions are roughly centered around the identity line shows there is no considerable systematic error in the predictions and no strong overfitting of the ML model. The SOAP descriptor yields slightly better results than ACSF and BLBA, particularly for negative formation energies.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

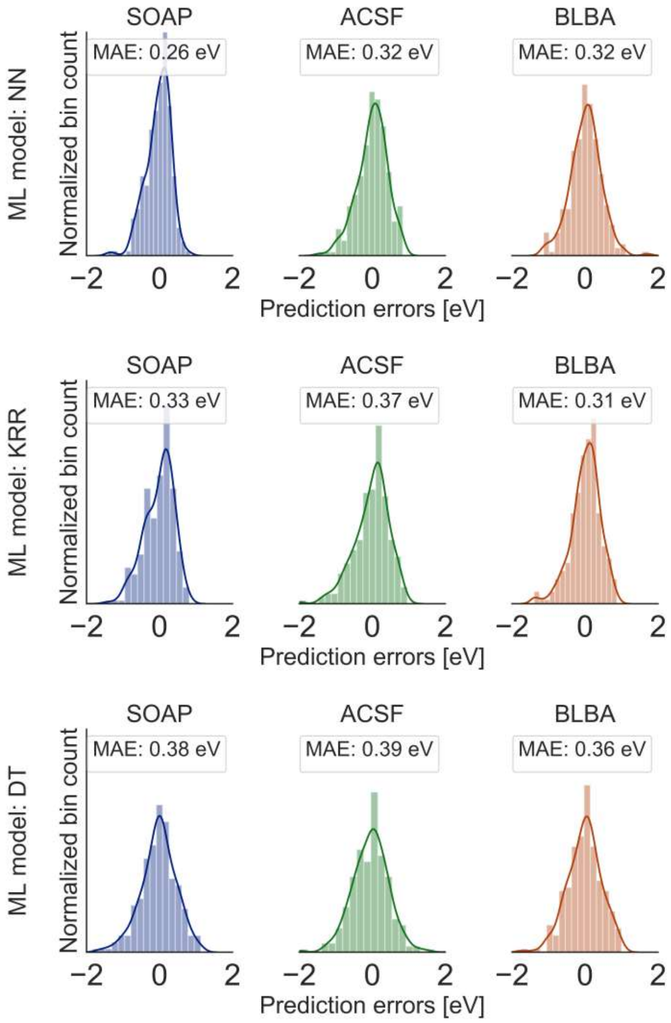

Figure 4.7: Error distributions and MAE in the prediction of formation energy of hydroxyl E' center defects in a-SiO$_2$ for the different combinations of descriptors (SOAP, ACSF and BLBA) and ML models (neural network, KRR and DT).

The study of reliability in modern microelectronic devices requires the use of simulation techniques in order to assess the effects of new materials and changes in the fabrication processes. In this context, there are several practical applications in which such DFT calculations could be aided or even replaced by computationally inexpensive ML models combined with well-established descriptors, as shown in this work and in other examples in the literature [51][144][145]. Moreover, novel descriptors can be developed for specific applications, in the same way in which the BLBA descriptor was developed to be used in the ML-based prediction of formation energies of hydroxyl E' center defects in a-SiO$_2$. This approach could prove particularly useful in the study of amorphous materials, where large statistics are needed. Moreover, this solution could be applied to other defects, as well as to new materials, potentially aiding the development of novel microelectronic devices.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

Finally, apart from providing computationally inexpensive solutions to practical problems in the study of modern electronic device reliability, descriptor-based ML solutions have the potential to provide a deeper understanding of the mechanisms by which defects are formed and of the relationships between atomic structure properties and the formation of such defects. The next section extends this work to predict the defect structure of defects in a-SiO_2.

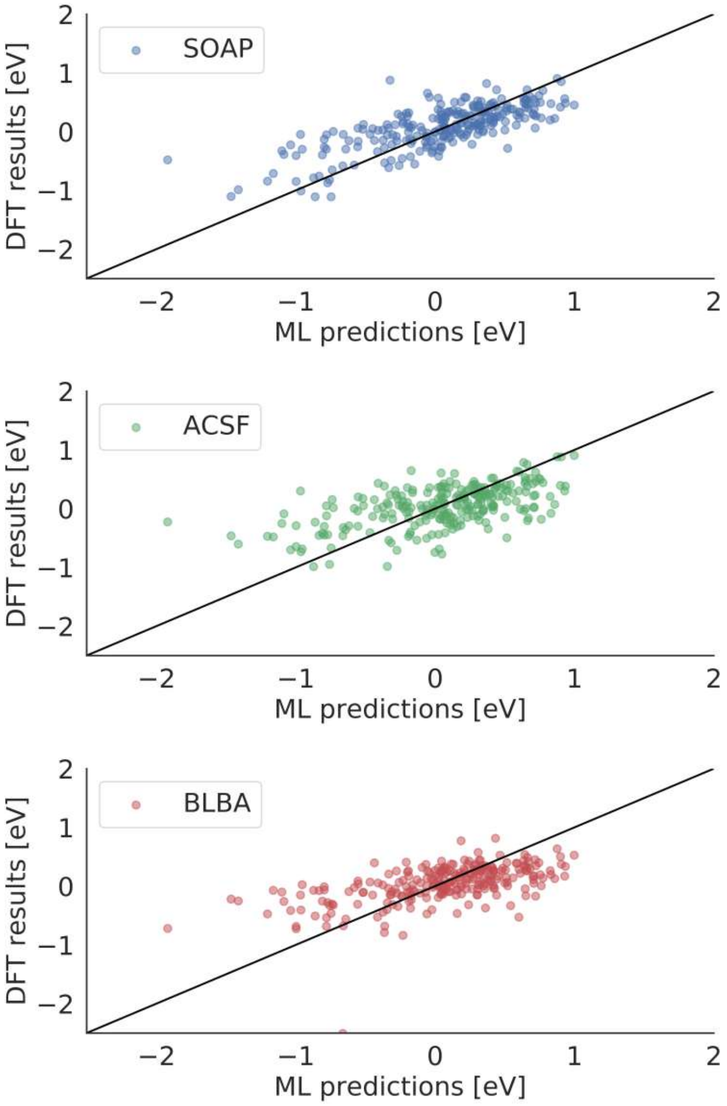

Figure 4.8: Parity plots for the DFT calculation results and the ML-based predictions using SOAP (top), ACSF (middle) and BLBA (bottom) descriptors (in all cases, combined with a neural network). All descriptors perform similarly. However, SOAP shows a slightly better correlation in general, particularly for negative values.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

4.2 Prediction of Defect Structures

The results and figures presented in this section have been partially published at the ESSDERC 2021 conference [146].

This study can be considered as an extension of the work presented in Section 4.1, where a solution to directly estimate the formation energy of hydroxyl E' center defects in a-SiO_2 was proposed. In this section, another solution was developed to predict the structure of such defects. As previously mentioned, the study of these defects is particularly important in the fields of modern micro- and nano-electronic device reliability, since they are suspected to be responsible for BTI and RTN in MOS transistors [137][138][139].

The proposed workflow is graphically summarized in Figure 4.9. The core concept is to train an ML model to find the relationship between the defect-free and the defect atomic structures in the descriptor space. Then, the defect atomic structure is reconstructed from the predicted descriptor, by optimizing its atomic coordinates, until its descriptor matches the predicted descriptor, within a convergence criterion. 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 convergence criteria of 5% of the initial loss function value was used.
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. The geometrical distance between the ML predicted and DFT relaxed defect site was used as a measure for the prediction quality.

The first step towards an efficient ML-based prediction is a consistent training dataset. For this work, the dataset described in the previous section was used. Calculations were performed using the PBE functional [112] in the CP2K software package [111], which is a computationally expensive process. Typically, 4 nodes with 48 cores each require several hours to complete a single geometry relaxation. Once the dataset was created, the next step was to represent these structures using SOAP [54], since this descriptor showed the best performance in the previous section. However, given the flexibility of the proposed approach, any other descriptor, even individually designed ones, can be used within the proposed workflow.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

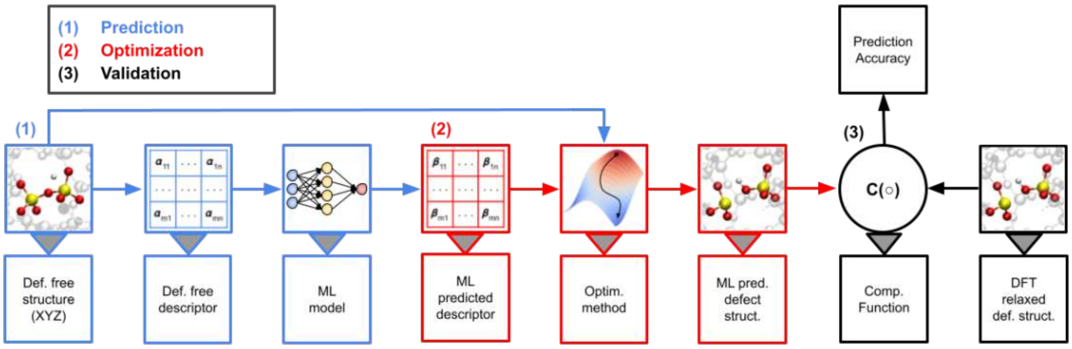

Figure 4.9: Proposed workflow to predict the structure of defects in atomic models. First, the local environment around the atom of interest in the defect-free structure is represented by a descriptor. This matrix contains a vector for every atom near the site of interest, and is used as an input for an 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).

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 dataset and 83 in the testing dataset. 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. To analyze the data, a KRR model was employed, as implemented in the Scikit-learn package [141]. This is a cost-efficient model, which was presented in the previous section.

The ML model is trained to find the relationship between the SOAP descriptions of the hydroxyl E' centers and those of 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.

Given the modular approach of this technique, the same generic workflow could be used in the prediction of other defects in new materials. Details of the particular application in this work, namely the prediction of hydroxyl E' centers in an initially defect-free a-SiO_2 structure, and the individual steps are shown in Figure 4.10. As an example, 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: $r_{\text{cut}} = 3.0$ Å and $n_{\max} = l_{\max} = 4$, as shown in Figure 4.11.

The resulting matrix is used as an input for the ML model, which was previously trained with the training dataset. The 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 MSE between the predicted SOAP and the defect-free structure with one additional H atom placed in the direct vicinity, 1 Å,

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

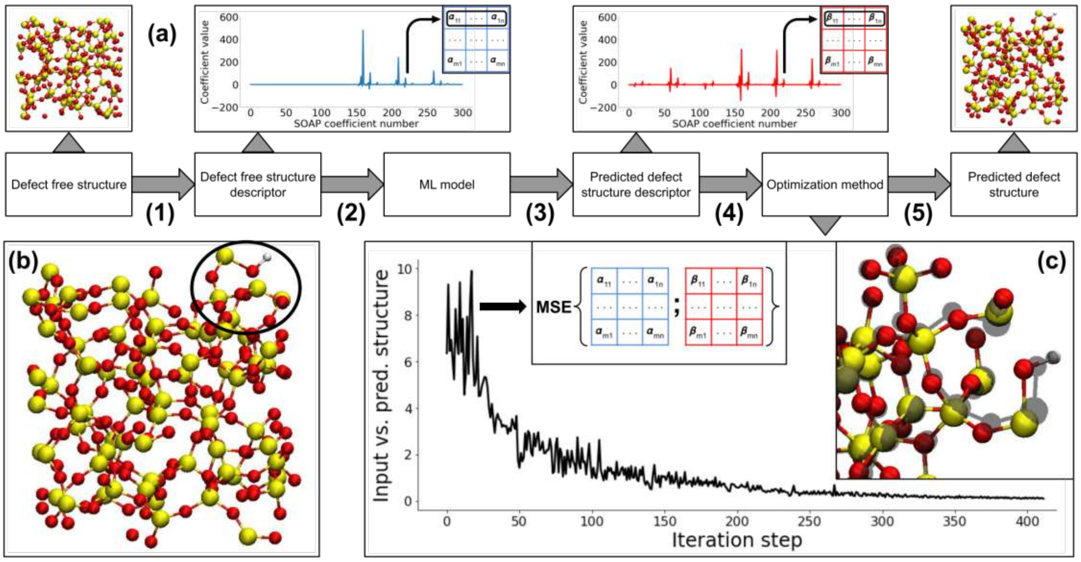

Figure 4.10: (a) Prediction of the formation of a hydroxyl E' center in an a-SiO_2 structure. The local environments of the atoms around the oxygen atom of interest are described with a matrix of SOAP descriptors (1). This matrix is used as input for the ML model (2) to predict the SOAP description matrix 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 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.

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 [147]. However, given the flexibility of this 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 Figure 4.10.

The complete prediction and optimization process takes 0.5 seconds on a typical desktop computer (Intel Core i7 2.2 GHz and 8 GB of RAM). Given the stochastic nature of the amorphous network, the accuracy and efficiency of the proposed framework must be analyzed on a statistical scale. It was therefore benchmarked against the full testing dataset of a-SiO_2. 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 with respect to the number of atoms involved, in order to obtain a normalized quantity.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

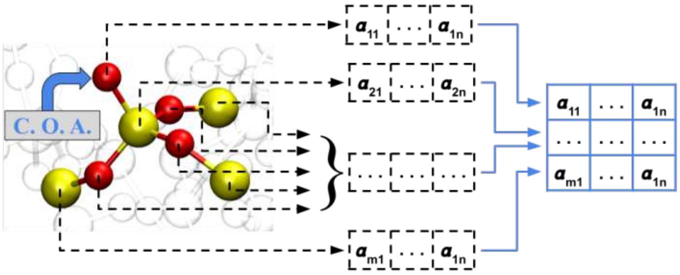

Figure 4.11: Defect-free structure descriptor matrix. Each row of this matrix is the SOAP vector representation of the local environment of a specific atom. Only those atoms within a certain cutoff distance from the central oxygen atom (C. O. A.) of interest are considered.

Apart from the initial loss function tested, loss function A: MSE between the SOAP descriptor of the input and predicted structures, a second loss function was defined, loss function B: loss function A plus a penalty term if the H-O bond-length deviates from 1 Å. The process of predicting the hydroxyl E' center defect structures in the testing dataset was repeated under identical conditions for both loss functions. The results can be seen in Figure 4.12, which shows the distribution of the deviations of the ML based approach compared to the DFT results across the testing dataset. For the 83 structures in the testing dataset, 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.

There is a wide range of applications in which ML based techniques can provide computationally inexpensive but accurate solutions, as shown by several examples in the literature [51][144][145]. In this work, a solution was developed to study the formation of hydroxyl E' center defects in a-SiO_2. The results clearly demonstrate a competitive level of accuracy, while being inexpensive when compared to DFT.

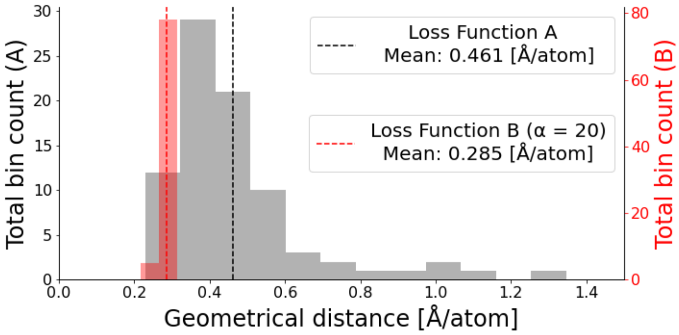

Figure 4.12: Distribution of the deviations of the ML predicted defect structures compared to its equivalent DFT-relaxed structures, together with its mean value. Atoms within a 6.0 Å cutoff distance from the central oxygen atom are considered.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

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).

This approach is of particular importance for defect studies in situations where the application of DFT and other *ab initio* methods is too expensive, such as defect studies in large simulation cells and investigations on a large statistical scale. Furthermore, applications with a demand for on-the-fly calculations of certain data, such as KMC simulations [140], benefit from the presented framework. Currently, reaction rates in such simulations have to be predefined. However, with this approach, structural and energetic information can be calculated almost instantaneously.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Chapter 5

Neuromorphic Hardware Accelerator for Machine Learning-Based Molecular Dynamics

This chapter will present the results of using neuromorphic hardware to accelerate ML-based MD simulations. An overview of the framework used in this work is presented in Section 5.1. In Section 5.2, the use case on which this framework was applied is presented. Finally, in Section 5.3, the obtained results are discussed. The results and figures presented in this chapter are a product of my research stay in the Computational Nanoelectronics Group at Eidgenössische Technische Hochschule Zürich (ETHZ) [27] from March to May 2023. The scanning electron microscope (SEM) images of the memristors and the plots depicting their RV-curves were provided by Till Zellweger (ETHZ).

5.1 Framework

One of the main limitations that commonly used ML interatomic potentials face when employed in running MD simulations is that their performance is severely limited by the hardware currently used to implement them. Presently, ML-based MDs are universally deployed in general-purpose computers, based on the von Neumann architecture. In this architecture, the central processing unit (CPU) is separated from the main memory and is connected to it by a slow buffer, therefore suffering from the so-called *von Neumann bottleneck* [148]. Running ML-based MD simulations is a demanding task, as it requires to process data of thousands of atoms over millions of time-steps. When performing these simulations in a von Neumann computer, 90% of the computing time (and energy) is consumed for the repeated movement of data between the CPU and memory [149].

As an alternative to von Neumann computers, *neuromorphic* architectures stand out for their co-location of CPU and main memory [150]. This characteristic makes neuromorphic computing an ideal candidate as a hardware-accelerator of ML-based

MD simulations. Neuromorphic computing architectures are inspired and based on the human brain, which consumes only approximately 20 W [151], while processing vast amounts of information in a highly parallelized manner, therefore making it an ideal role model for highly energy efficient computing architectures. Recent tests on prototypes of neuromorphic-based hardware accelerators yielded highly promising results when running ML-based MD simulation for Au [152]. For the simulation of more complex systems, such as GeTe, and $Li_{10}GeP_2S_{12}$, the computational times could be reduced by 2 orders of magnitude [149].

Running an ML-based MD simulation requires to repeat a set of specific tasks for every time-step [17]. First, a descriptor of the atomic structure of interest is computed. It relies on a mathematical representation of the local atomic environments and depends on a set of material-specific parameters. Second, a suitable ML model is set up to compute the potential energies and forces of the atomic system from the descriptor. Third, the forces are used to solve the Newton equation. Finally, the coordinates and velocities of every atom are updated, and the process is repeated for the next time-step, thereby creating a trajectory. The framework used in this work is the one proposed by Mo *et al.* [149], which consists of implementing a neural network in a field programmable gate array (FPGA)-based neuromorphic architecture, while performing the remaining computational kernels in a von Neumann computer, as schematized in Fig. 5.1.

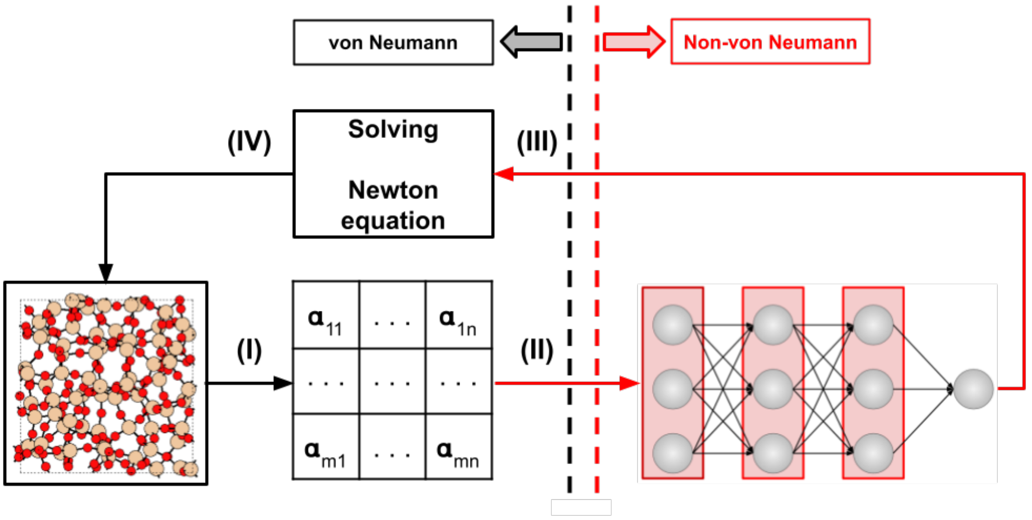

Figure 5.1: Alternative computing architecture schematic, where the MD simulations run on a hybrid platform. The typical steps of an ML-based MD simulation are followed: The descriptors of the local atomic environment of the system are computed (I) and passed to the ML model (II) to compute the energy and forces (III) which are then used to update the atomic coordinates for the following time-step (IV). However, while the PES is evaluated within a neuromorphic accelerator (II-III), the remaining computation requirements are fulfilled by a standard computer (I-IV).

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

The field of neuromorphic computing has also witnessed remarkable progress in the last few decades beyond the development of ML-based MD simulation accelerators. Coined by Carver Mead in the late 1980s [153][154], the term originally referred to mixed analogue-digital implementations of brain-inspired computing and has since extended to encompass a wide variety of approaches. Currently, the development of neuromorphic computing technologies is being spearheaded by both academia and industry. Examples of these projects include Neurogrid [155], BrainScales [156], DYNAPs [157], IFAT [158], Intel's Loihi [159] and IBM's TrueNorth [160].

These pioneering projects have not only advanced the theoretical foundations of neuromorphic computing but have also catalyzed its practical application in various domains. Neuromorphic computing solutions have experienced increased adoption in applications such as pattern recognition [161], where the brain-inspired architecture excels at processing complex visual data. Moreover, in sensing applications [162] and brain-computer-interfaces [163], neuromorphic approaches have demonstrated their potential in enhancing data acquisition and processing capabilities. For a comprehensive understanding of the evolving landscape of neuromorphic computing, numerous reviews in the literature provide in-depth analyses [164][165][166][167].

5.2 Use Case

The previously described framework was employed to develop an ML interatomic potential for amorphous germanium, with the objective of simulating a memristor device fabricated at ETHZ. The ultimate goal is to shed light on the physical mechanism behind this device, which is currently poorly understood. Memristors are a promising class of microelectronic devices that have the potential to revolutionize the field of non-volatile memory and in-memory computing. These devices were first conceptualized by Leon Chua in 1971 [168], and first demostrated by HP Laboratories in 2008 [169]. They exhibit a unique property, where resistance can be tuned in a multi-bit or even analog manner by the application of electrical stimuli, and this change can be retained even after the stimulus is removed. This makes them ideal candidates to implement artificial synapses in neuromorphic computing and brain-inspired information processing technologies.

In a basic model of a current-controlled memristor, the resistance of the device is a function of its state variable, which, in turn, depends on the current that has flowed through the memristor,

$$\begin{aligned} v &= R(w)i \\ \frac{\mathrm{d}w}{\mathrm{d}t} &= i, \end{aligned} \tag{5.1}$$

where v is the Voltage across the terminals of the memristor, R the resistance, i the current and w the state variable. Therefore, a memristor can be seen as the fourth fundamental two-terminal circuit element (with the other three being resistor, capacitor and inductor) [169]. These fundamental elements are shown in Fig. 5.2.

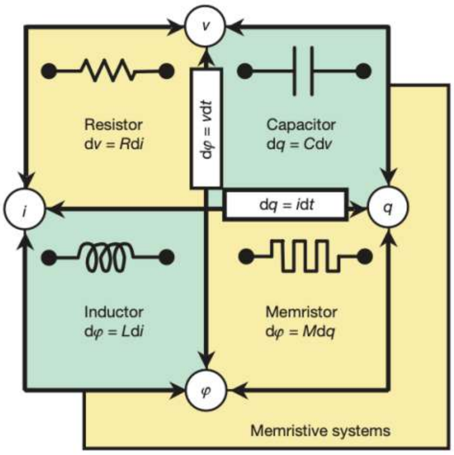

Figure 5.2: The four fundamental two-terminal circuit elements: resistor, capacitor, inductor and memristor. Reprinted from Nature 453, 80-83 (2008), with permission from the journal. Credits: Dmitri B. Strukov *et al.*

The memristor devices fabricated at ETHZ can be seen in Fig. 5.3 and Fig. 5.4. There are many mechanisms that are exploited for memristors, some of the most typical ones are: (I) Oxygen exchange with a metal electrode to create oxygen vacancies that form the connection between electrodes, used in resistive random-access memory (ReRAM) [170]; (II) Phase change, in which the dramatically different electrical resistivity values of the crystalline and amorphous phase of certain materials is exploited as a means to store information, used in phase-change memory (PCM) [171]; (III) Redox reactions and drift or diffusion of ions, also known as electrochemical

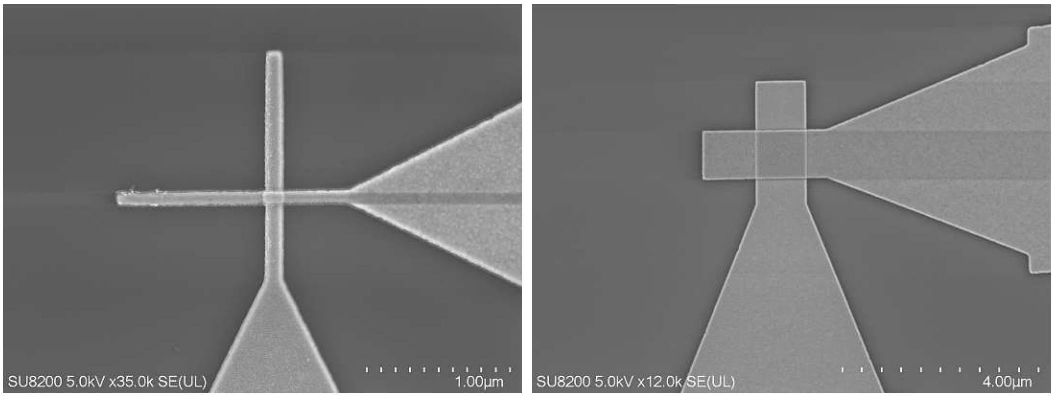

Figure 5.3: SEM image of a novel germanium-based memristor fabricated at ETHZ (left), and zoom to the device (right). The device is composed of a vertical and a horizontal metallic electrode, and amorphous germanium in between them. In both figures, the triangle-like shapes are tapered to connect the small device to larger wires that then connect to electrical pads such that the devices can be contacted with probes, which can be measured. The dark lines that extend from the device area to the right are artifacts of the visualization technique.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

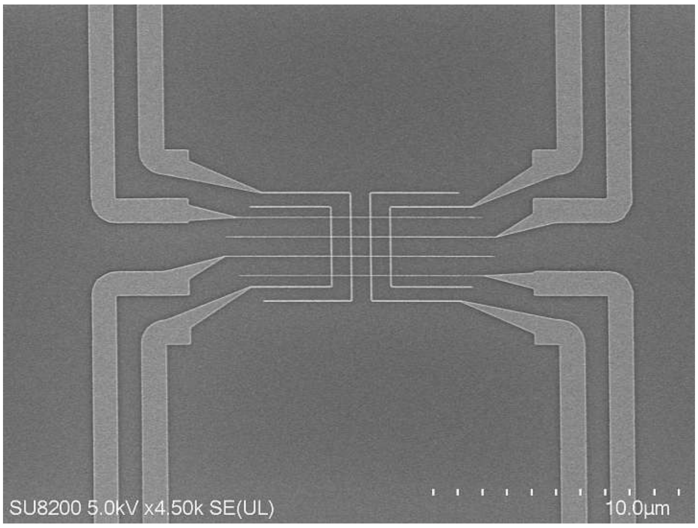

Figure 5.4: SEM image of an array of 16 memristors fabricated at ETHZ. These arrays allow performing matrix multiplication in an analog manner.

metallization mechanism, in which a redox reactions form and dissolve a conductive filament, used in conductive-bridging random-access memory (CBRAM) [172]; and (IV) Ferroelectric polarization switching, used in ferroelectric random-access memory (FeRAM) [173].

In recent years, memristor devices have gained attention due to their attractive properties, including high endurance, which means that the devices can withstand a large number of write-erase cycles without degradation, low power consumption, and compatibility with existing technologies. By applying different electric fields, the device can be switched between different resistance states, making it a promising candidate for non-volatile memory applications.

In this context, the Computational Nanoelectronics Group at ETHZ [27] fabricated a novel memristor based on germanium, a material widely used in electronics due to its CMOS-compatibility and high carrier mobility that allows faster transistor operations. More specifically, this new memristor consists of a thin germanium layer sandwiched between two metal electrodes. These devices show electrical characteristics promising for neuromorphic computing applications, as shown in Fig. 5.5. However, the physical mechanism behind these memristors is poorly understood, and the MD simulations needed to provide a deeper understanding of their operation are computationally expensive. Generating amorphous germanium models of this scale requires running melt-and-quench MD simulations for particularly large system sizes ($>$ 100,000 atoms), as shown in Fig. 5.6, and extended simulation times ($>$ 1,000,000 time-steps); a feat out of reach for *ab initio* methods, and a task particularly well suited for hardware-accelerated ML-based MD simulations. In Section 5.3, the results of this development are presented and validated against DFT.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

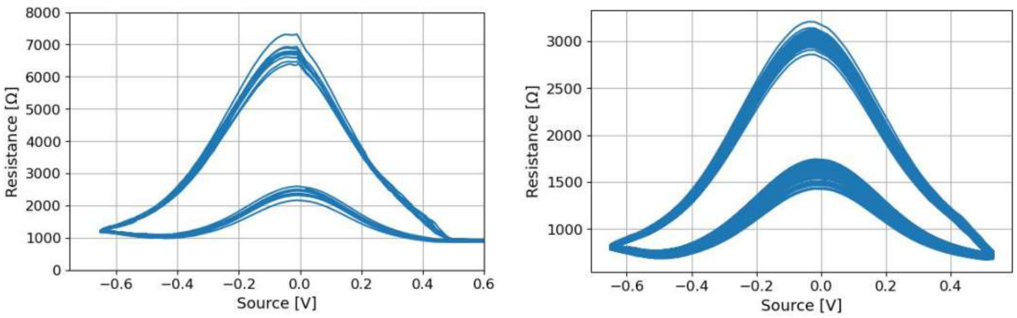

Figure 5.5: RV-curves of two novel germanium-based memristors fabricated at ETHZ. Note that for both devices the transition from the higher resistance to the lower resistance is gradual, meaning that the switching is not digital but analog. The device on the left presents a wider gap between the high- and low-resistant states, while the device on the right presents more cycles with less variance. These are highly desired and useful properties in memristors used in neuromorphic computing applications.

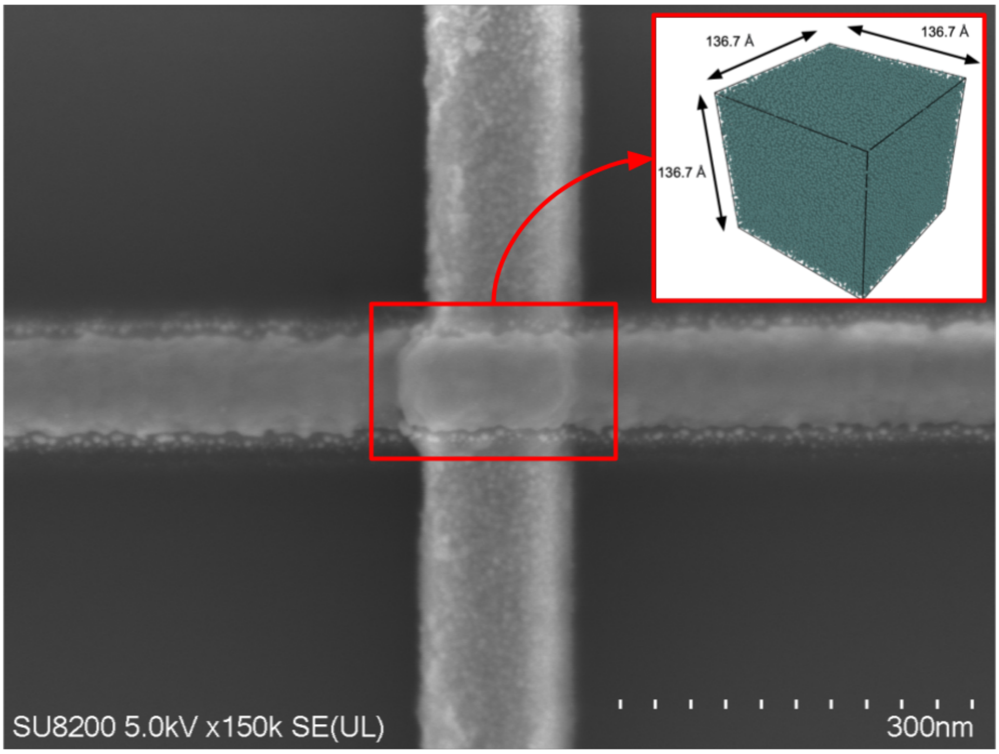

Figure 5.6: Zoom to the SEM image of a novel germanium-based memristor fabricated at ETHZ, together with the representative simulation cell used to perform MD simulations of the device. This is a cubic cell with a lattice parameter of 13.67 nm, which contains 110,592 atoms. The physical device is approximately 10 times larger than this cell.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

5.3 Results

This section presents the results of using non-von Neumann molecular dynamics (NVNMD) to model solid-, liquid- and amorphous germanium. First, a dataset of crystal, liquid and amorphous configurations was created using the melt-and-quench technique [120]. An MD simulation was performed in QuantumATK [110] using the Stillinger-Weber empirical potential [119], starting from a bulk system of 216 germanium atoms. The system was initially heated well above its melting temperature, to 5,000 K, and thereafter quenched back to room temperature with a quench rate of 4.7 K/ps. The resulting trajectory was thereafter sub-sampled to 1,500 atomic configurations, an isolated germanium atom and 50 Ge-Ge dimmers with equally-spaced interatomic distances ranging from 0.5 Å to 5 Å were added to this batch of data. Thereafter, their DFT energies and forces were recalculated using the PBE functional [112] in the CP2K software package [111]. The 1,379 atomic configurations for which the DFT calculations converged were collected, together with their DFT energies and forces into a training dataset.

Once the training dataset was build, the code provided by Mo *et al.* [174], which is a modified version of the commonly used DeepMD package [58], was used to train a continuous neural network (CNN) and thereafter a quantized neural network (QNN). The most relevant parameters used to train these neural networks are presented in Table 5.1, together with a brief definition of each parameter. Once the QNN was trained, it was used to run a melt-and-quench MD simulation on the FPGA-based hardware accelerator framework proposed by Mo *et al.* [175]. This framework allows running MD simulations on a modified version of LAMMPS [132], which provides compatibility with the FPGA hardware accelerator. The test MD simulation started from an initial crystal composed of 216 germanium atoms, which was first melted at 5,000 K and thereafter quenched back to room temperature using over 1,000,000 time-steps. The time-step used was 0.5 fs, therefore resulting in a quench rate of 9.4 K/ps. The final snapshot of this trajectory was saved, geometry-relaxed with the CNN, and validated by comparing it against a DFT reference.

The DFT reference was build by geometry-relaxing the last snap of the initial melt-and-quench MD using DFT in the CP2K software package. This was done in order to avoid running an *ab initio* MD simulation, therefore saving a substantial amount of computational resources and time.

Table 5.1: Most relevant parameters used to train the QNN, which performed the MD simulations on the neuromorphic-based hardware accelerator.

Parameter	**Definition**	**Value**
r_{cut}	Descriptor radial cutoff	4.5Å
r_{Δ}	Smooth cutoff parameter	0.5Å
Net size	Size of neural network	128
sel	Maximum number of neighbors	80
Learning type	Learning rate variant type	Exponential

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

The first validation test consisted in comparing the RDFs of these atomic structures, and the results can be seen in Fig. 5.7, where both functions are in agreement, presenting a large first peak around 2.5 Å and a secondary peak at about 4 Å. The second validation test consisted in computing the structure factor obtained from neutron scattering for both atomic structures. This calculation was performed using the QuantumATK software package [110]. The results are presented in Fig. 5.8. As can be seen, the scattering factors are virtually identical for both atomic structures, therefore indicating a successful training process. The third and final validation test consisted in comparing the bond-length distributions for both atomic structures. The

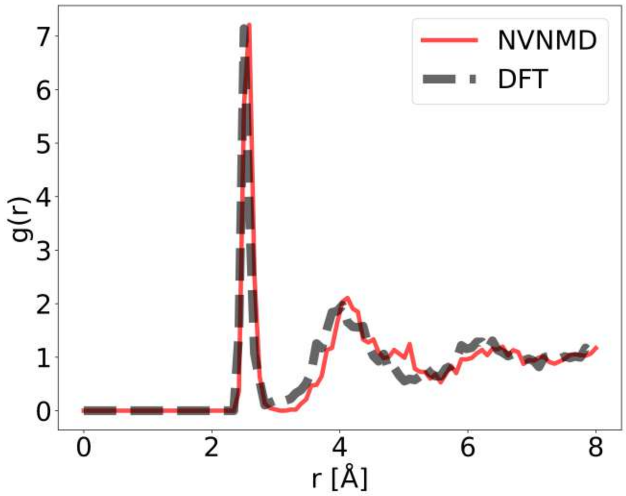

Figure 5.7: Comparison of the RDF for the amorphous germanium generated with NVNMD (red) and DFT (black).

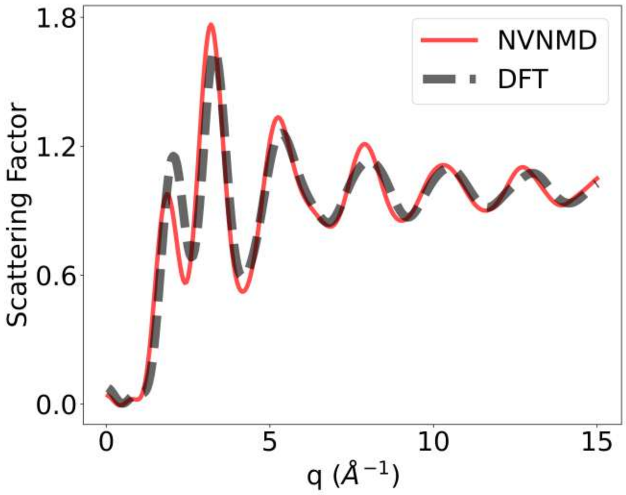

Figure 5.8: Comparison of the structure factor obtained from neutron scattering for the amorphous germanium generated with NVNMD (red) and DFT (black).

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

results are presented in Fig. 5.9. As can be seen, the NVNMD amorphous germanium model resulted in slightly elongated bond-lengths, but an overall good agreement with the DFT reference.

The previously discussed results validated the accuracy of the ML interatomic potential against DFT in MD simulations. However, as previously mentioned, the main motivation to make use of a hardware accelerator is to increase the performance of the MD simulations. Therefore, to evaluate the performance of this framework, a GAP was trained on the same training dataset and both interatomic potentials were used to run the same MD simulation, consisting of running 10,000 time-steps at 1,000 K for a system of 216 germanium atoms. The results are presented in Table 5.2, where it can be seen that using the neuromorphic-based hardware accelerator is over two orders of magnitude faster.

Once the proposed ML interatomic potential was validated for atomic structures composed of 216 germanium atoms, it was used to run a melt-and-quench MD simulation on an atomic system composed of 110,592 atoms, considered to be large enough to simulate the memristor device. The resulting amorphous germanium structure was then validated against the DFT reference and the results are presented in Fig. 5.10. The results are in good agreement with DFT, with both atomic structures presenting a first peak at around 2.5 Å and a secondary peak at about 4 Å.

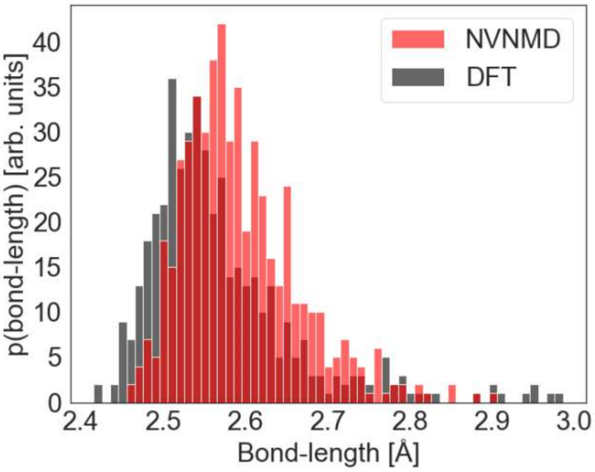

Figure 5.9: Comparison of the bond-length distributions in the amorphous germanium structures generated with NVNMD (red) and DFT (black).

Table 5.2: Comparison of the performance of NVNMD to a traditional GAP interatomic potential, by using both methods to run the same MD simulation in LAMMPS.

Interatomic potential	**Performance**
NVNMD	407 [time-steps/s/core]
GAP	2 [time-steps/s/core]

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

The high accuracy and remarkable efficiency of this ML interatomic potential will be exploited in future work, by running extensive MD simulations that will provide the basis for a better understanding of the physical mechanisms behind these novel germanium-based memristors. Future work will include re-training the ML interatomic potential to include the metallic electrodes, in order to ensure more realistic simulations of the entire device.

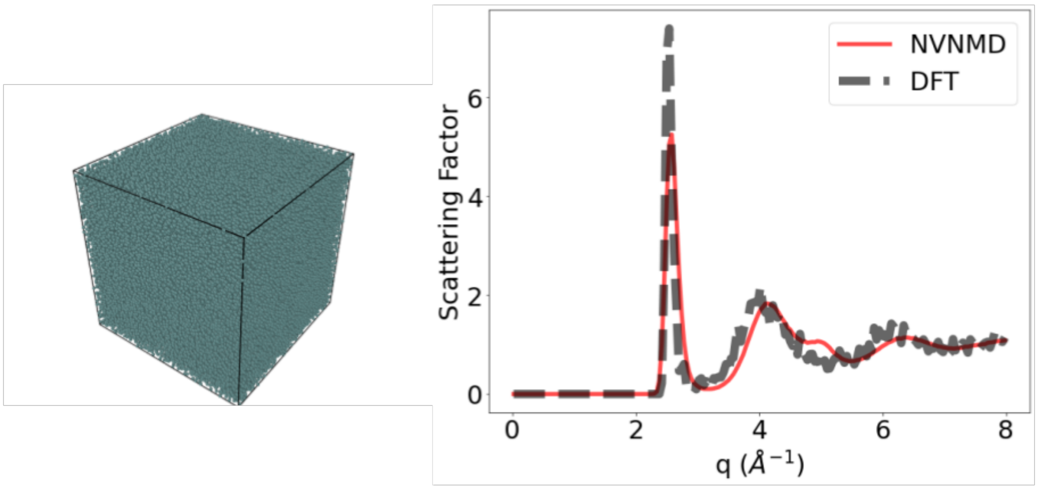

Figure 5.10: Results of running a melt-and-quench MD simulation on a cell containing 110,592 atoms (left), and RDF validation against a DFT reference (right). Note that the main peak of the NVNMD structure is shorter than that of the DFT reference, this is because the latter is markedly smaller, with only 216 atoms.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Chapter 6

Transdisciplinary Application to Biomedical Signals Processing

The results and figures presented in this chapter have been partially published in Nature Scientific Reports [176] and are the product of a collaboration between TUW and Aalto University (Helsinki, Finland) [28].

The object of this study was motor evoked potentials (MEPs), which are biomedical electrical signals measured with electrodes on the target muscle of a patient. MEPs are typically characterized by their latency (i.e., the time elapsed between the stimulation of the target neurons and the beginning of the muscle movement which these neurons control). The characterization of MEPs is used in research and clinical applications to evaluate the corticospinal excitability of a patient. However, the processing of these signals is a laborous manual task, since in a single session thousands of MEPs are recorded, which must be then manually annotated, one at a time.

Automating the analysis of MEPs would save clinicians and researchers time, as well as reduce human errors. However, the amplitudes of MEPs are low, which means that their recordings are highly sensible to noise. Moreover, completely uncorrelated MEPs might yield virtually the same latency. In other words, an algorithm to automate the annotation of MEPs would face the problem that highly different inputs might produce highly similar outputs, therefore making it a complex task.

In this work, the hypothesis was that MEPs belonged to a highly-dimensional space in which relevant symmetries were missing. This space was considered to be analog to the XYZ-coordinates of atomic configurations, in which translation, rotation and permutation symmetries are missing [177]. The thesis was that, in the same way in which a descriptor is used to transform atomic configurations into mathematical representations suitable for ML regression, the same approach could be used in the processing of MEPs. Therefore, MEPs could be transformed into a new space, more suitable for ML regression, thereafter, a neural network could be trained to find the relationship between the MEPs in this new space and their latencies. The result of this work was the first ML-based algorithm to automate the annotation of MEPs. An introduction to this work can be found in Section 6.1, the methodology is described in Section 6.2, and the results are presented in Section 6.3.

6.1 Introduction and Motivation

The MEP generated by transcranial magnetic stimulation (TMS) is a crucial neurophysiological signal in research and clinical practice. MEP amplitude and latency allow one to quantitatively assess the corticospinal excitability. This is necessary to evaluate patients undergoing surgery and to monitor neuromotor diseases, such as the progression of multiple sclerosis [178], idiopathic generalized epilepsy [179] and the recovery of stroke patients [180]. MEPs are commonly characterized by their latency, which is defined as the time elapsed between the stimulation and the onset of the MEP (Figure 6.1). The MEP latency is usually annotated manually after visual inspection of the electromyography (EMG) recording, making the process time-consuming, operator-dependent, and prone to errors [179][181]. An algorithm to automate the characterization of MEPs would not only save time and reduce human errors, but would also boost the development of brain-state-dependent and closed-loop brain stimulation protocols, by allowing accurate real-time MEP assessment [182].

Several attempts have been made to develop algorithms to automate the MEP latency annotation. These algorithms are based either on absolute hard threshold estimation (AHTE) [183] or on statistical measures [184][185]. A review and comparison of previous methods is presented by Šoda *et al.* [186], together with their own algorithm named Squared Hard Threshold Estimator (SHTE). In general, the previously presented algorithms require the user to specify a set of so-called *magic numbers*. These are hyperparameters with a large impact on the algorithm performance, which are empirically derived and depend on the user's knowledge and experience [186]. Contrary to these traditional algorithms, Bigoni's method [187], which is a derivative-based algorithm, does not require the user to specify magic numbers.

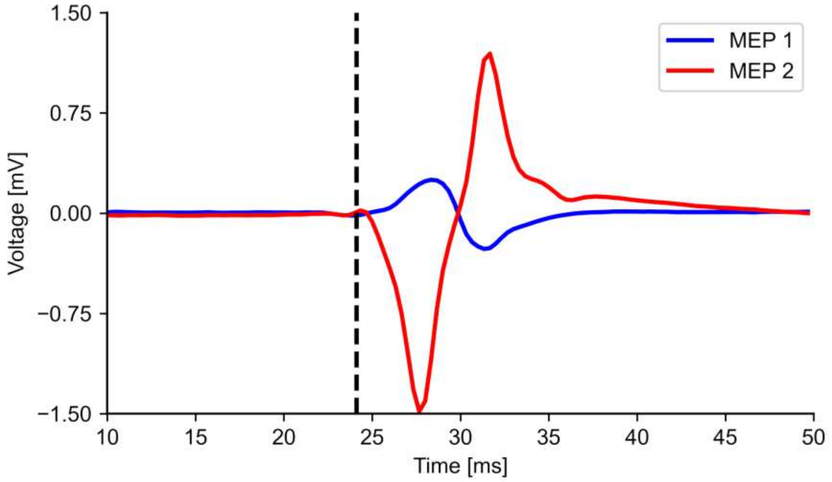

Figure 6.1: Two different MEP waveforms (red and blue curves) with virtually identical latencies (black vertical line), defined as the time elapsed between the TMS pulse and the beginning of the MEP trace. The epoch starts at the time the TMS pulse is delivered.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Developing an algorithm to automate the annotation of MEPs is not trivial. Even in ideal conditions of high signal-to-noise ratio (SNR), MEPs are highly variable, presenting significant inter- and within-subject amplitude variability [188][189][190]. Similarly, the MEP latency variability is well known and has been previously documented for neurosurgical patients [188][191]. Under similar circumstances, the amplitude of two MEPs can differ up to an order of magnitude and present totally different shapes, as shown in Figure 6.1. Furthermore, low-amplitude MEPs, commonly recorded in inhibitory stimulation paradigms with paired-pulse TMS have inherently lower SNR than high-amplitude MEPs, thus adding a new layer of complexity. These factors make it demanding to assess the MEP latency automatically and reliably.

In this context, ML-based algorithms, particularly those employing deep learning techniques, offer a promising approach to provide an accurate and reliable solution. The MEP latency annotation is a pattern recognition problem, where deep learning methods have already demonstrated their potential [192]. In this work, the DELMEP algorithm was developed, which relies on deep learning for automated MEP latency annotation. It can be argued that DELMEP has the potential to expedite data analysis procedures and streamline the creation of closed-loop brain stimulation protocols, along with the generation of personalized solutions. This is the first ML-based solution to the problem of automating the MEP latency estimation.

6.2 Methodology

A dataset was collected from 9 healthy volunteers (3 women and 6 men, mean age: 30 years, range 24-41) for two previous studies [193][194], which describe the detailed experimental protocol and stimulation paradigms. Experiments were performed in accordance with the Declaration of Helsinki and approved by the Coordinating Ethics Committee of the Hospital District of Helsinki and Uusimaa. All participants gave written informed consent before their participation.

The application of TMS was done with a monophasic trapezoidal waveform by a custom-made multi-channel TMS (mTMS) power electronics [195] connected to a 2-coil transducer capable of electronically rotating the peak induced electric field [193]. EMG signals were digitized using an eXimia EMG 3.2 system (Nexstim Plc, Finland; sampling frequency 3 kHz; 10-500 Hz band-pass filter). MEPs were collected with single-pulse and paired-pulse paradigms. The paired-pulse stimuli were delivered with interstimulus intervals of 0.5 and 1.5 ms (short-interval intracortical inhibition, low-amplitude MEPs) and 6.0 and 8.0 ms (intracortical facilitation, high-amplitude MEPs). The conditioning stimulus intensity was 80% of the resting motor threshold and the test stimulus and single pulse intensity were both 110% of the resting motor threshold. MEPs were recorded from the abductor pollicis brevis, abductor digiti minimi and first dorsal interosseous muscles. EMG recordings showing muscle pre-activation or movement artifacts greater than ±15 µV within 1 s before the TMS pulse were removed from the analysis. The raw MEPs were visually inspected, and the latency was manually annotated by a single expert (doctoral candidate; 7 years of

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

experience) and quality-checked by a second expert (postdoctoral researcher; 10 years of experience) who confirmed the latency annotations. The aforementioned experts were part of this study. However, the dataset was collected and the latencies were annotated for two prior studies [193][194] conducted before the conceptualization and development of DELMEP. Therefore, the annotations were performed independently of the development of the DELMEP algorithm. An additional validation was performed on an external MEP dataset annotated by three experts. Data preprocessing and annotation was performed with custom-made scripts written in MATLAB R2017a (MathWorks Inc, USA). A total of 33,060 MEPs were recorded, i.e., 11,020 from each muscle group. From all MEPs, 232 (0.7%) were discarded because of pre-activation and 11,548 (34.9%) were discarded because of noise or no-response. Out of the remaining 21,244 MEPs, the validator discarded 4,569 (21.5%) and approved 16,675 (78.5%). Therefore, in total, the dataset is composed of 16,675 MEPs together with their peak-to-peak amplitudes and latencies.

To automate the MEP latency assessment, the DELMEP algorithm was developed in Python 3.8 and its pipeline is composed of the following steps (Figure 6.2): (I) pre-processing and (II) latency estimation with a neural network:

Pre-processing (step I): The pre-processing simplifies the training and use of the neural network. Without the pre-processing, the high variability of the MEPs would

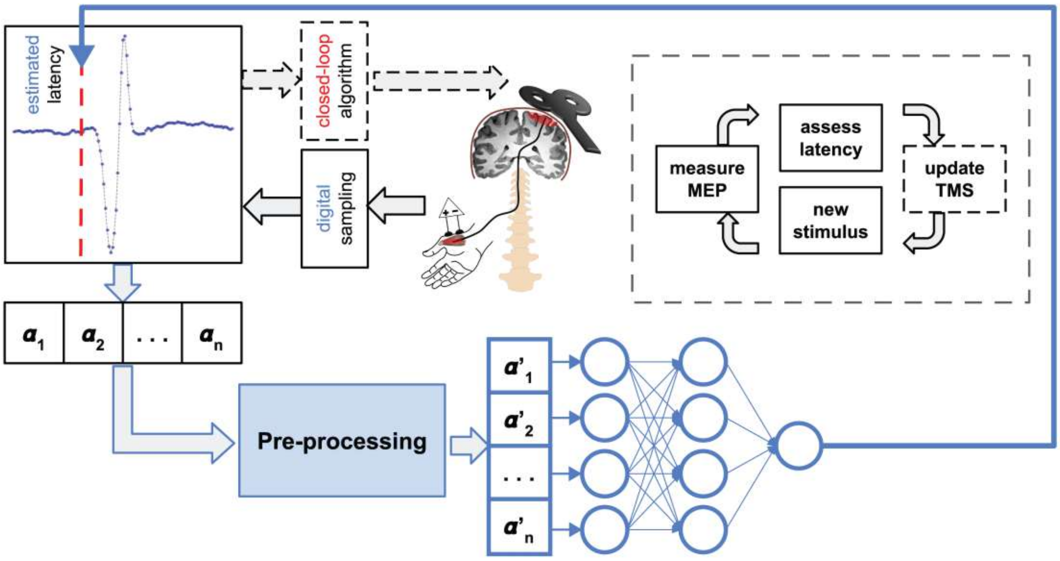

Figure 6.2: Workflow for automated assessment of MEP latencies in a possible closed-loop TMS set-up. MEPs are measured with electrodes placed on the target muscle and stored in a 120-dimensional vector. The pre-processing is done by trimming, smoothing, centering, and normalizing the MEP. The resulting vector is used as an input to the neural network for the latency estimation. The dashed arrows show how the DELMEP algorithm could be applied to a closed-loop protocol (dashed box), in which the brain stimulation parameters are modified depending on the MEP responses.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

require that the neural network "learns" different inputs (MEP traces) corresponding to similar outputs (latencies). Here, the MEPs are represented by a mathematical vector of the raw voltage measurements, significantly reducing the complexity and increasing the speed of deep learning algorithms necessary to process the data. Hence, in this step, the data are (1) trimmed, (2) smoothed, (3) centered, and (4) amplitude normalized. The MEPs are trimmed from 10 to 50 ms after the TMS (120 samples). This is done to reduce and standardize their length, because in the resting condition, the measurements shortly after the TMS and much later than the end of the MEP do not carry relevant information. On the contrary, their inherent noise could pose a problem to the training and use of the neural network, since it would unnecessarily increase the dimensions of the input vectors. After trimming, the MEPs are smoothed with a moving average filter with a window length of 3 samples, to reduce the high-frequency noise of the recordings [196]. Next, the MEPs are centered by computing their mean value in the first 15 samples (5 ms) and then subtracting it from every sample. This step reduces the impact of low-frequency noise in the measurements, by counteracting the shifting it produces in the mean MEP value. The window length in the smoothing step and the time window in the centering step were tuned employing a grid search algorithm. Lastly, the MEPs are normalized so that their minimum and maximum values correspond to 0 and 1, respectively, to mitigate the effects of the large variations in amplitude. A detailed representation of this preprocessing is illustrated in Figure 6.3, where the changes on the two MEPs presented in Figure 6.1 are shown.

Deep learning algorithm (step II): The pre-processed MEPs are used as inputs to the neural network, which produces a latency prediction as its output. This neural network is built as a multi-layer fully connected perceptron layout with two hidden layers of 30 artificial neurons each, and an output layer. A rectified linear unit activation function was used, and the network was trained with the Adam optimizer [71] (early stopping criteria: 200 epochs; batch size: 32), as implemented in the software package Keras 2.4.3 [197].

From all the MEPs available, 2,113 (13%) are low amplitude (peak-to-peak amplitude (V_{PP}) <100 µV, 2,995 (18%) are medium amplitude 100 µV $<$ V_{PP} <200 µV while 11,565 (69%) are high amplitude (V_{PP} $>$ 200 µV. The MEPs were first divided randomly into a training (13,340 MEPs) and testing (3,335 MEPs) dataset, with a training/testing ratio of 80/20. The accuracy and repeatability of the method was verified by using it to evaluate the latency of the 13,340 and 3,335 MEPs in the training and testing datasets, respectively, and comparing these results with the manual assessment of the expert. The comparison was made by computing the MAE between the latencies provided by the method and those provided by the expert. The latency prediction error was analyzed by computing the correlation of the automated latency estimate with the two main MEP features: V_{PP} and the manually annotated latency. The computational times were also estimated for the DELMEP algorithm using a standard computer (CPU Intel Core i7-5650U 2.2 GHz and 8 GB of RAM).

For comparison, the following algorithms were used to estimate the latency of the MEPs in the testing dataset: Signal Hunter [185], AHTE [186], SHTE [186] and Bigoni's method [187]. These estimations were then compared to the manual anno-

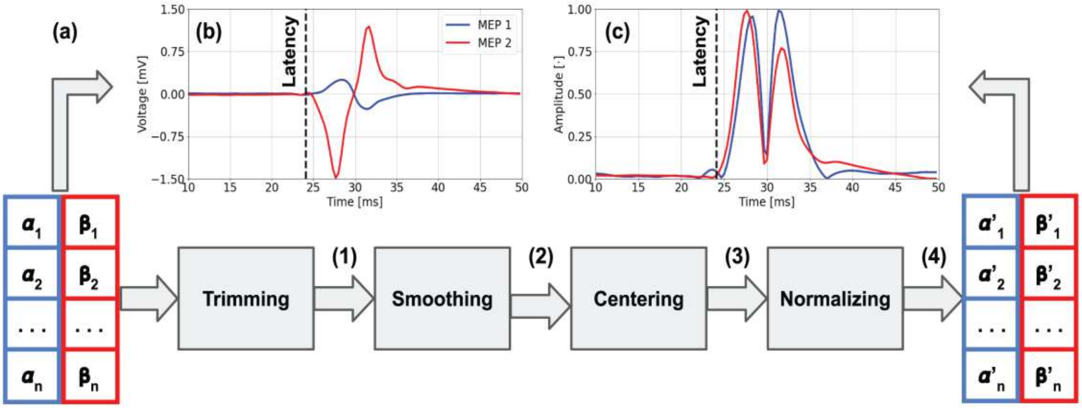

Figure 6.3: (a) Pre-processing of the two MEPs shown in Figure 6.1, divided into: trimming (1), smoothing (2), centering (3) and normalizing (4). Panels (b) and (c) show the MEPs before and after the pre-processing, respectively. When comparing (b) and (c), note that the different MEPs look similar after the pre-processing, thus facilitating the training and later use of the neural network.

tations of the expert and the MAE was computed for every method. Signal Hunter is an open source software for MEP analysis having a latency estimation algorithm based on statistical measures, it performs a moving average filtering on the MEP, differentiates the smoothed signal, calculates the standard deviation (SD), and finds the index value for which the difference between the absolute differentiated MEP value and the SD is the largest; thereafter, it estimates the MEP latency by subtracting a user-selected magic number from that index value. For these tests, Signal Hunter was employed with a magic number equal to 5, following the author's original implementation [185]. The AHTE algorithm performs an absolute value operation on the MEP, finds its maximum amplitude and determines the threshold value ($\mathrm{V_{thr}}$), marks ± 10% around the mean value of the MEP, and finds the index value where the marked line is crossed by the MEP for the first time. The latency estimation is obtained by subtracting a user-selected magic number from that index value. The SHTE algorithm is based on the same principle as the AHTE algorithm, but it works by squaring the MEP coefficients, instead of performing an absolute value operation. The AHTE and SHTE algorithms ($\mathrm{V_{thr}} = 10\%$ and magic number $= 5$) were implemented as done in [186]. Bigoni's method is a derivative-based method, it reduces the MEP to a window of 10-50 ms after the stimulation, finds the peak and trough of the MEP, performs an absolute value operation, computes the approximate first derivative of the MEP until the peak, finds the longest vector of consecutive samples having a positive derivative, and estimates the latency as the first sample of this vector. All algorithms were implemented in Python 3.8.

To evaluate the generalizability of the DELMEP algorithm, a cross-validation (CV) was performed both within and across subjects. In a within-subject test, the data of each subject was split in 5 folds; 80% of the MEPs were used as training dataset and the remaining 20% as the validation set, interchangeably. Final results were obtained by computing the average MAE and SD of MAE across folds for each

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

subject separately. The inter-subject variability of the DELMEP algorithm was tested using a leave-one-subject-out approach. In this test, the data from all but one subject was used to train the model and the MAE was estimated using the data from the left-out subject. This was repeated for all subjects and the MAE was computed in every iteration.

An additional validation was performed in which the DELMEP algorithm and Bigoni's method were used to estimate the latency of the MEPs in an independent dataset, which is composed of 1,561 MEPs and described in detail in the study by Bigoni *et al.* [187]. This dataset was collected from 16 healthy volunteers (eight women and eight men; age: 26.7 ± 2.6 years). The latencies were manually annotated by three different experts (with 0.5, 5 and 14 years of experience) who did not take part in the development of DELMEP. For validating DELMEP, the ground truth (GT) latency was computed as the mean value from the three annotations. About 99% of the MEPs in this dataset have a high amplitude ($V_{PP} > 100$ µV), as these MEPs were collected using a single-pulse paradigm, with a test-intensity chosen to produce an MEP amplitude of 0.50 mV.

6.3 Results and Discussion

Training the neural network on a dataset of 13,340 MEPs required about 2 minutes and pre-processing an MEP trace required 1.2 ms. On average, annotating a pre-processed MEP required 65 µs. The estimated MEP latencies by the DELMEP algorithm and corresponding MAE, for the testing and training datasets are illustrated in Figure 6.4. The similarity between the automated DELMEP and the manual expert annotation suggests a successful training process, since the MEPs in the testing dataset were not used to train the neural network.

To provide a practical example of the DELMEP performance, Figure 6.5 shows eight MEPs and their corresponding automated and manually annotated latencies.

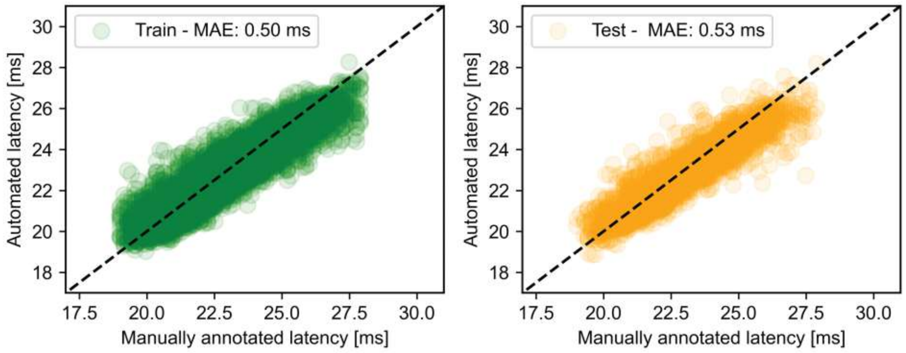

Figure 6.4: Automated MEP latency annotations with the proposed DELMEP algorithm in the training dataset (green) and testing dataset (orange). The results are compared to the manually assessed values. The MAE is presented for both datasets.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Although such a small sample can only be considered as an illustrative example, it provides a notion of how the proposed algorithm performs when used to replace a human expert in MEP annotations.

Figure 6.6 illustrates the error associated with the DELMEP algorithm for the corresponding MEP V_{PP} and manually annotated latency. The sub-panel shows the correlation between the DELMEP latency estimation error and the MEP V_{PP}, for MEPs with an estimation error equal or higher than 1 ms. From the 3,335 MEPs assessed by DELMEP in the testing dataset, 1,895 (57%) had an error lower than 0.5 ms and 2,924 (88%) had an error lower than 1 ms.

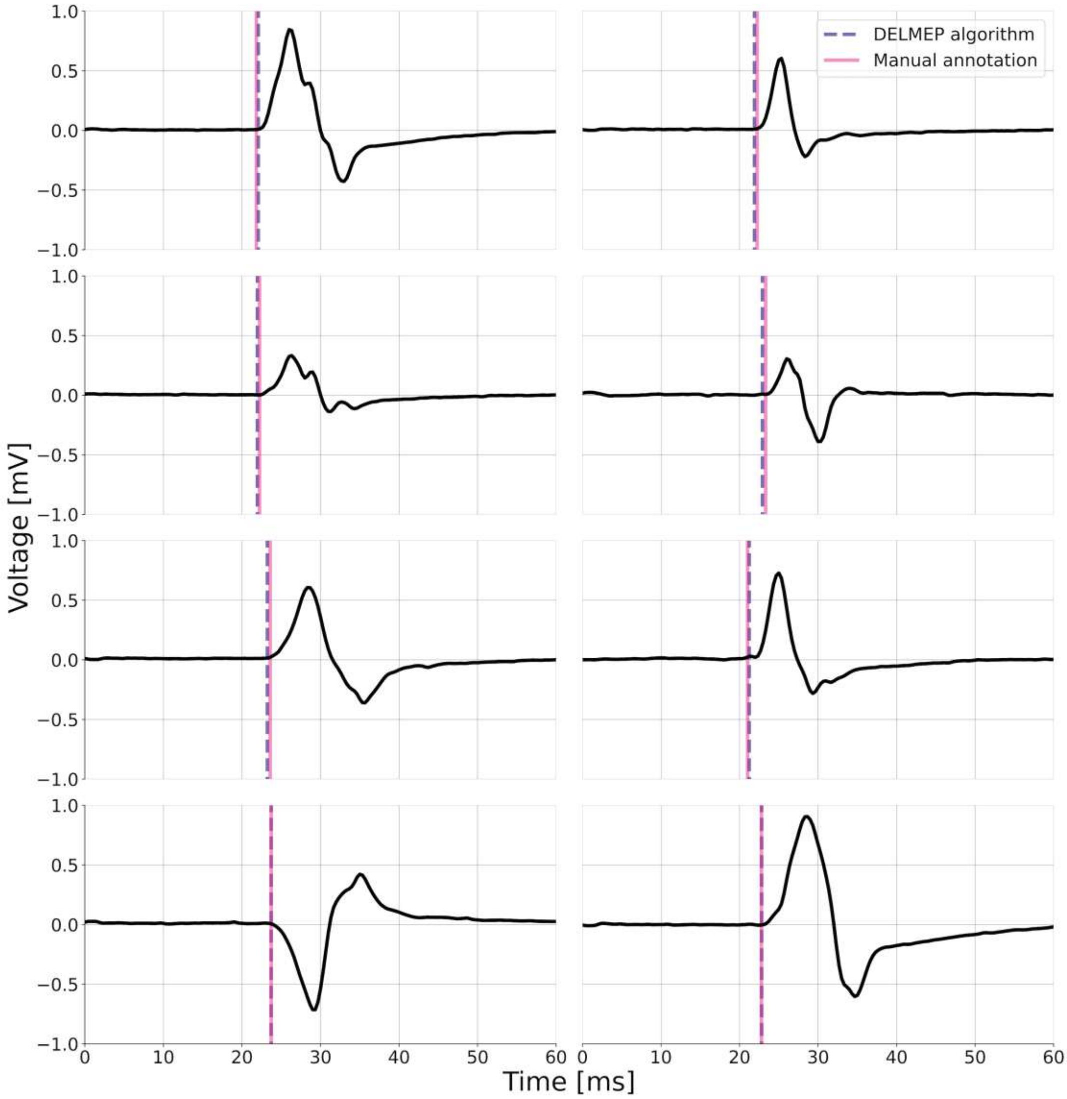

Figure 6.5: Illustrative MEPs from the testing dataset and their corresponding automated (dashed violet vertical line) and manually assessed (purple vertical line) latencies. These MEPs were not used to train the neural network. The similarity between both latencies indicates a good performance of the DELMEP algorithm.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

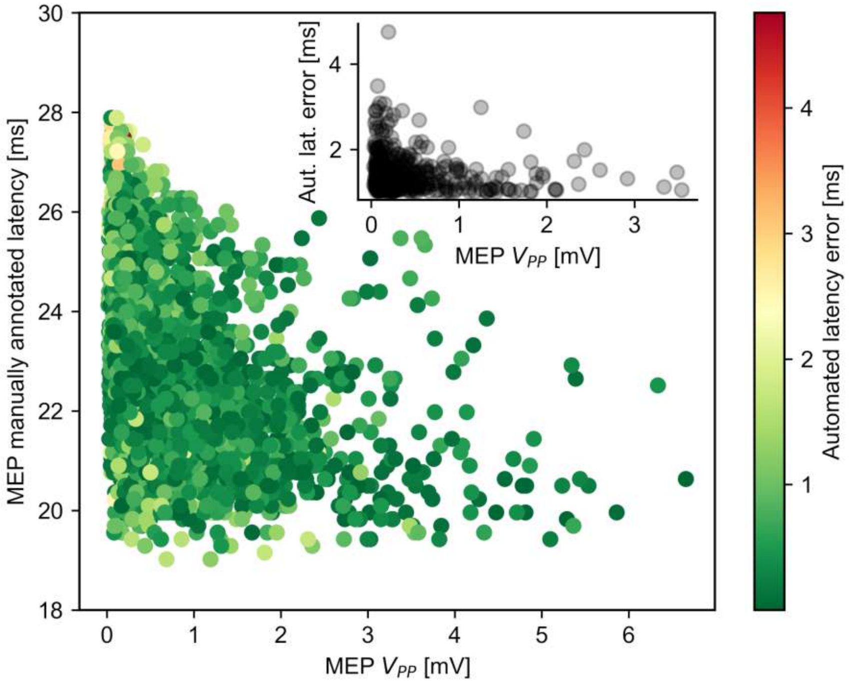

Figure 6.6: Map of the DELMEP algorithm errors in estimating MEP latencies, as a function of the MEP V_{PP} and manually annotated latency. The upper-right panel shows the DELMEP estimation error versus the MEP V_{PP}, for MEPs with latency estimation error higher than 1 ms.

The results from the comparison between DELMEP and Signal Hunter, AHTE, SHTE, and Bigoni's method are shown in Table 6.1, where the MAE is reported for the entire testing dataset and also divided between high- and low-amplitude MEPs. It is important to notice that Bigoni's method discards MEPs when it is not able to find a long enough vector of samples with positive derivatives. The minimum number of samples with positive derivatives in this implementation was set to five, following the original author's implementation [187]. This resulted in 483 out of 3,335 MEPs in the testing dataset (15%), most of which corresponded to low-amplitude MEPs, being discarded by Bigoni's method. To make a direct comparison, only the remaining MEPs were considered to compute the MAE of every method. However, for the CV, in Table 6.2, the entire dataset was used. The resulting MAE from the five-fold CV when using each batch for testing is reported in Table 6.2, together with the average MAE for all tests and its SD.

The intra-subject variability was analyzed by computing a 5-fold CV using data from one subject at a time, and repeating the process for all subjects. The MAE for each data batch of each subject is reported in Table 6.3, together with the average and SD for all tests. Furthermore, the correlation between error and dataset size for every subject is shown in Figure 6.7, together with a fitted curve.

The inter-subject variability was analyzed by using the data of one of the subjects for testing and the data from the remaining eight subjects for training; this process

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Table 6.1: Comparison of MAE obtained with DELMEP, Signal Hunter, AHTE, SHTE, and Bigoni's method for the MEPs in the testing dataset. The MAE is reported for the entire dataset and for low ($V_{PP} < 100$ µV) and high ($V_{PP} > 100$ µV) amplitude MEPs. MEPs discarded by Bigoni's method were also discarded on the other methods for a direct comparison.

Method	MAE low amp. [ms]	MAE high amp. [ms]	**MAE all MEPs [ms]**
Signal Hunter	22.9	6.3	**7.7**
AHTE	16.5	2.8	**4.0**
SHTE	7.3	1.3	**1.8**
Bigoni's method	1.0	0.8	**0.8**
DELMEP	0.6	0.5	**0.5**

Table 6.2: CV of DELMEP across the entire dataset. The available data was divided into five different batches. In each test, one of the batches was used to test the algorithm and the remaining batches to train it. The MAE is reported for every batch (CV 1-5), as well as its average and SD for all batches.

CV 1	CV 2	CV 3	CV 4	CV 5	**Avg MAE [ms]**	SD [ms]
0.5	0.5	0.5	0.5	0.6	**0.5**	0.03

was repeated for each subject. The MAE for each subject together with the average and SD for all tests are reported in Table 6.4.

The DELMEP algorithm performed better than traditional hard-threshold based algorithms across different MEP amplitude ranges. This improved performance is especially noticeable for low-amplitude MEPs, commonly recorded at low stimulation intensities, when computing the motor threshold, as well as in inhibitory paired-pulse paradigms [198][199][194]. For example, with the same low-amplitude MEPs ($V_{PP} < 100$ µV) in the testing dataset, DELMEP, Bigoni's method, SHTE, AHTE and Signal Hunter yielded an MAE of 0.6, 1.0, 7.3, 16.5 and 22.9 ms, respectively; with DELMEP being about one order of magnitude more accurate than these algorithms. On the other hand, with the same high-amplitude MEP ($V_{PP} > 100$ µV) in the testing dataset, DELMEP, Bigoni's method, SHTE, AHTE and Signal Hunter yielded an MAE of 0.5, 0.8, 1.3, 2.8 and 6.3 ms, respectively. This is possibly due to the consistent accuracy of the DELMEP algorithm regardless of the MEP amplitude. Such higher prediction errors correlated with lower MEP amplitudes can be explained by the inherently lower SNR, which has a stronger effect on methods relying on hard-threshold estimators [186].

From the user point of view, both DELMEP and Bigoni's algorithms work by providing the MEP trace as an input and obtaining the estimated latency as an output. The ML nature of DELMEP makes it a more complex algorithm than Bigoni's method. However, this does not translate into a disadvantage for the user, since the open-access code is ready-to-use and no experience in ML is required for using it in research and/or clinical applications. Re-training DELMEP on a new training dataset

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 6.3: Intra-subject variability of DELMEP. A 5-fold CV was performed using data from one subject at a time, the process was repeated for all subjects. The table contains the MAE for every data batch of each subject, together with the average and SD for every subject.

Nº	CV 1	CV 2	CV 3	CV 4	CV 5	**Avg MAE**	SD
1	0.6	0.7	0.6	0.8	0.6	**0.7**	0.08
2	0.5	0.5	0.5	0.5	0.5	**0.5**	0.01
3	0.6	0.6	0.6	0.6	0.6	**0.6**	0.02
4	0.7	0.7	1.0	0.9	1.0	**0.9**	0.14
5	0.5	0.6	0.6	0.6	0.5	**0.5**	0.03
6	0.5	0.6	0.5	0.6	0.6	**0.6**	0.03
7	0.5	0.5	0.5	0.5	0.6	**0.5**	0.03
8	0.6	0.6	0.6	0.6	0.7	**0.6**	0.05
9	0.4	0.5	0.5	0.4	0.5	**0.5**	0.04

Table 6.4: Inter-subject variability of DELMEP. The data from one subject was used for testing and the data from the remaining subjects for training. This process was repeated for every subject. The MAE for every subject, together with its average and SD are reported.

CV 1	CV 2	CV 3	CV 4	CV 5	CV 6	CV 7	CV 8	CV 9	**Avg MAE**	SD
0.7	0.6	0.7	0.5	0.6	0.6	0.5	0.7	0.5	**0.6**	0.1

requires just minor changes to the source code and a few seconds of running time in a regular desktop computer. Both algorithms require minimal human labor time and, unless modifications to the code are intended, minimal interventions and technical knowledge as well.

From a technical point of view, the main difference between DELMEP and Bigoni's method is that the former is an ML algorithm, which "learns" how to annotate MEPs through a dataset of examples; while the latter is a rule-based algorithm, which finds the latency of MEPs by following a static set of steps. This makes Bigoni's method simpler and a more explicable algorithm than DELMEP. However, an important advantage of the deep learning approach is the possibility to pre-train and apply the neural networks on application-specific datasets. For instance, separate models can be created for MEPs from the leg, forearm, and hand muscles, which naturally have distinct latencies [200][201]. Therefore, this approach may provide more accurate automated annotations for a wider set of applications. Deep learning algorithms can also be used in active-learning processes to constantly and automatically improve the accuracy of their annotations [202][203][204][94], by periodically retraining them on data generated during their utilization. This is of special importance for applications on personalized medicine. As depicted in Figure 6.7, when training and testing on data from a single subject, the latency estimation errors were noticeably reduced as the size of the available dataset was increased. Thus, the proposed DELMEP algorithm could be trained on already-available annotated MEPs of one particular

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

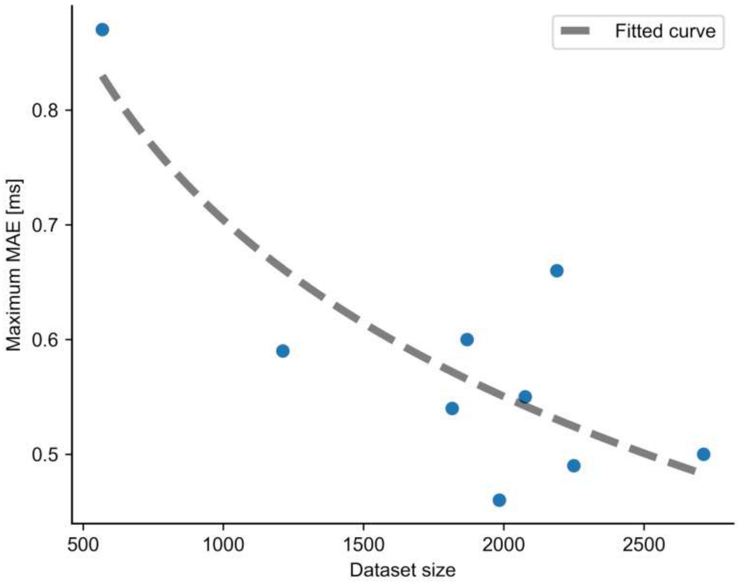

Figure 6.7: Relation between the average MAE obtained during the intra-subject CV tests and the dataset size of every subject (number of MEPs). Each data point represents a different subject in the dataset.

subject, and thereafter used to automate the annotation of MEPs of this subject, in order to ensure the best possible accuracy. This is a feature that non-ML algorithms do not have, due to their static set of rules.

A deep learning-based algorithm requires a large dataset for training. However, for a research lab already performing experiments using MEPs, there might be a suitable dataset available, since just a few sessions can produce thousands of MEPs. Data from previous studies are useful even if they were recorded on different muscles and using a different setup (e.g., with a different sampling frequency or stimulation paradigm). Moreover, if more MEPs are required, there is no need for the same expert to annotate them. However, DELMEP would benefit from different experts annotating different sections of the dataset, as that would reduce the chance of the algorithm overfitting to biases that could be present in a single expert (e.g., a tendency to under- or over-estimate MEPs latencies). In this regard, it should be noted that DELMEP was trained on MEPs annotated by a single expert. There was a 0.20 ms increase in the MAE (0.50 to 0.70 ms) when comparing the results from testing against the same expert used for the training, and a committee of three independent experts on a different dataset. This increase could be partially caused by having used a single expert to annotate the MEPs in the training dataset. However, this is still accurate enough for the intended application. Moreover, using a single expert to annotate the latencies facilitates and drastically speeds up the process, since only one person has to visually inspect the EMGs recordings. On the other hand, the MAE of Bigoni's method when tested on their own dataset versus on the dataset used in this work increased 0.40 ms (from 0.40 to 0.80 ms), indicating that a variation in accuracy of this magnitude is possible even if the algorithm is not based on ML. As a reference,

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Bigoni *et al.* found a difference of about 0.40 ms when comparing the estimation of two experts [187]. A comparison of DELMEP and Bignoni's mehtod for the Bigoni *et al.* dataset can be found in Figure 6.8. In addition, the low computational cost associated with the DELMEP algorithm allows it to be efficiently used in real-time closed-loop brain stimulation protocols [205][206] and combined with multi-coil TMS electronic targeting for fast and automated cortical mappings [207][208][209], as it requires roughly 1 ms to process an MEP.

In conclusion, a deep learning-based algorithm was developed to annotate MEP latencies automatically without the need for a human expert intervention. The main difference between this algorithm and previously reported solutions is that the deep learning nature of DELMEP allows it to learn and improve based on the available data, making it a candidate for personalized clinical applications. The accuracy of the DELMEP algorithm was practically independent of the amplitude of the MEP, a feature only found in Bigoni's method, as all threshold-based algorithms considered in this study failed this test. It was demonstrated that DELMEP had a high accuracy on two independent datasets. The millisecond-level automated annotation in the proposed DELMEP algorithm opens the possibility for real-time assessment of MEP latencies in closed-loop brain stimulation protocols. The Python implementation of the DELMEP algorithm is currently hosted as a public repository at [210].

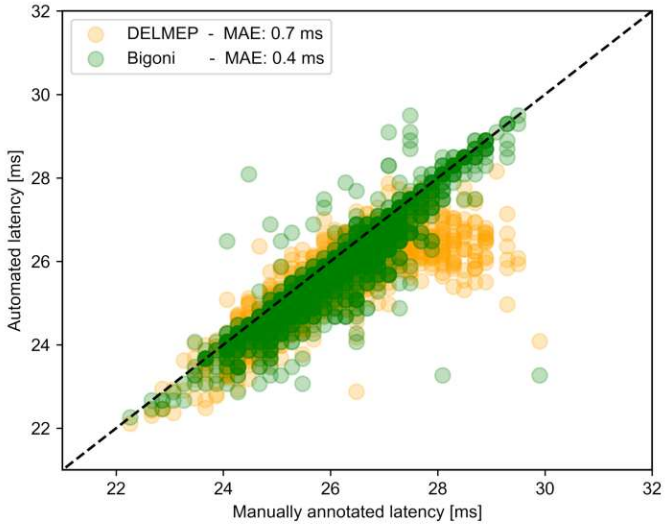

Figure 6.8: Latencies annotated with DELMEP and Bignoni's method vs. manually annotated latencies for the MEPs in Bigoni's dataset. The manually annotated latencies are the mean value of the annotations of three independent experts.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Chapter 7

Summary and Outlook

This thesis addresses the computational costs associated with molecular dynamics (MD) calculations at *ab initio* level accuracy. These calculations are essential in various fields, including drug discovery [211], quantum chemistry [212], and biophysics [213]. The computational costs of traditional methods, such as density functional theory (DFT), are so high that their use is limited to small systems and short simulation times [15]. Consequently, the range of atomic systems and processes that can be studied is severely limited, which motivates the development of alternative solutions. The first attempts to address this issue led to the development of empirical potentials, which are approximate mathematical models of the potential energy surface (PES) tailored to specific applications [40][41][42]. While empirical potentials have reduced computational times, their accuracy is often questionable.

In the past decade, machine learning (ML) interatomic potentials have attracted interest due to their favorable cost-to-accuracy ratio. Since then, they have been rapidly popularized. In this work, four ML interatomic potentials were developed: (I) to model amorphous silicon-nitride [94][95], a material with applications ranging from dielectrics in microelectronic devices to armorial-bearing in NASA's Space Shuttle; (II) to simulate the effects of neutron irradiation in silicon-germanium [115][116], relevant for studying microelectronic components exposed to radiative environments, such as outer space; (III) to investigate the oxidation process of silicon [121][214], crucial for the semiconductor industry; and (IV) to model amorphous silicon-dioxide [129], extensively used as an insulator in microelectronic devices.

These developments were based on the Gaussian approximation potential (GAP) method [55], which reduces the amount of training data required for ML interatomic potentials. As a result, development times were shortened, and computational resources were used more efficiently. The accuracy of these ML interatomic potentials was generally comparable to that of *ab initio* methods, with mean absolute errors (MAEs) in the range of 5-10 meV/atom compared to DFT, while being three to four orders of magnitude faster than this method. This allowed for the study of atomic system sizes and simulation times previously unattainable by orders of magnitude.

During the development of these ML interatomic potentials, the creation of the required datasets emerged as the main bottleneck. This is a common challenge in the development of ML interatomic potentials, arising from the complex and high-dimensional nature of the target PESs. In order to mitigate this issue, the controlled dynamical sampling technique was developed. This simple yet effective active learning technique reduces the number of calculations needed to train ML interatomic potentials. Additionally, a second active learning technique was developed based on selecting atomic configurations using a distance matrix to compare their local atomic environments. However, the testing and publication of this technique fall outside the scope of this thesis and are considered as future work.

The biggest drawback to the ML interatomic potentials' performance was found to be the hardware architecture of the scientific clusters in which they run. These systems are based on the von Neumann architecture, where the memory is separated from the central processing unit (CPU) and connected to it by a slow buffer. This gives rise to the well-known *von Neumann bottleneck* [148], which is particularly notorious in MD simulations, since the positions and forces on thousands of atoms must be shuffled back-and-forth millions of times. In order to mitigate this problem, the use of a neuromorphic hardware accelerator was studied. Following the novel work of Mo *et al.* [149], a neural network based ML interatomic potential for germanium was developed. The application for this interatomic potential was the simulation of germanium-based memristors, with cells composed of over 100,000 atoms. The results indicate that this technique allows for MD simulations about two orders of magnitude faster than the previously employed GAP method, with virtually identical accuracy.

The most common use of ML in the field of MD is in the development of interatomic potentials. However, in this thesis, two alternative uses for this versatile tool were proposed: (I) An ML model to estimate the formation energy of hydrogen-based defects in amorphous silicon-dioxide [135], and (II) an ML model to predict the atomic structure of these defects [146]. The advantage of using these ML models is that neither MD simulations nor geometry optimizations are needed, therefore greatly reducing the computational costs and times associated with these tasks. The design of these algorithms was modular, and they can be used for other defects and/or materials.

Finally, the techniques developed to reduce the computational costs of MD calculations were adapted and employed to the automation of the analysis of biomedical data [176]. This transdisciplinary application consisted in the development of an ML algorithm to automate the evaluation of motor evoked potentials (MEPs) produced by transcranial magnetic stimulation (TMS). MEPs are electromyography (EMG) traces measured with electrodes on the target muscle of a patient, while TMS is a technique that allows to non-invasively stimulate specific areas of the motor cortex in the brain. The combination of these two techniques is widely used in the treatment of neurological disorders, such as multiple sclerosis, and the treatment of a single patient might require thousands of stimulations. Traditionally, a clinician must manually evaluate each of these recordings, making this a slow and error-prone process. The ML algorithm developed in this thesis automates this task, therefore accelerating the work of medical doctors. Moreover, by analyzing the MEPs on-the-fly, this algorithm paves the way to future closed-loop brain stimulation protocols [205][206].

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Bibliography

[1] R. R. Kline, “Cybernetics, automata studies, and the dartmouth conference on artificial intelligence,” *IEEE Annals of the History of Computing*, vol. 33, no. 4, pp. 5–16, 2011. doi: 10.1109/MAHC.2010.44.

[2] F.-H. Hsu, *Behind Deep Blue*. Princeton University Press, 2002. ISBN 0-691-09065-3

[3] J. G. Greener, S. M. Kandathil, L. Moffat, and D. T. Jones, “A guide to machine learning for biologists,” *Nature Reviews Molecular Cell Biology*, vol. 23, pp. 40–55, 2022. doi: 10.1038/s41580-021-00407-0.

[4] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, “Physics-informed machine learning,” *Nature Reviews Physics*, vol. 3, pp. 422–440, 2021. doi: 10.1038/s42254-021-00314-5.

[5] M. Molina and F. Garip, “Machine learning for sociology,” *Annual Review of Sociology*, vol. 45, pp. 27–45, 2019. doi: 10.1146/annurev-soc-073117-041106.

[6] J. I. Glaser, A. S. Benjamin, R. Farhoodi, and K. P. Kording, “The roles of supervised machine learning in systems neuroscience,” *Progress in Neurobiology*, vol. 175, pp. 126–137, 2019. doi: 10.1016/j.pneurobio.2019.01.008.

[7] C. Janiesch, P. Zschech, and K. Heinrich, “Machine learning and deep learning,” *Electronic Markets*, vol. 31, pp. 685–695, 2021. doi: 10.1007/s12525-021-00475-2.

[8] T. Caelli and W. F. Bischof, “Machine learning paradigms for pattern recognition and imagine understanding,” *Spatial Vision*, vol. 10, no. 1, pp. 87–103, 1995. doi: 10.1163/156856896x00079.

[9] P.-L. Tu and J.-Y. Chung, “A new decision-tree classification algorithm for machine learning,” *Proceedings Fourth International Conference on Tools with Artificial Intelligence TAI ’92*, pp. 370–371, 1992. doi: 10.1109/TAI.1992.246431.

[10] D. W. Otter, J. R. Medina, and J. K. Kalita, “A survey of the usages of deep learning for natural language processing,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 2, pp. 604–624, 2021. doi: 10.1109/TNNLS.2020.2979670.

[11] X. Feng, Y. Jiang, X. Yang, M. Du, and X. Li, “Computer vision algorithms and hardware implementations: A survey,” *Integration*, vol. 69, pp. 309–320, 2019. doi: 10.1016/j.vlsi.2019.07.005.

[12] A. Rajkomar, J. Dean, and I. Kohane, “Machine learning in medicine,” *The New England Journal of Medicine*, vol. 380, pp. 1347–1358, 2019. doi: 10.1056/NEJMra1814259.

[13] P. Gogas and T. Papadimitriou, “Machine learning in economics and finance,” *Computational Economics*, vol. 57, pp. 1–4, 2021. doi: 10.1007/s10614-021-10094-w.

[14] S. Zhong *et al.*, “Machine learning: New ideas and tools in environmental science and engineering,” *Environmental Science and Technology*, vol. 55, pp. 12 741–12 754, 2021. doi: 10.1021/acs.est.1c01339.

[15] R. Car and M. Parrinello, “Unified approach for molecular dynamics and density-functional theory,” *Physical Review Letters*, vol. 55, pp. 2471–2474, 1985. doi: 10.1103/PhysRevLett.55.2471.

[16] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, “Development and testing of a general amber force field,” *The Journal of Computational Chemistry*, vol. 25, pp. 1157–1174, 2004. doi: 10.1002/jcc.20035.

[17] J. Behler and M. Parrinello, “Generalized neural-network representation of high-dimensional potential-energy surfaces,” *Physical Review Letters*, vol. 98, p. 146401, 2007. doi: 10.1103/PhysRevLett.98.146401.

[18] J. Behler, “Representing potential energy surfaces by high-dimensional neural network potentials,” *Journal of Physics: Condensed Matter*, vol. 26, p. 183001, 2014. doi: 10.1088/0953-8984/26/18/183001.

[19] V. L. Deringer *et al.*, “Realistic atomistic structure of amorphous silicon from machine-learning-driven molecular dynamics,” *The Journal of Physical Chemistry Letters*, vol. 9, pp. 2879–2885, 2018. doi: 10.1021/acs.jpclett.8b00902.

[20] W. Jia *et al.*, “Pushing the limit of molecular dynamics with ab initio accuracy to 100 million atoms with machine learning,” *SC20: International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. 1–14, 2020. doi: 10.1109/SC41405.2020.00009.

[21] A. G. Kusne *et al.*, “On-the-fly machine-learning for high-throughput experiments: search for rare-earth-free permanent magnets,” *Nature Scientific Reports*, vol. 4, p. 6367, 2014. doi: 10.1038/srep06367.

[22] V. Di Lazzaro *et al.*, “The diagnostic value of motor evoked potentials,” *Clinical Neurophysiology*, vol. 110, pp. 1297–1307, 1999. doi: 10.1016/S1388-2457(99)00060-7.

[23] Institute for Microelectronics, TU Wien, https://www.iue.tuwien.ac.at/home/.

[24] MUNDFAB, https://www.mundfab.eu.

[25] Argonne National Laboratory, https://www.anl.gov.

[26] Laboratoire d’Analyse et d’Architecture des Systèmes (LAAS) - Centre National de la Recherche Scientifique (CNRS), https://www.laas.fr/public/.

[27] Computational Nanoelectronics Group, ETH Zürich, https://nano-tcad.ee.ethz.ch.

[28] Aalto University, Department of Neuroscience and Biomedical Engineering, https://www.aalto.fi/en/department-of-neuroscience-and-biomedical-engineering.

[29] M. Born and R. Oppenheimer, “Zur Quantentheorie der Molekeln,” *Annalen der Physik*, vol. 389, pp. 457–484, 1927. doi: 10.1002/andp.19273892002.

[30] J. A. Steckel and D. S. Sholl, *Density Functional Theory - A Practical Introduction.* John Wiley and Sons, 2009. ISBN 978-0-470-37317-0

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

[31] V. Bapst *et al.*, “Unveiling the predictive power of static structure in glassy systems,” *Nature Physics*, vol. 16, pp. 448–454, 2020. doi: doi.org/10.1038/s41567-020-0842-8.

[32] M. Karplus and G. A. Petsko, “Molecular Dynamics simulations in biology,” *Nature*, vol. 347, pp. 631–639, 1990. doi: 10.1038/347631a0.

[33] Y. Wang *et al.*, “Dynamic deformability of individual PbSe nanocrystals during superlattice phase transitions,” *Science Advances*, vol. 5, 2019. doi: 10.1126/sciadv.aaw5623.

[34] H. J. C. Berendsen, *Simulating the physical world: hierarchical modeling from quantum mechanics to fluid dynamics.* Cambridge University Press, 2007. ISBN 9780521835275

[35] R. C. Tolman, “A general theory of energy partition with applications to quantum theory,” *Physical Review*, vol. 11, pp. 261–275, 1918. doi: 10.1103/PhysRev.11.261.

[36] J. Behler, “Perspective: Machine learning potentials for atomistic simulations,” *The Journal of Chemical Physics*, vol. 145, p. 170901, 2016. doi: 10.1063/1.4966192.

[37] S. Nosé, “A unified formulation of the constant temperature molecular dynamics methods,” *The Journal of Chemical Physics*, vol. 81, pp. 511–519, 1984. doi: 10.1063/1.447334.

[38] W. G. Hoover, “Canonical dynamics: Equilibrium phase-space distributions,” *Physical Review A*, vol. 31, pp. 1695–1697, 1985. doi: 10.1103/PhysRevA.31.1695.

[39] T. Schneider and E. Stoll, “Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions,” *Physical Review B*, vol. 17, pp. 1302–1322, 1978. doi: 10.1103/PhysRevB.17.1302.

[40] J. E. Jones, “On the determination of molecular fields. —II. from the equation of state of a gas,” *Proceedings of the Royal Society A*, vol. 106, pp. 463–477, 1924. doi: 10.1098/rspa.1924.0082.

[41] S. Munetoh, T. Motooka, K. Moriguchi, and A. Shintani, “Interatomic potential for Si-O systems using Tersoff parameterization,” *Computational Materials Science*, vol. 39, pp. 334–339, 2007. doi: 10.1016/j.commatsci.2006.06.010.

[42] T. P. Senftle *et al.*, “The ReaxFF reactive force-field: development, applications and future directions,” *npj Computational Materials*, vol. 2, no. 15011, 2016. doi: 10.1038/npjcompumats.2015.11.

[43] W. D. Cornel *et al.*, “A second generation force field for the simulation of proteins, nucleic acids, and organic molecules,” *Journal of the American Chemical Society*, vol. 117, pp. 5179–5197, 1994. doi: 10.1021/ja00124a002.

[44] W. L. Jorgensen, J. D. Madura, and C. J. Swenson, “Optimized intermolecular potential functions for liquid hydrocarbons,” *Journal of the American Chemical Society*, vol. 106, pp. 6638–6646, 1984. doi: 10.1021/ja00334a030.

[45] S. J. Stuart *et al.*, “A reactive potential for hydrocarbons with intermolecular interactions,” *The Journal of Chemical Physics*, vol. 112, pp. 6472–6486, 2000. doi: 10.1063/1.481208.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[46] Z. Fan, Y. Wang, X. Gu, P. Qian, Y. Su, and T. Ala-Nissila, “A minimal Tersoff potential for diamond silicon with improved descriptions of elastic and phonon transport properties,” *Journal of Physics: Condensed Matter*, vol. 32, no. 13, p. 135901, 2020. doi: 10.1088/1361-648X/ab5c5f.

[47] J. Tersoff, “Empirical interatomic potential for carbon, with applications to amorphous carbon,” *Physical Review Letters*, vol. 61, p. 2879, 1988. doi: 10.1103/PhysRevLett.61.2879.

[48] K. Chenoweth, A. C. T. van Duin, and W. A. Goddard, “Reaxff reactive force field for molecular dynamics simulations of hydrocarbon oxidation,” *The Journal of Physical Chemistry A*, vol. 112, pp. 1040–1053, 2008. doi: 10.1021/jp709896w.

[49] S. Agrawalla and A. C. T. van Duin, “Development and application of a ReaxFF reactive force field for hydrogen combustion,” *The Journal of Physical Chemistry A*, vol. 115, pp. 960–972, 2011. doi: 10.1021/jp108325e.

[50] S. Fadida *et al.*, “Hf-based high-k dielectrics for p-Ge MOS gate stacks,” *Journal of Vacuum Science and Technology B*, vol. 32, no. 3, 2014. doi: 10.1116/1.4837295.

[51] F. Faber, A. Lindmaa, O. A. von Lilienfeld, and R. Armiento, “Crystal structure representations for machine learning models of formation energies,” *International Journal of Quantum Chemistry*, vol. 115, pp. 1094–1101, 2015. doi: 10.1002/qua.24917.

[52] H. Huo and M. Rupp, “Unified representation of molecules and crystals for machine learning,” *Machine Learning: Science and Technology*, vol. 3, no. 4, 2022. doi: 10.48550/arXiv.1704.06439.

[53] K. T. Schütt, H. E. Sauceda, A. Kindermans, P.-J. Tkatchenko, and K.-R. Müller, “Schnet - a deep learning architecture for molecules and materials,” *The Journal of Chemical Physics*, vol. 148, p. 241722, 2018. doi: 10.1063/1.5019779.

[54] A. P. Bartók, R. Kondor, and G. Csányi, “On representing chemical environments,” *Physical Review B*, vol. 87, p. 184115, 2013. doi: 10.1103/PhysRevB.87.184115.

[55] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, “Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons,” *Physical Review Letters*, vol. 104, p. 136403, 2010. doi: 10.1103/PhysRevLett.104.136403.

[56] A. P. Bartók and G. Csányi, “Gaussian approximation potentials: A brief tutorial introduction,” *International Journal of Quantum Chemistry*, vol. 115, pp. 1051–1057, 2015. doi: 10.1002/qua.24927.

[57] V. L. Deringer and G. Csányi, “Machine learning based interatomic potential for amorphous carbon,” *Physical Review B*, vol. 95, p. 094203, 2017. doi: 10.1103/PhysRevB.95.094203.

[58] H. Wang, L. Zhang, J. Han, and W. E, “DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics,” *Computer Physics Communications*, vol. 228, pp. 178–184, 2018. doi: 10.1016/j.cpc.2018.03.016.

[59] R. Lot, F. Pellegrini, Y. Shaidu, and E. Küçükbenli, “Panna: Properties from artificial neural network architectures,” *Computer Physics Communications*, vol. 256, p. 107402, 2020. doi: 10.1016/j.cpc.2020.107402.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

[60] A. Khorshidi and A. A. Peterson, "AMP: A modular approach to machine learning in atomistic simulations," *Computer Physics Communications*, vol. 207, pp. 310–324, 2016. doi: 10.1016/j.cpc.2016.05.010.

[61] What's GAP, https://gap-ml.org/whats-gap/.

[62] Y. Wang, Z. Fan, P. Qian, T. Ala-Nissila, and M. A. Caro, "Structure and pore size distribution in nanoporous carbon," *Chemistry of Materials*, vol. 34, pp. 617–628, 2022. doi: 10.1021/acs.chemmater.1c03279.

[63] J. Kloppenburg, A. Pedersen, K. Laasonen, M. A. Caro, and H. Jónsson, "Reassignment of magic numbers for icosahedral Au clusters: 310, 564, 928 and 1426," *Nanoscale*, vol. 14, pp. 9053–9060, 2022. doi: 10.1039/D2NR01763F.

[64] A. P. Bartók, J. Kermode, N. Bernstein, and G. Csányi, "Machine learning a general-purpose interatomic potential for silicon," *Physical Review X*, vol. 8, p. 041048, 2018. doi: 10.1103/PhysRevX.8.041048.

[65] G. Sivaraman *et al.*, "A combined machine learning and high-energy x-ray diffraction approach to understanding liquid and amorphous metal oxides," *The Journal of the Physical Society of Japan*, vol. 91, p. 091009, 2022. doi: 10.7566/JPSJ.91.091009.

[66] D. Unruh, R. V. Meidanshahi, S. M. Goodnick, G. Csányi, and G. T. Zimányi, "Gaussian approximation potential for amorphous Si : H," *Physical Review Materials*, vol. 6, no. 065603, 2022. doi: 10.1103/PhysRevMaterials.6.065603.

[67] J. Kloppenburg, L. B. Pártay, H. Jónsson, and M. A. Caro, "A general-purpose machine learning Pt interatomic potential for an accurate description of bulk, surfaces and nanoparticles," *The Journal of Chemical Physics*, vol. 158, p. 134704, 2023. doi: 10.1063/5.0143891.

[68] R. Jana and M. A. Caro, "Searching for iron nanoparticles with a general-purpose gaussian approximation potential," *Physical Review B*, vol. 107, p. 245421, 2023. doi: 10.1103/PhysRevB.107.245421.

[69] Available GAPs, https://gap-ml.org/available-gaps/.

[70] T. Mueller, A. Hernandez, and C. Wang, "Machine learning for interatomic potential models," *The Journal of Chemical Physics*, vol. 152, p. 050902, 2020. doi: 10.1063/1.5126336.

[71] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017. [Online]. Available: arXiv:1412.6980

[72] T. W. Ko, J. A. Finkler, S. Goedecker, and J. Behler, "A fourth-generation high-dimensional neural network potential with accurate electrostatics including non-local charge transfer," *Nature Communications*, vol. 12, no. 398, 2021. doi: 10.1038/s41467-020-20427-2.

[73] T. B. Blank, S. D. Brown, A. W. Calhoun, and D. J. Doren, "Neural network models of potential energy surfaces," *The Journal of Chemical Physics*, vol. 103, pp. 4129–4137, 1995. doi: 10.1063/1.469597.

[74] L. M. Raff, M. Malshe, M. Hagan, D. I. Doughan, M. G. Rockley, and R. Komanduri, "Ab initio potential-energy surfaces for complex, multichannel systems using modified novelty sampling and feedforward neural networks," *The Journal of Chemical Physics*, vol. 122, p. 084104, 2005. doi: 10.1063/1.1850458.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[75] K. T. No, B. H. Chang, S. Y. Kim, M. S. Jhon, and H. A. Scheraga, "Description of the potential energy surface of the water dimer with an artificial neural network," *Chemical Physics Letters*, vol. 271, pp. 152–156, 1997. doi: 10.1016/S0009-2614(97)00448-X.

[76] S. Manzhos and T. Carrington, "Using redundant coordinates to represent potential energy surfaces with lower-dimensional functions," *The Journal of Chemical Physics*, vol. 127, p. 014103, 2007. doi: 10.1063/1.2746846.

[77] F. V. Prudente and J. J. Soares Neto, "The fitting of potential energy surfaces using neural networks. application to the study of the photodissociation processes," *Chemical Physics Letters*, vol. 287, pp. 585–589, 1998. doi: 10.1016/S0009-2614(98)00207-3.

[78] M. Malshe *et al.*, "Development of generalized potential-energy surfaces using many-body expansions, neural networks, and moiety energy approximations," *The Journal of Chemical Physics*, vol. 130, p. 184102, 2009. doi: 10.1063/1.3124802.

[79] S. Manzhos, X. Wang, R. Dawes, and T. Carrington, "A nested molecule-independent neural network approach for high-quality potential fits," *The Journal of Physical Chemistry A*, vol. 110, pp. 5295–5304, 2006. doi: 10.1021/jp055253z.

[80] H. Gassner, M. Probst, A. Lauenstein, and K. Hermansson, "Representation of intermolecular potential functions by neural networks," *The Journal of Physical Chemistry A*, vol. 102, pp. 4596–4605, 1998. doi: 10.1021/jp972209d.

[81] H. Eshet, R. Z. Khaliullin, T. D. Kühne, J. Behler, and M. Parrinello, "Ab initio quality neural-network potential for sodium," *Physical Review B*, vol. 81, p. 184107, 2010. doi: 10.1103/PhysRevB.81.184107.

[82] R. Z. Khaliullin, H. Eshet, T. D. Kühne, J. Behler, and M. Parrinello, "Graphite-diamond phase coexistence study employing a neural-network mapping of the ab initio potential energy surface," *Physical Review B*, vol. 81, p. 100103, 2010. doi: 10.1103/PhysRevB.81.100103.

[83] J. Behler, R. Martoňák, D. Donadio, and M. Parrinello, "Metadynamics simulations of the high-pressure phases of silicon employing a high-dimensional neural network potential," *Physical Review Letters*, vol. 100, p. 185501, 2008. doi: 10.1103/PhysRevLett.100.185501.

[84] J. Behler, R. Martoňák, D. Donadio, and M. Parrinello, "Pressure-induced phase transitions in silicon studied by neural network-based metadynamics simulations," *Physica Status Solidi B*, vol. 245, pp. 2618–2629, 2008. doi: 10.1002/pssb.200844219.

[85] N. Artrith, T. Morawietz, and J. Behler, "High-dimensional neural-network potentials for multicomponent systems: Applications to zinc oxide," *Physical Review B*, vol. 83, p. 153101, 2011. doi: 10.1103/PhysRevB.83.153101.

[86] T. Morawietz, V. Sharma, and J. Behler, "A neural network potential-energy surface for the water dimer based on environment-dependent atomic energies and charges," *The Journal of Chemical Physics*, vol. 136, p. 064103, 2012. doi: 10.1063/1.3682557.

[87] E. Kocer, T. W. Ko, and J. Behler, "Neural network potentials: A concise overview of methods," *Annual Review of Physical Chemistry*, vol. 73, pp. 163–186, 2022. doi: 10.1146/annurev-physchem-082720- 034254.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

[88] D. Duvenaud *et al.*, “Convolutional networks on graphs for learning molecular fingerprints,” in *Proceedings of the 28th International Conference on Neural Information Processing Systems*, 2015, pp. 2224–2232.

[89] D. Rogers and M. Hahn, “Extended-connectivity fingerprints,” *Journal of Chemical Information and Modeling*, vol. 50, pp. 742–754, 2010. doi: 10.1021/ci100050t.

[90] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, “Neural message passing for quantum chemistry,” in *Proceedings of the 34th International Conference on Machine Learning, PMLR*, 2017, pp. 1263–1272.

[91] Z. Hu *et al.*, “Aisnet: A universal interatomic potential neural network with encoded local environment features,” *Journal of Chemical Information and Modeling*, vol. 36, pp. 1756–1765, 2023. doi: 10.1021/acs.jcim.3c00077.

[92] K. Lee, D. Yoo, W. Jeong, and S. Han, “SIMPLE-NN: An efficient package for training and executing neural-network interatomic potentials,” *Computer Physics Communications*, vol. 242, pp. 95–103, 2019. doi: 10.1016/j.cpc.2019.04.014.

[93] A. V. Shapeev, “Moment tensor potentials: A class of systematically improvable interatomic potentials,” *Multiscale Modeling and Simulation*, vol. 14, pp. 1153–1173, 2016. doi: 10.1137/15M1054183.

[94] D. Milardovich, C. Wilhelmer, D. Waldhoer, L. Cvitkovich, G. Sivaraman, and T. Grasser, “Machine learning interatomic potential for silicon-nitride (Si_3N_4) by active learning,” *The Journal of Chemical Physics*, vol. 158, no. 19, p. 194802, 2023. doi: 10.1063/5.0146753.

[95] C. Wilhelmer, D. Waldhoer, L. Cvitkovich, D. Milardovich, M. Waltl, and T. Grasser, “Over- and undercoordinated atoms as a source of electron and hole traps in amorphous silicon nitride (a-Si_3N_4),” *Nanomaterials*, vol. 13, p. 2286, 2023. doi: 10.3390/nano13162286.

[96] J. H. You, H. W. Kim, D. H. Kim, T. W. Kim, and K. W. Lee, “Effect of the trap density and distribution of the silicon nitride layer on the retention characteristics of charge trap flash memory devices,” *2011 International Conference on Simulation of Semiconductor Processes and Devices, Osaka, Japan*, pp. 199–202, 2011. doi: 10.1109/SISPAD.2011.6035085.

[97] F. L. Riley, “Silicon nitride and related materials,” *Journal of the American Ceramic Society*, vol. 83, no. 2, pp. 245–265, 2004. doi: 10.1111/j.1151-2916.2000.tb01182.x.

[98] R. Sharma, “Silicon nitride as antireflection coating to enhance the conversion efficiency of silicon solar cells,” *Turkish Journal of Physics*, vol. 42, no. 4, pp. 350–355, 2018. doi: 10.3906/fiz-1801-28.

[99] NASA, “Space shuttle main engine enhancements. (accessed: 09.12.2022).” [Online]. Available: https://www.nasa.gov/centers/marshall/news/background/facts/ssme.html

[100] H. O. Pierson, *Handbook of chemical vapor deposition (CVD).* Noyes Publications, 1992. ISBN 978-0-8155-1300-1

[101] J. W. Mellor, “A comprehensive treatise on inorganic and theoretical chemistry,” *Nature*, vol. 124, no. 3133, p. 757, 1929. doi: 10.1038/124757c0.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

[102] A. López-Suárez *et al.*, "Modification of the nonlinear optical absorption and optical kerr response exhibited by nc-Si embedded in a silicon-nitride film," *Optics Express*, vol. 17, pp. 10 056–10 068, 2009. doi: 10.1364/OE.17.010056.

[103] F. de Brito Mota, J. F. Justo, and A. Fazzio, "Hydrogen role on the properties of amorphous silicon nitride," *Journal of Applied Physics*, vol. 86, pp. 1843–1847, 1999. doi: 10.1063/1.370977.

[104] S. R. Billeter, A. Curioni, D. Fischer, and W. Andreoni, "Ab initio derived augmented Tersoff potential for silicon oxynitride compounds and their interfaces with silicon," *Physical Review B*, vol. 73, p. 155329, 2006. doi: 10.1103/PhysRevB.73.155329.

[105] R. P. Vedula, N. L. Anderson, and A. Strachan, "Effect of topological disorder on structural, mechanical, and electronic properties of amorphous silicon nitride: An atomistic study," *Physical Review B*, vol. 85, p. 205209, 2012. doi: 10.1103/PhysRevB.85.205209.

[106] O. Böhm, "SiCNH parameter set," 2015. [Online]. Available: https://docs.quantumatk.com/manual/ForceField.html#reaxff-sicnh-c-n-si-h

[107] L. C. Erhard, J. Rohrer, K. Albe, and V. L. Deringer, "A machine-learned interatomic potential for silica and its relation to empirical models," *npj Computational Materials*, vol. 8, no. 90, 2022. doi: 10.1038/s41524-022-00768-w.

[108] J. R. Kermode, "f90wrap: an automated tool for constructing deep Python interfaces to modern Fortran codes," *Journal of Physics: Condensed Matter*, vol. 32, no. 30, p. 305901, 2020. doi: 10.1088/1361-648X/ab82d2.

[109] Gaussian Approximation Potential for Silicon-Nitride, https://github.com/dmilardovich/GAP_Si3N4.

[110] S. Smidstrup *et al.*, "QuantumATK: An integrated platform of electronic and atomic-scale modelling tools," *Journal of Physics: Condensed Matter*, vol. 32, no. 1, p. 015901, 2020. doi: 10.1088/1361-648X/ab4007.

[111] J. Hutter, M. Iannuzzi, F. Schiffmann, and J. VandeVondele, "CP2K: Atomistic simulations of condensed matter systems," *WIREs Computational Molecular Science*, vol. 4, pp. 15–25, 2013. doi: 10.1002/wcms.1159.

[112] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Physical Review Letters*, vol. 77, pp. 3865–3868, 1996. doi: 10.1103/PhysRevLett.77.3865.

[113] S. Goedecker, M. Teter, and J. Hutter, "Separable dual-space gaussian pseudopotentials," *Physical Review B*, vol. 54, pp. 1703–1710, 1996. doi: 10.1103/PhysRevB.54.1703.

[114] T. Aiyama, T. Fukunaga, K. Niihara, T. Hirai, and K. Suzuki, "An x-ray diffraction study of the amorphous structure of chemically vapor-deposited silicon nitride," *Journal of Non-Crystalline Solids*, vol. 33, pp. 131–139, 1979. doi: 10.1016/0022-3093(79)90043-7.

[115] D. Milardovich *et al.*, "Machine learning interatomic potential for silicon-germanium," November 2022, poster presented at Toulousaines du Calcul Atomique et Moléculaire (TouCAM).

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[116] D. Milardovich *et al.*, "General purpose machine learning interatomic potential for silicon-germanium," *Spring Meeting, European Materials Research Society*, 2023.

[117] J. R. Srour and J. W. Palko, "Displacement damage effects in irradiated semiconductor devices," *IEEE Transactions on Nuclear Science*, vol. 60, pp. 1740–1766, 2013. doi: 10.1109/TNS.2013.2261316.

[118] T. Jarrin, A. Jay, A. Hémeryck, and N. Richard, "Parametric study of the two-temperature model for molecular dynamics simulations of collisions cascades in Si and Ge," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 485, pp. 1–9, 2020. doi: 10.1016/j.nimb.2020.09.025.

[119] F. H. Stillinger and T. A. Weber, "Computer simulation of local order in condensed phases of silicon," *Physical Review B*, vol. 31, pp. 5262–5271, 1985. doi: 10.1103/PhysRevB.31.5262.

[120] A.-M. El-Sayed, M. B. Watkins, A. L. Shluger, and V. V. Afanas'ev, "Identification of intrinsic electron trapping sites in bulk amorphous silica from ab initio calculations," *Microelectronic Engineering*, vol. 109, pp. 68–71, 2013. doi: 10.1016/j.mee.2013.03.027.

[121] L. Cvitkovich *et al.*, "Multi-scale modeling of thermal oxidation of Si," *The 14th international conference on SiO2, advanced dielectrics and related devices*, 2023.

[122] V. Y. Vasilyev, "Review-atomic layer deposition of silicon dioxide thin films," *ECS Journal of Solid State Science and Technology*, vol. 10, no. 5, p. 053004, 2021. doi: 10.1149/2162-8777/abffab.

[123] F. Palumbo *et al.*, "A review on dielectric breakdown in thin dielectrics: Silicon dioxide, high-k, and layered dielectrics," *Advanced Functional Materials*, vol. 30, p. 1900657, 2019. doi: 10.1002/adfm.201900657.

[124] K. J. Åström and T. Hägglund, *PID Controllers: Theory, Design, and Tuning.* Research Triangle Park, NC: Instrumentation Society of America, 1995. ISBN 1-55617-516-7

[125] T. Mikolajick *et al.*, "Silicon nanowires - a versatile technology platform," *Physica Status Solidi RRL*, vol. 7, pp. 793–799, 2013. doi: 10.1002/pssr.201307247.

[126] K.-Q. Peng, X. Wang, L. Li, Y. Hu, and S.-T. Lee, "Silicon nanowires for advanced energy conversion and storage," *Nano Today*, vol. 8, pp. 75–97, 2013. doi: 10.1016/j.nantod.2012.12.009.

[127] M. Akbari-Saatlu *et al.*, "Silicon nanowires for gas sensing: A review," *Nanomaterials*, vol. 10, p. 2215, 2020. doi: 10.3390/nano10112215.

[128] A. Gao, S. Chen, Y. Wang, and T. Li, "Silicon nanowire field-effect-transistor-based biosensor for biomedical applications," *Sensors and Materials*, vol. 30, pp. 1619–1628, 2018. doi: 10.18494/SAM.2018.1829.

[129] D. Milardovich, D. Waldhoer, M. Jech, A.-M. El-Sayed, and T. Grasser, "Building robust machine learning force fields by composite gaussian approximation potentials," *Solid-State Electronics*, vol. 200, p. 108529, 2022. doi: 10.1016/j.sse.2022.108529.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[130] F. C. Mocanu *et al.*, "Modeling the phase-change memory material, $Ge_2Sb_2Te_5$, with a machine-learned interatomic potential," *Journal of Physical Chemistry B*, vol. 122, pp. 8998–9006, 2018. doi: 10.1021/acs.jpcb.8b06476.

[131] G. Csányi *et al.*, "Expressive programming for computational physics in fortran 95+," *IoP Computational Physics Newsletter*, 2007. [Online]. Available: https://github.com/libAtoms/QUIP

[132] A. P. Thompson *et al.*, "LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales," *Computer Physics Communications*, vol. 271, no. 108171, 2022. doi: 10.1016/j.cpc.2021.108171.

[133] A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard, "ReaxFF: A reactive force field for hydrocarbons," *Journal of Physical Chemistry A*, vol. 105, pp. 9396–9409, 2001. doi: 10.1021/jp004368u.

[134] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, "Phonons and related crystal properties from density-functional perturbation theory," *Reviews of Modern Physics*, vol. 73, pp. 515–562, 2001. doi: 10.1103/RevModPhys.73.515.

[135] D. Milardovich, M. Jech, D. Waldhoer, M. Waltl, and T. Grasser, "Machine learning prediction of defect formation energies in a-SiO2," *2020 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 339–342, 2020. doi: 10.23919/SISPAD49475.2020.9241609.

[136] A.-M. El-Sayed, M. B. Watkins, T. Grasser, V. V. Afanas'ev, and A. L. Shluger, "Hydrogen-induced rupture of strained Si-O bonds in amorphous silicon dioxide," *Physical Review Letters*, vol. 114, p. 115503, 2015. doi: 10.1103/PhysRevLett.114.115503.

[137] T. Grasser *et al.*, "On the microscopic structure of hole traps in pMOSFETs," *2014 IEEE International Electron Devices Meeting*, pp. 21.1.1–21.1.4, 2014. doi: 10.1109/IEDM.2014.7047093.

[138] Y. Wimmer, A.-M. El-Sayed, W. Goes, T. Grasser, and A. L. Shluger, "Role of hydrogen in volatile behaviour of defects in SiO_2-based electronic devices," *Proceedings of the Royal Society A*, vol. 472, p. 20160009, 2016. doi: 10.1098/rspa.2016.0009.

[139] W. Goes *et al.*, "Identification of oxide defects in semiconductor devices: A systematic approach linking DFT to rate equations and experimental evidence," *Microelectronics Reliability*, vol. 87, pp. 286–320, 2018. doi: 10.1016/j.microrel.2017.12.021.

[140] A. La Magna *et al.*, "Simulation of the growth kinetics in group IV compound semiconductors," *Physica Status Solidi A*, vol. 216, no. 10, p. 1800597, 2018. doi: 10.1002/pssa.201800597.

[141] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011. doi: 10.48550/arXiv.1201.0490.

[142] J. Behler, "Atom-centered symmetry functions for constructing high-dimensional neural network potentials," *Journal of Chemical Physics*, vol. 134, no. 7, p. 074106, 2011. doi: 10.1063/1.3553717.

[143] L. Himanen *et al.*, "Dscribe: Library of descriptors for machine learning in materials science," *Computer Physics Communications*, vol. 247, p. 106949, 2019. doi: 10.1016/j.cpc.2019.106949.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

[144] S. Kiyohara, H. Oda, T. Miyata, and T. Mizoguchi, "Prediction of interface structures and energies via virtual screening," *Science Advances*, vol. 2, p. e1600746, 2016. doi: 10.1126/sciadv.1600746.

[145] A. Seko, H. Hayashi, K. Nakayama, A. Takahashi, and I. Tanaka, "Representation of compounds for machine-learning prediction of physical properties," *Physical Review B*, vol. 95, p. 144110, 2017. doi: 10.1103/PhysRevB.95.144110.

[146] D. Milardovich, M. Jech, D. Waldhoer, A.-M. El-Sayed, and T. Grasser, "Machine learning prediction of defect structures in amorphous silicon dioxide," *ESSDERC 2021 - IEEE 51st European Solid-State Device Research Conference (ESSDERC)*, pp. 239–242, 2021. doi: 10.1109/ESSDERC53440.2021.9631837.

[147] P. Virtanen *et al.*, "Scipy 1.0: fundamental algorithms for scientific computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020. doi: 10.1038/s41592-020-0772-5.

[148] J. Backus, "Can programming be liberated from the von Neumann style? A functional style and its algebra of programs," *Communications of the ACM*, vol. 21, no. 8, pp. 613–641, 1978. doi: 10.1145/359576.359579.

[149] P. Mo *et al.*, "Accurate and efficient molecular dynamics based on machine learning and non von Neumann architecture," *npj Computational Materials*, vol. 8, no. 107, 2022. doi: 10.1038/s41524-022-00773-z.

[150] H. Hendy and C. Merkel, "Review of spike-based neuromorphic computing for brain-inspired vision: Biology, algorithms, and hardware," *Journal of Electronic Imaging*, vol. 31, p. 010901, 2022. doi: 10.1117/1.JEI.31.1.010901.

[151] D. Marković, A. Mizrahi, D. Querlioz, and J. Grollie, "Physics for neuromorphic computing," *Nature Reviews Physics*, vol. 2, pp. 499–510, 2020. doi: 10.1038/s42254-020-0208-2.

[152] S. S. Bulusu and S. Vasudevan, "FPGA accelerator for machine learning interatomic potential-based molecular dynamics of gold nanoparticles," *IEEE Access*, vol. 10, pp. 40 338–40 347, 2022. doi: 10.1109/ACCESS.2022.3165650.

[153] C. Mead, "Neuromorphic electronic systems," *Proceedings of the IEEE*, vol. 78, pp. 1629–1636, 1990. doi: 10.1109/5.58356.

[154] C. Mead, "How we created neuromorphic engineering," *Nature Electronics*, vol. 3, pp. 434–435, 2020. doi: 10.1038/s41928-020-0448-2.

[155] B. V. Benjamin *et al.*, "Neurogrid: A mixed-analog-digital multichip system for large-scale neural simulations," *Proceedings of the IEEE*, vol. 102, pp. 699–716, 2014. doi: 10.1109/JPROC.2014.2313565.

[156] J. Schemmel, D. Brüderle, A. Grübl, M. Hock, K. Meier, and S. Millner, "A wafer-scale neuromorphic hardware system for large-scale neural modeling," *2010 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1947–1950, 2010. doi: 10.1109/ISCAS.2010.5536970.

[157] S. Moradi, N. Qiao, F. Stefanini, and G. Indiveri, "A scalable multicore architecture with heterogeneous memory structures for dynamic neuromorphic asynchronous processors (dynaps)," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 12, pp. 106–122, 2018. doi: 10.1109/TBCAS.2017.2759700.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[158] C. S. Thakur *et al.*, “Large-scale neuromorphic spiking array processors: A quest to mimic the brain,” *Frontiers in Neuroscience*, vol. 12, no. 891, 2018. doi: 10.3389/fnins.2018.00891.

[159] M. Davies *et al.*, “Loihi: A neuromorphic manycore processor with on-chip learning,” *IEEE Micro*, vol. 38, pp. 82–99, 2018. doi: 10.1109/MM.2018.112130359.

[160] P. A. Merolla *et al.*, “Artificial brains. a million spiking-neuron integrated circuit with a scalable communication network and interface,” *Science*, vol. 345, pp. 668–673, 2014. doi: 10.1126/science.1254642.

[161] T. Yokouchi *et al.*, “Pattern recognition with neuromorphic computing using magnetic field-induced dynamics of skyrmions,” *Science Advances*, vol. 8, no. 39, p. eabq5652, 2022. doi: 10.1126/sciadv.abq5652.

[162] M. Zeng, Y. He, C. Zhang, and Q. Wan, “Neuromorphic devices for bionic sensing and perception,” *Frontiers in Neuroscience*, vol. 15, no. 690950, 2021. doi: 10.3389/fnins.2021.690950.

[163] Y. Qi, J. Chen, and Y. Wang, “Neuromorphic computing facilitates deep brain-machine fusion for high-performance neuroprosthesis,” *Frontiers in Neuroscience*, vol. 17, no. 1153985, 2023. doi: 10.3389/fnins.2023.1153985.

[164] C. D. Schuman, S. R. Kulkarni, M. Parsa, J. P. Mitchell, P. Date, and B. Kay, “Opportunities for neuromorphic computing algorithms and applications,” *Nature Computational Science*, vol. 2, pp. 10–19, 2022. doi: 10.1038/s43588-021-00184-y.

[165] S. Furber, “Large-scale neuromorphic computing systems,” *Journal of Neural Engineering*, vol. 13, no. 5, p. 051001, 2016. doi: 10.1088/1741-2560/13/5/051001.

[166] N. Rathi *et al.*, “Exploring neuromorphic computing based on spiking neural networks: Algorithms to hardware,” *ACM Computing Surveys*, vol. 55, no. 243, pp. 1–49, 2023. doi: 10.1145/3571155.

[167] J. B. Aimone *et al.*, “A review of non-cognitive applications for neuromorphic computing,” *Neuromorphic Computing and Engineering*, vol. 2, no. 3, p. 032003, 2022. doi: 10.1088/2634-4386/ac889c.

[168] L. Chua, “Memristor-the missing circuit element,” *IEEE Transactions on Circuit Theory*, vol. 18, pp. 507–519, 1971. doi: 10.1109/TCT.1971.1083337.

[169] J. J. Yang *et al.*, “Memristive switching mechanism for metal/oxide/metal nanodevices,” *Nature Nanotechnology*, vol. 3, pp. 429–433, 2008. doi: 10.1038/nnano.2008.160.

[170] C. Ye *et al.*, “Physical mechanism and performance factors of metal oxide based resistive switching memory: A review,” *Journal of Materials Science and Technology*, vol. 32, pp. 1–11, 2016. doi: 10.1016/j.jmst.2015.10.018.

[171] M. Le Gallo and A. Sebastian, “An overview of phase-change memory device physics,” *Journal of Physics D: Applied Physics*, vol. 53, no. 21, p. 213002, 2020. doi: 10.1088/1361-6463/ab7794.

[172] S. López-Soriano, J. M. Purushothama, A. Vena, and E. Perret, “CBRAM technology: transition from a memory cell to a programmable and non-volatile impedance for new radiofrequency applications,” *Scientific Reports*, vol. 12, no. 4105, 2022. doi: 10.1038/s41598-022-08127-x.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

[173] T. Mikolajick *et al.*, “FeRAM technology for high density applications,” *Microelectronics Reliability*, vol. 41, pp. 947–950, 2001. doi: 10.1016/S0026-2714(01)00049-X.

[174] NVNMD GitHub, https://github.com/LiuGroupHNU/nvnmd-example.

[175] NVNMD website, http://nvnmd.picp.vip.

[176] D. Milardovich *et al.*, “DELMEP: A deep learning algorithm for automated annotation of motor evoked potential latencies,” *Nature Scientific Reports*, vol. 13, no. 8225, 2023. doi: https://doi.org/10.1038/s41598-023-34801-9.

[177] L. Zhang, J. Han, H. Wang, W. Saidi, R. Car, and W. E, “End-to-end symmetry preserving inter-atomic potential energy model for finite and extended systems,” *32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montreal, Canada*, 2018.

[178] R. G. Emerson, “Evoked potentials in clinical trials for multiple sclerosis,” *Journal of Clinical Neurophysiology*, vol. 15, pp. 109–116, 1998. doi: 10.1097/00004691-199803000-00003.

[179] K. E. Brown *et al.*, “The reliability of commonly used electrophysiology measures,” *Brain Stimulation*, vol. 10, pp. 1102–1111, 2017. doi: 10.1016/j.brs.2017.07.011.

[180] R. A. Macdonell, G. A. Donnan, and P. F. Bladin, “A comparison of somatosensory evoked and motor evoked potentials in stroke,” *Annals of Neurology*, vol. 25, pp. 68–73, 1989. doi: 10.1002/ana.410250111.

[181] S. C. Livingston and C. D. Ingersoll, “Intra-rater reliability of a transcranial magnetic stimulation technique to obtain motor evoked potentials,” *International Journal of Neuroscience*, vol. 118, pp. 239–256, 2008. doi: 10.1080/00207450701668020.

[182] S. M. Krieg *et al.*, “Protocol for motor and language mapping by navigated TMS in patients and healthy volunteers; workshop report,” *Acta Neurochirurgica*, vol. 159, pp. 1187–1195, 2017. doi: 10.1007/s00701-017-3187-z.

[183] S. R. Giridharan, D. Gupta, A. Pal, A. M. Mishra, N. J. Hill, and J. B. Carmel, “Motometrics: A toolbox for annotation and efficient analysis of motor evoked potentials,” *Frontiers in Neuroinformatics*, vol. 13, 2019. doi: 10.3389/fninf.2019.00008.

[184] S. Harquel, L. Beynel, N. Guyader, O. Marendaz, C. David, and A. Chauvin, “CortExTool: a toolbox for processing motor cortical excitability measurements by transcranial magnetic stimulation,” 2016. [Online]. Available: https://hal.science/hal-01390016

[185] Signal Hunter algorithm, https://github.com/biomaglab/signalhunter.

[186] J. Šoda, M. R. Vidaković, J. Lorincz, A. Jerković, and I. Vujović, “A novel latency estimation algorithm of motor evoked potential signals,” *IEEE Access*, vol. 8, pp. 193 356–193 374, 2020. doi: 10.1109/ACCESS.2020.3033075.

[187] C. Bigoni, A. Cadic-Melchior, P. Vassiliadis, T. Morishita, and F. C. Hummel, “An automatized method to determine latencies of motor-evoked potentials under physiological and pathophysiological conditions,” *Journal of Neural Engineering*, vol. 19, no. 2, p. 024002, 2022. doi: 10.1088/1741-2552/ac636c.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[188] N. Sollmann, L. Bulubas, N. Tanigawa, C. Zimmer, B. Meyer, and S. M. Krieg, “The variability of motor evoked potential latencies in neurosurgical motor mapping by preoperative navigated transcranial magnetic stimulation,” *BMC Neuroscience*, vol. 18, no. 5, 2017. doi: 10.1186/s12868-016-0321-4.

[189] L. Kiers, D. Cros, K. H. Chiappa, and J. Fang, “Variability of motor potentials evoked by transcranial magnetic stimulation,” *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, vol. 89, pp. 415–423, 1993. doi: 10.1016/0168-5597(93)90115-6.

[190] E. M. Wassermann, “Variation in the response to transcranial magnetic brain stimulation in the general population,” *Clinical Neurophysiology*, vol. 113, pp. 1165–1171, 2002. doi: 10.1016/S1388-2457(02)00144-X.

[191] T. Picht *et al.*, “Assessing the functional status of the motor system in brain tumor patients using transcranial magnetic stimulation,” *Acta Neurochirurgica*, vol. 154, pp. 2075–2081, 2012. doi: 10.1007/s00701-012-1494-y.

[192] J. Schmidhuber, “Deep learning in neural networks: An overview,” *Neural Networks*, vol. 61, pp. 85–117, 2014. doi: 10.1016/j.neunet.2014.09.003.

[193] V. H. Souza, J. O. Nieminen, S. Tugin, L. M. Koponen, O. Baffa, and R. J. Ilmoniemi, “TMS with fast and accurate electronic control: Measuring the orientation sensitivity of corticomotor pathways,” *Brain Stimulation*, vol. 15, pp. 306–315, 2022. doi: 10.1016/j.brs.2022.01.009.

[194] V. H. Souza, J. O. Nieminen, S. Tugin, L. Koponen, O. Baffa, and R. J. Ilmoniemi, “Probing the orientation specificity of excitatory and inhibitory circuitries in the primary motor cortex with multi-channel TMS,” *bioRxiv*, 2021. doi: 10.1101/2021.08.20.457101.

[195] L. M. Koponen, J. O. Nieminen, and R. J. Ilmoniemi, “Multi-locus transcranial magnetic stimulation—theory and implementation,” *Brain Stimulation*, vol. 11, pp. 849–855, 2018. doi: 10.1016/j.brs.2018.03.014.

[196] S. G. Makridakis, S. C. Wheelwright, and R. J. Hyndman, *Forecasting: Methods and Applications*. Wiley, 1998.

[197] F. Chollet *et al.*, “Keras,” 2015. [Online]. Available: https://github.com/fchollet/keras

[198] U. Ziemann, J. C. Rothwell, and M. C. Ridding, “Interaction between intracortical inhibition and facilitation in human motor cortex,” *The Journal of Physiology*, vol. 496, pp. 873–881, 1996. doi: 10.1113/jphysiol.1996.sp021734.

[199] T. V. Ilić, F. Meintzschel, U. Cleff, D. Ruge, K. R. Kessler, and U. Ziemann, “Short-interval paired-pulse inhibition and facilitation of human motor cortex: the dimension of stimulus intensity,” *The Journal of Physiology*, vol. 545, pp. 153–167, 2002. doi: 10.1113/jphysiol.2002.030122.

[200] E. M. Wassermann, A. V. Peterchev, U. Ziemann, S. H. Lisanby, H. R. Siebner, and V. Walsh, *Oxford Handbook of Transcranial Stimulation*. Oxford University Press, 2021.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[201] P. M. Rossini *et al.*, “Non-invasive electrical and magnetic stimulation of the brain, spinal cord, roots and peripheral nerves: Basic principles and procedures for routine clinical and research application. an updated report from an I.F.C.N. Committee,” *Clinical Neurophysiology*, vol. 126, pp. 1071–1107, 2015. doi: 10.1016/j.clinph.2015.02.001.

[202] P. Ren *et al.*, “A survey of deep active learning,” *ACM Computing Surveys*, vol. 54, no. 9, pp. 1–40, 2021. doi: 10.1145/3472291.

[203] Y. Shen, H. Yun, Z. C. Lipton, Y. Kronrod, and A. Anandkumar, “Deep active learning for named entity recognition,” in *Proceedings of the 2nd Workshop on Representation Learning for NLP*. Association for Computational Linguistics, 2017, pp. 252–256. doi: 10.18653/v1/W17-2630.

[204] L. Zhang, D.-Y. Lin, H. Wang, R. Car, and W. E, “Active learning of uniformly accurate interatomic potentials for materials simulation,” *Physical Review Materials*, vol. 3, p. 023804, 2019. doi: 10.1103/PhysRevMaterials.3.023804.

[205] B. Zrenner *et al.*, “Brain oscillation-synchronized stimulation of the left dorsolateral prefrontal cortex in depression using real-time EEG-triggered TMS,” *Brain Stimulation*, vol. 13, pp. 197–205, 2019. doi: 10.1016/j.brs.2019.10.007.

[206] C. Zrenner, D. Desideri, P. Belardinelli, and U. Ziemann, “Real-time EEG-defined excitability states determine efficacy of TMS-induced plasticity in human motor cortex,” *Brain Stimulation*, vol. 11, pp. 374–389, 2018. doi: 10.1016/j.brs.2017.11.016.

[207] A. E. Tervo, J. Metsomaa, J. O. Nieminen, J. Sarvas, and R. J. Ilmoniemi, “Automated search of stimulation targets with closed-loop transcranial magnetic stimulation,” *NeuroImage*, vol. 220, p. 117082, 2020. doi: 10.1016/j.neuroimage.2020.117082.

[208] A. E. Tervo *et al.*, “Closed-loop optimization of transcranial magnetic stimulation with electroencephalography feedback,” *Brain Stimulation*, vol. 15, pp. 523–531, 2022. doi: 10.1016/j.brs.2022.01.016.

[209] J. O. Nieminen *et al.*, “Multi-locus transcranial magnetic stimulation system for electronically targeted brain stimulation,” *Brain Stimulation*, vol. 15, pp. 116–124, 2022. doi: 10.1016/j.brs.2021.11.014.

[210] D. Milardovich *et al.*, “DELMEP: A deep learning algorithm for automated annotation of motor evoked potential latencies.” [Online]. Available: https://github.com/connect2brain/delmep

[211] O. Salo-Ahen *et al.*, “Molecular dynamics simulations in drug discovery and pharmaceutical development,” *MDPI Processes*, vol. 9, no. 71, 2021. doi: 10.3390/pr9010071.

[212] A. D. Boese, A. Chandra, J. M. L. Martin, and D. Marx, “From ab initio quantum chemistry to molecular dynamics: The delicate case of hydrogen bonding in ammonia,” *The Journal of Chemical Physics*, vol. 119, pp. 5965–5980, 2003. doi: 10.1063/1.1599338.

[213] J. Jung *et al.*, “Scaling molecular dynamics beyond 100,000 processor cores for large-scale biophysical simulations,” *The Journal of Computational Chemistry*, vol. 40, pp. 1919–1930, 2019. doi: 10.1002/jcc.25840.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[214] L. Cvitkovich, D. Waldhoer, A.-M. El-Sayed, M. Jech, C. Wilhelmer, and T. Grasser, "Dynamic modeling of Si(100) thermal oxidation: Oxidation mechanisms and realistic amorphous interface generation," *Applied Surface Science*, vol. 610, no. 155378, 2023. doi: 10.1016/j.apsusc.2022.155378.

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

List of Publications

Journal Articles

[1] C. Wilhelmer, D. Waldhoer, L. Cvitkovich, **D. Milardovich**, M. Waltl, T. Grasser, "Over- and Undercoordinated Atoms as a Source of Electron and Hole Traps in Amorphous Silicon Nitride (a-Si3N4)", *Nanomaterials*, vol. 13, 2286, 2023, doi: 10.3390/nano13162286.

[2] **D. Milardovich**, V. H. Souza, I. Zubarev, S. Tugin, J. O. Nieminen, C. Bigoni, F. C. Hummel, J. T. Korhonen, D. B. Aydogan, P. Lioumis, N. Taherinejad, T. Grasser, R. J. Ilmoniemi, "DELMEP: A deep learning algorithm for automated annotation of motor evoked potential latencies", *Nature Scientific Reports*, vol. 13, 8225, 2023, doi: 10.1038/s41598-023-34801-9.

[3] **D. Milardovich**, C. Wilhelmer, D. Waldhoer, L. Cvitkovich, G. Sivaraman, T. Grasser, "Machine learning interatomic potential for silicon-nitride (Si3N4) by active learning", *The Journal of Chemical Physics*, vol. 158, 194802, 2023, doi: 10.1063/5.0146753.

[4] **D. Milardovich**, D. Waldhoer, M. Jech, A.-M. El-Sayed, T. Grasser, "Building robust machine learning force fields by composite Gaussian approximation potentials", *Solid-State Electronics*, vol. 200, 108529, 2023, doi: 10.1016/j.sse.2022.108529.

[5] J. O. Nieminen, H. Sinisalo, V. H. Souza, M. Malmi, M. Yuryev, A. E. Tervo, M. Stenroos, **D. Milardovich**, J. T. Korhonen, L. M. Koponen, "Multi-locus transcranial magnetic stimulation system for electronically targeted brain stimulation", *Brain Stimulation*, vol. 15, 116-124, 2022, doi: 10.1016/j.brs.2021.11.014.

[6] J. O. Nieminen, H. Sinisalo, V. H. Souza, M. Malmi, M. Yuryev, A. E. Tervo, M. Stenroos, **D. Milardovich**, J. Korhonen, "Agile TMS: A multi-locus system for rapid and automatic spatial targeting and mapping", *Brain Stimulation*, vol. 14, 6, 2021, doi: 10.1016/j.brs.2021.10.545.

[7] J. O. Nieminen, M. Malmi, **D. Milardovich**, H. Sinisalo, V. H. Souza, A. E. Tervo, M. Yuryev, R. Ilmoniemi, "P78 Multi-locus TMS system for electronically controlled stimulation within a cortical region", *Clinical Neurophysiology*, vol. 131, 4, 2020, doi: 10.1016/j.clinph.2019.12.189.

Conference Contributions

[8] L. Cvitkovich, **D. Milardovich**, C. Wilhelmer, D. Waldhoer, T. Grasser, "Multiscale modeling of thermal oxidation of Si", *The 14th international conference on SiO2, advanced dielectrics and related devices*, Palermo, Italy, 2023.

[9] **D. Milardovich**, D. Waldhoer, A. Jay, A. Hémeryck, T. Grasser, "General Purpose Machine Learning Interatomic Potential for Silicon-Germanium", *2023 Spring Meeting, European Materials Research Society*, Strassbourg, France, 2023.

[10] C. Wilhelmer, **D. Milardovich**, D. Waldhoer, L. Cvitkovich, M. Waltl, T. Grasser, "Intrinsic charge trapping sites in amorphous Si_3N_4", *2023 Spring Meeting, European Materials Research Society*, Strassbourg, France, 2023.

[11] **D. Milardovich**, D. Waldhoer, A. Jay, A. Hémeryck, T. Grasser, "Machine Learning Interatomic Potential for Silicon-Germanium", *4-ème éd. des Journées Toulousaines du Calcul Atomique et Moléculaire*, Toulouse, France, 2022.

[12] **D. Milardovich**, D. Waldhoer, M. Jech, A.-M. El-Sayed, T. Grasser, "Building Robust Machine Learning Force Fields by Composite Gaussian Approximation Potentials", Talk: International Conference on Simulation of Semiconductor Processes and Devices (SISPAD 2022), Granada, Spain; 2022-09-06 - 2022-09-08; in "SISPAD 2022: International Conference on Simulation of Semiconductor Processes and Devices - Conference Abstract Booklet", (2022), 61 - 62.

[13] **D. Milardovich**, M. Jech, D. Waldhoer, A.-M. El-Sayed, T. Grasser, "Machine Learning Prediction of Defect Structures in Amorphous Silicon Dioxide"; Talk: Psi-K Conference (Psi-K) 2022, Lausanne, Switzerland; 2022-08-22 - 2022-08-25; in "PSI-K 2022: abstracts book", (2022), 138.

[14] **D. Milardovich**, M. Jech, D. Waldhoer, A.-M. El-Sayed, T. Grasser, "Machine Learning Prediction of Defect Structures in Amorphous Silicon Dioxide"; Talk: European Solid-State Device Research Conference (ESSDERC), Grenoble, France; 2021-09-13 - 2021-09-22; in "Proceedings of the European Solid-State Device Research Conference (ESSDERC)", (2021), ISBN: 978-1-6654-3748-6, 239-242, doi: 10.1109/ESSDERC53440.2021.9631837.

[15] **D. Milardovich**, M. Jech, D. Waldhoer, M. Waltl, T. Grasser, "Machine Learning Prediction of Formation Energies in a-SiO_2"; Talk: International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Kobe, Japan - virtual; 2020-09-23 - 2020-10-06; in "Proceedings of the International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)", (2020), 339-342, 10.23919/SISPAD49475.2020.9241609.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Curriculum Vitae

Education

02/2020 – 05/2023

Doctoral Candidate and Research Assistant
Institute for Microelectronics, TU Wien
Vienna, Austria

05/2015 – 03/2019

Master's Degree in Electronic Engineering
Thesis titled "Adaptation and Simulation of a Platform for Measuring Losses of Power Electronic Devices"
National University of Rosario
Rosario, Argentina

09/2013 – 02/2014

Exchange Student
University of Zagreb
Zagreb, Croatia

03/2011 – 05/2015

Bachelor's Degree in Engineering Sciences
National University of Rosario
Rosario, Argentina

Research Projects

03/2023 – 05/2023

"Neuromorphic Hardware Accelerator for Machine Learning-Based Molecular Dynamics"
Integrated Systems Laboratory, ETH Zurich
Zurich, Switzerland

09/2022 – 11/2022

"Machine Learning Interatomic Potential for Silicon-Germanium with Application to Neutron Irradiation Simulations"
LAAS, CNRS
Toulouse, France

05/2019 – 01/2020

"Deep Learning Algorithm for Automated Annotation of Motor Evoked Potential Latencies";
"Hardware Developments for a Novel Multi-Coil Transcranial Magnetic Stimulation System"
Department of Neuroscience and Biomedical Engineering, Aalto University
Helsinki, Finland

01/2019 – 04/2019

"Machine Learning Model for Prediction of Steel Slab Temperatures"
Techint
Buenos Aires, Argentina

Awards and Scholarships

2023

PhD-in-kind program: Award granted to fund conference trips during the PhD
Infineon, Austria
International Communication: Award granted to fund a research stay at ETH Zurich
ÖFG, Austria

2022

KUWI: Award granted to fund a research stay at ETH Zurich
TU Wien, Austria

2019

Best project award in the sector during the summer 2019 period
Techint, Argentina

2018

Winterkurs: Six weeks German language and culture course in IIK Berlin
DAAD, Germany

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

2010

Croaticum: One semester of Croatian language course at the University of Zagreb
MVEP, Croatia

Industry and Laboratory Experience

05/2023 – Present

BrightMind.AI: AI Specialist
Vienna, Austria

03/2017 – 01/2019

Nestlé: Engineering Officer
Firmat, Argentina

08/2015 – 01/2017

National University of Rosario: Laboratory Technician
Rosario, Argentina

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.