UNIVERSITY OF LILLE

Doctoral School **ED 104 - Sciences de la Matière, du Rayonnement et de l'Environnement**

University Department **Laboratoire PhLAM**

Thesis defended by **Gilberto Antonio ALOU ANGULO**

Defended on **October 2, 2025**

In order to become Doctor from University of Lille

Academic Field **Physics**

Speciality **Dilute Media and Fundamental Optics**

From Ab Initio Molecular Dynamics to Machine Learning Potential Energy Surface-Based Simulations: NO Oxidation on Oxygen-Functionalized Graphite

Committee members

Referees	Cédric CRESPOS	Professeur–ISM, Université de Bordeaux
	Carine CLAVAGUERA	Directrice de recherche CNRS–ICP, Université Paris-Saclay
Examiners	Denis PETITPREZ (PRESIDENT OF THE JURY)	Professeur–PC2A, Université de Lille
	Pascal LARREGARAY	Directeur de recherche CNRS–ISM, Université de Bordeaux
	Sabine MORISSET	Junior Researcher–ISMO, Université Paris-Saclay
Supervisors	Maurice MONNERVILLE	Professeur–PhLAM, Université de Lille
	Céline TOUBIN	Professeure–PhLAM, Université de Lille
	Alejandro RIVERO-SANTAMARIA	Maître de conférences–PhLAM, Université de Lille

UNIVERSITY OF LILLE

Doctoral School **ED 104 - Sciences de la Matière, du Rayonnement et de l'Environnement**

University Department **Laboratoire PhLAM**

Thesis defended by **Gilberto Antonio ALOU ANGULO**

Defended on **October 2, 2025**

In order to become Doctor from University of Lille

Academic Field **Physics**

Speciality **Dilute Media and Fundamental Optics**

From Ab Initio Molecular Dynamics to Machine Learning Potential Energy Surface-Based Simulations: NO Oxidation on Oxygen-Functionalized Graphite

Committee members

Referees	Cédric CRESPOS	Professeur–ISM, Université de Bordeaux
	Carine CLAVAGUERA	Directrice de recherche CNRS–ICP, Université Paris-Saclay
Examiners	Denis PETITPREZ (PRESIDENT OF THE JURY)	Professeur–PC2A, Université de Lille
	Pascal LARREGARAY	Directeur de recherche CNRS–ISM, Université de Bordeaux
	Sabine MORISSET	Junior Researcher–ISMO, Université Paris-Saclay
Supervisors	Maurice MONNERVILLE	Professeur–PhLAM, Université de Lille
	Céline TOUBIN	Professeure–PhLAM, Université de Lille
	Alejandro RIVERO-SANTAMARIA	Maître de conférences–PhLAM, Université de Lille

UNIVERSITY OF LILLE

École doctorale **ED 104 - Sciences de la Matière, du Rayonnement et de l'Environnement**

Unité de recherche **Laboratoire PhLAM**

Thèse présentée par **Gilberto Antonio ALOU ANGULO**

Soutenue le **2 octobre 2025**

En vue de l'obtention du grade de docteur de l'University of Lille

Discipline **Physique**

Spécialité **Milieux Dilues et Optique Fondamentale**

Des simulations de dynamique moléculaire ab initio aux surfaces d'énergie potentielle basées sur l'apprentissage automatique : Oxydation de NO sur une surface de graphite oxydée

Composition du jury

Rapporteurs	Cédric CRESPOS	Professeur à l'ISM, Université de Bordeaux
	Carine CLAVAGUERA	Directrice de recherche CNRS à l'ICP, Université Paris-Saclay
Examinateurs	Denis PETITPREZ (PRESIDENT OF THE JURY)	Professeur au PC2A, Université de Lille
	Pascal LARREGARAY	Directeur de recherche CNRS à l'ISM, Université de Bordeaux
	Sabine MORISSET	chargée de recherche à l'ISMO, Université Paris-Saclay
Directeurs de thèse	Maurice MONNERVILLE	Professeur au PhLAM, Université de Lille
	Céline TOUBIN	Professeure au PhLAM, Université de Lille
	Alejandro RIVERO-SANTAMARIA	Maître de conférences au PhLAM, Université de Lille

This thesis has been prepared at

Laboratoire PhLAM

Laboratoire PhLAM
CNRS UMR 8523
Université de Lille
Bâtiment P5
59655 Villeneuve d'Ascq
France



☎ (33)(0)3 20 33 60 37

✉ phlam-contact@univ-lille.fr

Web Site http://phlam.univ-lille.fr

List of publications and participation in conferences

Peer-reviewed journal articles produced as part of the thesis project:

1. Gilberto A. Alou Angulo, Alejandro Rivero Santamaría, Céline Toubin, and Maurice Monnerville, "Ab Initio Molecular Dynamics Calculations on NO Oxidation over Oxygen-Functionalized Highly Oriented Pyrolytic Graphite." *The Journal of Physical Chemistry C*, 128, 17894–17904 (2024).

2. Maxime Infuso, Samuel Del Fré, Gilberto A. Alou Angulo, Mathieu Bertin, Jean-Hugues Fillion, Alejandro Rivero Santamaría, Maurice Monnerville, "Deep Potential-Driven Molecular Dynamics of CO Ice Analogues: Investigating Desorption Following Vibrational Excitation." *J. Chem. Phys.*, 163, 084303 (2025).

Participation in scientific events during the thesis period:

1. *Journée des doctorants du PhLAM*.
   Lille, France. April 7, 2023. **Poster:** "Ab Initio Molecular Dynamics calculation on NO oxidation over graphite oxides: Preliminary results."

2. *EGU Training School on Multiscale Characterization of Atmospheric New Particle Formation, EPFL*.
   Lausanne, Switzerland. October 17-19, 2023. **Poster:** "Ab Initio Molecular Dynamics calculation on NO oxidation over graphite oxides."

3. *Journée Scientifique de labex CaPPA*.
   Lille, France. March 12, 2024. **Poster:** "Ab Initio Molecular Dynamics calculations on NO oxidation over oxygen functionalized Highly Oriented Pyrolytic Graphite."

4. *Molecular-Level Understanding of Atmospheric Aerosols (MUOAA 2024).*
   Cargèse, Corsica, France. April 1-5, 2024. **Poster:** "Reactivity, scattering, and energetic distribution of collisions between nitric oxide and oxidized graphite: Insights from AIMD calculations."

5. *Journée des doctorants du PhLAM*.
   Lille, France. April 17, 2024. **Talk:** "Ab Initio Molecular Dynamics calculations on NO oxidation over oxygen functionalized Highly Oriented Pyrolytic Graphite."

6. *Journées de chimie théorique du pôle Nord/IdF de ThéMoSiA*.
   Paris, France. May 21-22, 2024. **Talk:** "Ab Initio Molecular Dynamics calculations on NO oxidation over oxygen functionalized Highly Oriented Pyrolytic Graphite."

7. *Journées Théorie, Modélisation et Simulations (JTMS 2024).*
   Toulouse, France. October 8-9, 2024. **Talk:** "Ab Initio Molecular Dynamics calculations on NO oxidation over oxygen functionalized Highly Oriented Pyrolytic Graphite."

8. *XII International Meeting on Photodynamics and Related Aspects*.
   Santiago, Chile. November 11-15, 2024. **Talk:** "Ab Initio Molecular Dynamics calculations on NO oxidation over oxygen functionalized Highly Oriented Pyrolytic Graphite."

9. *Journée scientifique IrePSE*.
   Lille, France. November 19, 2024. **Talk:** "Ab Initio Molecular Dynamics calculations on NO oxidation over oxygen functionalized Highly Oriented Pyrolytic Graphite."

10. *Workshop CaPPA Area*.
    Lille, France. March 6, 2025. **Poster:** "Theoretical study of the molecular dynamics of complex gas-surface systems using Artificial Intelligence methods."

11. *Journée des doctorants du PhLAM*.
    Lille, France. March 20, 2025. **Poster:** "Theoretical study of the molecular dynamics of complex gas-surface systems using Artificial Intelligence methods."

Acknowledgments

The completion of this doctoral thesis marks the end of an enriching journey that would not have been possible without the guidance, support, and encouragement of many people and institutions.

I would first like to express my deepest gratitude to the members of the jury for accepting to evaluate my doctoral work and for the time and attention they dedicated to reading and assessing this thesis. I am equally grateful to the Doctoral School SMRE and to the CaPPA project for their financial support throughout these years. I also wish to extend my sincere thanks to the PhLAM Laboratory for welcoming me during this period and providing a stimulating and collaborative research environment.

During this PhD, I have had the privilege of learning from inspiring mentors. I would like to express my heartfelt gratitude to Alejandro, Céline, and Maurice, without whom the realization of this thesis would not have been possible. Thank you for your trust, your expertise, your guidance, and above all, for your humanity.

I would also like to express my sincere gratitude to my entire family. If I am the person I am today, it is thanks to the education and affection I received from my father and my mother (whose memory and teachings continue to inspire me every day). I wish I could name each and every one of them, but I come from such a large family that it would be unfair to leave anyone out.

My deepest thanks also go to my beautiful wife Belle, for her unwavering love and support through every stage of this journey. Thank you for being my source of balance and motivation, for reminding me of what truly matters beyond the challenges of research, and for your patience and understanding during the many hours I devoted to this work instead of being by your side. I love you.

I also wish to thank the PCMT team for their warm welcome and support. Among them, I owe a special thanks to my dear friends Maxime and Samuel for their constant help and advice throughout the thesis, and for making my Cuban-to-(northern)French transition such a joyful experience, one beer and one “welsh” at a time.

Of course, none of this would have been possible without the support of those closest to me. Thank you to Magui for your help and advice, who together with Ale, Vivianne and Maurice, made me feel that I have my second family here in Lille. My thanks also to Yosvany, always there when I needed him, as well as to Manolito, Lali, and Alfa for their friendship and encouragement.

Finally, a huge thank you to all those I have not mentioned by name here — your support and friendship have been truly invaluable throughout this journey. I apologize if I have forgotten anyone, but please know that you all hold a special place in my heart.

Affidavit

I, undersigned, Gilberto Antonio Alou Angulo, hereby declare that the work presented in this manuscript is my own work, carried out under the scientific direction of Dr. Maurice Monerville, Dra, Céline Toubin and Dr. Alejandro Rivero Santamaría, in accordance with the principles of honesty, integrity and responsibility inherent to the research mission. The research work and the writing of this manuscript have been carried out in compliance with both the French National Charter for Research Integrity and the University of Lille charter on the fight against plagiarism.

This work has not been submitted previously either in this country or in another country in the same or in a similar version to any other examination body.

Lille, on November 16, 2025.



This work is made available under the terms of the license Creative Commons Attribution - Non Commercial - No Derivatives 4.0 International (CC BY-NC-ND 4.0).

From Ab Initio Molecular Dynamics to Machine Learning Potential Energy Surface-Based Simulations: NO Oxidation on Oxygen-Functionalized Graphite

Abstract

Oxygen-functionalized graphite surfaces (O@HOPG) are commonly found in the atmosphere due to exposure to oxidizing pollutants such as ozone (O_3) and hydrogen peroxide (H_2O_2). These surface-bound oxygen groups play a crucial role in atmospheric chemistry, as they can react with a broad range of pollutants, including hydrocarbons and nitrogen oxides (NO_x). Among these, NO_x are major contributors to smog and acid rain; nitric oxide (NO) is a precursor to nitrogen dioxide (NO_2), which also promotes the formation of secondary organic aerosols—known to negatively impact both human health and the environment.

This thesis presents a theoretical investigation of the oxidation of NO molecules on epoxy-functionalized HOPG, with relevance to heterogeneous atmospheric chemistry and pollution abatement. The study relies on an approach combining Density Functional Theory (DFT)-based ab initio Molecular Dynamics (AIMD) with a machine learning potential energy surface (ML-PES). Four collision energies (0.025, 0.05, 0.1, and 0.3 eV) and two different orientations were analyzed where the reaction, adsorption, and scattering probabilities were computed.

AIMD simulations reveal that NO_2 formation can occur even at the lowest collision energy investigated (0.025 eV), approximately equivalent to room temperature (300 K), which agrees qualitatively with the experimental results. Additionally, simulations show that scattered NO molecules exhibit low specular reflection, lose half of their initial translational energy, and remain vibrationally cold with minimal rotational excitation. Furthermore, a statistical analysis of all reactive trajectories, elucidated the structural requirements for the reaction to efficiently occur. Finally, this theoretical study demonstrates the potential of oxygen-doped carbon surfaces for the conversion of NO to NO_2 highlighting the importance of explicitly accounting for dynamical effects on the NO oxidation.

To circumvent the high computational cost of AIMD simulations and the limited statistical sampling, a ML-PES was constructed and trained on configurations extracted from AIMD trajectories. The accuracy of the ML-PES was rigorously validated by direct comparison with AIMD results, confirming its ability to reproduce both reactive and non-reactive events with high fidelity. The ML-PES successfully captures the essential dynamical and energetic features of the NO + O@HOPG $\longrightarrow$ NO_2 + HOPG reaction, accurately identifying reaction channels and characterizing energy transfer mechanisms, while reducing the computational cost by approximately two orders of magnitude compared to conventional AIMD simulations. The high efficiency of ML-PES simulation makes it possible to enhance the sampling of reactive events, thereby providing statistically robust insights into the underlying mechanisms of NO oxidation. Nevertheless, limitations persist in regions of the configurational space that are underrepresented in the training data. To overcome this, an active learning strategy is proposed to iteratively refine the ML-PES by incorporating dynamically important configurations, thereby improving the model's robustness and transferability.

This work establishes a robust computational framework that combines AIMD and ML-PES for the atomistic simulation of gas–surface reactions. While this thesis concentrated on the NO oxidation on O@HOPG, the developed methodology is broadly transferable and can be extended to study other gas-surface reactions of atmospheric relevance within the context of depollution or catalysis.

Keywords: ab initio molecular dynamics, machine learning potentials, gas–surface reactions, atmospheric chemistry, graphite, nitric oxide, dft

DES SIMULATIONS DE DYNAMIQUE MOLÉCULAIRE AB INITIO AUX SURFACES D'ÉNERGIE POTENTIELLE BASÉES SUR L'APPRENTISSAGE AUTOMATIQUE : OXYDATION DE NO SUR UNE SURFACE DE GRAPHITE OXYDÉE

Résumé

Suite aux émissions anthropogéniques, les particules de carbone sont abondamment présentes dans l'atmosphère et particulièrement sous forme oxydées par des oxydants tels que l'ozone (O_3) et le peroxyde d'hydrogène (H_2O_2). Les groupes oxygénés en surface jouent un rôle clé, réagissant avec divers polluants, notamment les hydrocarbures et les oxydes d'azote (NO_x). Ces derniers sont des précurseurs majeurs du smog et des pluies acides ; le monoxyde d'azote (NO), en particulier, se transforme en dioxyde d'azote (NO_2), favorisant la formation d'aérosols organiques secondaires, nocifs pour la santé humaine et l'environnement.

Cette thèse présente une étude théorique de l'oxydation des molécules de NO sur du graphite pyrolytique hautement orienté (HOPG) fonctionnalisé par des groupes époxy. L'étude repose sur une approche combinant la dynamique moléculaire ab initio (AIMD) basée sur la théorie de la fonctionnelle de la densité (DFT) et une surface d'énergie potentielle obtenue par apprentissage automatique (ML-PES). Les probabilités de réaction, d'adsorption et de diffusion ont été caractérisées pour quatre énergies de collision et deux orientations différentes. Les simulations AIMD révèlent que la collision réactive, conduisant à la formation de NO_2, peut se produire même à la plus faible énergie de collision étudiée (0.025 eV), correspondant approximativement à la température ambiante (300 K), ce qui s'accorde qualitativement avec les résultats expérimentaux. Ce travail met en évidence l'influence des effets dynamiques sur le processus d'oxydation du NO à la surface de graphite activée. Les distributions angulaires et énergétiques des produits révèlent que Les molécules de NO diffusées présentent une faible réflexion spéculaire, perdent la moitié de leur énergie translationnelle initiale et restent vibrationnellement froides, avec une excitation rotationnelle minimale. En outre, une analyse statistique de toutes les trajectoires réactives, a permis d'élucider les exigences structurales pour que la réaction ait lieu dans des conditions dynamiques. Cependant, le coût numérique élevé des simulations AIMD limite l'échantillonnage statistique des événements réactifs rares. Pour pallier à cette limitation, une ML-PES a été construite et entraînée sur des configurations extraites des trajectoires AIMD. L'exactitude de la ML-PES a été rigoureusement validée par une comparaison directe avec les résultats AIMD, confirmant sa capacité à reproduire avec une grande fidélité aussi bien les événements réactifs que non réactifs. La ML-PES capture avec succès les caractéristiques dynamiques et énergétiques essentielles de la réaction NO + O@HOPG $\longrightarrow$ NO_2 + HOPG, identifiant précisément les canaux réactionnels et caractérisant les mécanismes de transfert d'énergie, tout en réduisant le coût computationnel d'environ deux ordres de grandeur par rapport aux simulations AIMD conventionnelles. Cette réduction spectaculaire du coût numérique rend possible l'exploration d'un nombre bien plus important d'événements réactifs, offrant ainsi une vision statistiquement robuste des mécanismes sous-jacents de l'oxydation du NO. Néanmoins, certaines régions de l'espace des configurations restent sous-échantillonnées dans les données d'entraînement. Pour y remédier, une stratégie d'apprentissage actif est proposée afin d'affiner itérativement la ML-PES en intégrant des configurations dynamiquement importantes, améliorant ainsi la robustesse et la transférabilité du modèle. Ce travail établit une méthodologie solide combinant AIMD et ML-PES pour la simulation atomique des processus gaz–surface. Bien que cette thèse se concentre sur l'oxydation du NO sur les surfaces O@HOPG, les développements proposés sont largement transférables. En particulier, l'approche peut être facilement étendue à l'étude d'autres réactions avec des surfaces carbonées fonctionnalisées d'intérêt atmosphérique.

Mots clés : dynamique moléculaire ab initio, potentiels d'apprentissage automatique, réactions gaz–surface, chimie atmosphérique, graphite, monoxyde d'azote, dft

Contents

List of publications and participation in conferences i

Acknowledgments iii

Affidavit i

Abstract iii

Contents v

List of Figures ix

List of Tables xvii

Acronyms xix

1 Introduction and State of the Art 1

1.1 The role of NOx oxidation in atmospheric pollution 5

1.1.1 Soot particles as NO oxidants and graphite as a model system . . 8

1.2 NO Removal and Control Strategies 10

1.2.1 Non-Catalytic NOx control methods 11

1.2.2 Catalytic NOx control methods 12

1.2.3 Role of NO oxidation in emission control methods 13

1.3 Carbon-Based catalysts for NO oxidation: Mechanisms and Challenges 14

1.3.1 Insights from experimental investigations 14

1.3.2 Theoretical Perspectives and Computational Approaches 16

1.3.3 Unresolved challenges and knowledge gaps 17

1.4 Advanced computational methods for NO oxidation on carbon-based catalysts 18

1.5 Scientific challenges and thesis objectives 20
1.6 Thesis Structure 21

2 Modeling Gas–Surface Interactions with AIMD **23**
2.1 Why Ab Initio Molecular Dynamics? 24
2.2 Quantum foundations and principles of AIMD 27
2.3 DFT as the workhorse of AIMD 30
2.3.1 Dispersion Corrections in DFT: The DFT-D Method 41
2.3.2 Treatment of spin in the Kohn–Sham formalism 43
2.3.3 DFT for Periodic Systems: Plane-Wave Expansion and Bloch Formalism 45
2.4 Molecular dynamics on a quantum PES 48

3 NO oxidation on O@HOPG from AIMD approach **51**
3.1 Computational Setup 52
3.2 Computational Modeling of the O@HOPG surface 53
3.2.1 Unit cell lattice geometry 53
3.2.2 HOPG surface 59
3.2.3 O@HOPG surface 61
3.3 Validation of the NO–O@HOPG Surface Interaction Model 66
3.3.1 Dissociation of NO 66
3.3.2 NOx adsorption on the O@HOPG surface 71
3.4 AIMD's initial conditions 78
3.4.1 O@HOPG surface thermalization 78
3.4.2 Geometrical sampling of the NO molecule 80
3.4.3 Initial velocity vector of the NO molecule 82
3.4.4 Molecular dynamics simulations 86
3.5 Results and discussion 91
3.5.1 Reaction probabilities and reaction mechanism 91
3.5.2 Products angular distributions 98
3.5.3 Products energy distributions 100
3.6 Conclusion of the chapter 104

4 ML-PES for Studying Gas–Surface Interactions **109**
4.1 From AIMD to ML-PES: Bridging Accuracy and Efficiency in Gas–Surface Molecular Dynamics 110
4.2 ML-PESs for Gas– Surface Interactions 111
4.2.1 Representation of Atomic Configurations 112
4.2.2 Structure and Training of ML-PES 120
4.2.3 ML-PES validation 126
4.2.4 ML-PES driven MD 126

4.2.5 Active Learning 127

5 NO Oxidation on O@HOPG: MD Simulations with a ML–PES 131
5.1 Building a ML-PES for NO oxydation over O@HOPG 132
5.1.1 Conformation of the training set 133
5.1.2 Training Procedure 138
5.1.3 Validation of the ML-PES 139
5.2 Validating the ML-PES with MD simulations mirroring AIMD initial conditions 144
5.2.1 Computational Setup 145
5.2.2 Validation Results 146
5.3 Active learning methodology 154
5.4 Outlook on Large-Scale MD Applications 156
5.5 Conclusions of the chapter 156

Conclusions and perspectives 159

Bibliography 163

List of Figures

1.1 Representative adsorption mechanisms: (a) dissociative adsorption; (b) molecular adsorption (physisorption and chemisorption); (c) abstraction. (Source C. Díaz [1]) 3

1.2 Representative scattering mechanisms: (a) elastic scattering; (b) rotational excitation; (c) vibrational excitation; (d) diffraction. (Source C. Díaz [1]) 4

1.3 Surface-mediated reaction mechanisms: (a) Langmuir–Hinshelwood; (b) Eley–Rideal; (c) hot-atom mechanism. (Source C. Díaz [1]) 4

1.4 An infographic illustrating various sources and effects of environmental pollution, including NO_x emissions, acid rain, ozone depletion, and water contamination. (Source: *Britannica*) 6

1.5 Simplified schematic representation of tropospheric ozone formation involving NO_x, VOC, and CO under solar radiation. Photolysis of NO_2 produces ozone, while reactive radicals (HO, HO_2) from VOC and CO oxidation enhance NO to NO_2 conversion without ozone consumption, leading to net ozone accumulation. $PM_{2.5}$ and stratospheric exchange also influence background ozone levels. (Adapted from *Zhang et al.* [11]). 7

1.6 Schematic representation of soot formation and evolution in a flame. The molecular species and particles are not drawn to scale. (Source: *Michelsen et al.* [21]) 9

2.1 Functional categorization according to Perdew's "Jacob's ladder". ρ = electron density, τ = kinetic energy density, ϕ = molecular orbital, Fock exc. = Fock exchange. (Source Bursch et al. [116]) 36

2.2 Overview of commonly used dispersion correction methods in DFT where E_{KS} and V_{KS} correspond to the bare Kohn–Sham total energies and potentials, respectively. These scheme include density-based approaches (vdW-DF, parameterized functionals), C_6-based pairwise correction schemes (DFT-D), and effective one-electron potential modifications (DCACP, LAP/DCP). (Source Grimme [140]) 42

3.1 QR code linking to the publication by Alou et al. [153] 51

3.2 Crystal structure of highly oriented pyrolytic graphite. (A) Side view of the primitive hexagonal unit cell with lattice vectors a, b, and c; red and green spheres represent the two inequivalent carbon atoms occupying Wyckoff positions with fractional coordinates (0, 0, 1/4) and (1/3, 2/3, 1/4), respectively. The labels d_g and d_{C-C} indicate the interlayer distance and in-plane C–C bond length. (B) Top view along the (0001) direction, illustrating the hexagonal arrangement of the carbon atoms. (Source Lechner et al. [163]) 55

3.3 Energy versus volume curve for graphite obtained from first-principles calculations using the PBE-D3(BJ) functional. The red dots represent the computed total energies for different volumes of the fully relaxed unit cell. The solid lines correspond to fits using different equations of state: Murnaghan, Birch-Murnaghan, Birch, and Vinet. The equilibrium volume and bulk modulus were extracted from each fit and are reported in Table 3.2. 57

3.4 Crystallographic structure of graphite. 59

3.5 Graphite slab model used in surface simulations. The supercell consists of a $p(4 \times 4)$ surface unit, with three graphene layers (labeled L1, L2, and L3) stacked in the AB (Bernal) sequence. Each layer contains 32 carbon atoms, totaling 96 atoms in the slab. The interlayer spacing is 3.37 Å, and a vacuum region of 20 Å is added along the z-direction to eliminate spurious interactions between periodic images. The in-plane dimensions are 9.87 Å along both $\vec{a}_1$ and $\vec{a}_2$. 60

3.6 Top view of a graphene sheet illustrating the three high-symmetry adsorption sites commonly considered for O adsorption: *top* (above a carbon atom), *bridge* (above the midpoint of a C–C bond), and *hollow* (above the center of a hexagonal ring). 62

3.7 Dissociation energy profiles and corresponding magnetic moments for the NO molecule computed using two approaches: point-to-point (PtP) and constant-velocity (V-Const) dissociation. **(A)** Results for NO placed at 10 Å over the O@HOPG surface. **(B)** Results for NO in a periodic vacuum box of 20 Å; the inset illustrates the simulation cell. Solid lines represent total energies as a function of the N–O bond length, while dashed lines indicate the evolution of the system's magnetic moment. The color scheme is consistent across panels: blue for PtP and green for V-Const. 68

3.8 Comparison between spin-polarized (SP) and non-spin-polarized (NSP) DFT calculations of the NO dissociation profile over the O@HOPG surface. The energy curves are nearly indistinguishable in the bound region ($d \lesssim 1.65$ Å), with significant deviations arising only as the molecule approaches the dissociation limit. 70

3.9 Morse potential fit to the vacuum dissociation curve of NO. Red dots: DFT energy points. Blue line: analytical Morse fit. 71

3.10 Initial configurations considered for the spin-polarized PBE-D3(BJ) geometry optimizations of NO adsorbed on the O@HOPG surface. (**A**) NO horizontal with oxygen at top, (**B**) NO vertical with oxygen at top, (**C**) NO horizontal at mid-plane, (**D**) NO vertical with N at top, and (**E**) NO horizontal with N at top. 72

3.11 Optimized spin-polarized adsorption geometry of NO on the O@HOPG surface. The NO molecule is positioned above the epoxy group in a tilted configuration, with a N–O bond length of 1.17 Å and a N–O_{HOPG} distance of 2.70 Å. The nitrogen atom lies 3.11 Å above the surface plane, while the C–O bonds of the underlying epoxide measure 1.47 Å. This configuration corresponds to the global minimum among the adsorption structures explored. 73

3.12 Initial configurations considered for spin-polarized PBE-D3(BJ) geometry optimizations of NO_2 on the HOPG surface. (**A**) Horizontal alignment with the NO_2 molecular plane normal to the surface and the oxygen oriented up. (**B**) Horizontal alignment with the NO_2 molecular plane normal to the surface and the oxygen oriented down. (**C**) Vertical alignment with the NO_2 molecular plane normal to the surface; the closer oxygen atom to the surface was placed above the bridge, top, and hollow sites of the underlying carbon lattice to explore the influence of the local adsorption environment. 75

3.13 Optimized spin-polarized adsorption geometry of NO_2 on the HOPG surface. The NO_2 molecule is arranged in a triangular configuration above the HOPG surface, with the nitrogen atom tilted toward the surface. The N–O bond lengths are 1.23 Å, the O–N–O angle is 129.56°, and the vertical height of the molecule is 3.60 Å. This structure corresponds to the global minimum among the adsorption geometries explored. 76

3.14 Proposed reaction coordinate for the surface-mediated oxidation of NO by an epoxy group on O@HOPG. The start and end correspond to relaxed NO and NO_2 molecules in the gas phase at 10.0 Å above the surface. Adsorbed intermediates (NO AS and NO_2 AS) represent the most stable configurations identified in Sections 3.3.2.1 and 3.3.2.2. The transition state (TS) structure and activation energy (0.10 eV) were adapted from Hou et al. [70]. 77

3.15 Schematic representation of the initial conditions used to sample the gas-phase NO molecule in the AIMD simulations. **(A)** Sampling region for the NO center of mass ($x_{\text{C.M.}}, y_{\text{C.M.}}$) in the XY-plane. The red disk is centered on the position of the adsorbed oxygen atom (red sphere) on the graphite surface, and black dots indicate randomly selected center-of-mass positions. **(B)** Definition of the molecular orientation in spherical coordinates. The polar angle α and azimuthal angle β determine the orientation of the NO molecular axis with respect to a framework at its center of mass. 81

3.16 Schematic representation of the initial spatial arrangement of the NO molecule relative to the graphite surface. The vector $\vec{r}_{O_{\text{HOPG}}}$ represents the position of the oxygen atom adsorbed on the surface, while $\vec{r}_{\text{CM-O}}$ connects the center of mass of NO to this reactive site. The polar angle φ and azimuthal angle θ define the orientation of the velocity vector relative to $\vec{r}_{\text{CM-O}}$, with dotted lines indicating the projection of vectors onto the xy-plane. 84

3.17 Schematic representation of the initial configuration for the quasi-classical AIMD simulations. The center of mass of the NO molecule is randomly placed inside a disk of radius $R = 5$ Å centered on the surface oxygen atom and located at a height $H = 9$ Å above the O@HOPG surface. The surface is thermalized at $T = 300$ K. The direction of the NO translational velocity vector $\vec{v}_{\text{NO}_{\text{CM}}}$ defines the incidence conditions (normal (blue vector) and oriented (orange vector)). 88

3.18 Reaction probabilities as a function of the collision energy for both normal and oriented incidences. The probability errors were calculated by the usual formula for binomial distributions. 92

3.19 Top Panel: Time evolution of the total potential energy along a typical reactive trajectory. IS designates the initial state, marking the point when the potential energy of the system starts to decrease, while FS indicates the final state, where the potential energy of the system stabilizes and NO_2 is formed. The red line correspond to a Savitzky-Golay filter [220] applied to the potential energy. Panel A: Probability distribution of the energy released during the reaction. Panel B: Probability distribution of the reaction time computed as ($\Delta t = t_{FS} - t_{IS}$). Both distributions are normalized to all the reactive trajectories from the four initial kinetic energies (0.025, 0.05, 0.1, 0.3 eV) and the two investigated incidences (normal and oriented). The size of the bins are 0.13 eV and 16.6 fs respectively. 95

3.20 Probability distribution of the geometrical configuration regarding the NO and the epoxy group at the initial state (IS): distance between the oxygen atom of the epoxy group with respect to the first carbon atom (A) and the second carbon atom (B); angle defined by the O atom of NO, the N atom of NO, and the O atom of the epoxy group (C); distance between the center of mass of NO and the HOPG surface (D); distance between the N atom of NO and the O atom of the epoxy group (E); and a representative configuration at the average values of these geometrical parameters (F). All distributions are normalized to all the reactive trajectories from the four initial kinetic energies (0.025, 0.05, 0.1, 0.3 eV) and the two incidences (normal and oriented) of the NO. The size of the bins are 0.06 Å, 0.06 Å, 6.0°, 0.17 Å and 0.11 Å respectively. 97

3.21 Comparative analysis of the distributions of the angular deviations $\Delta\theta$ after reflection from the oxidized graphite surface at an initial kinetic energy of 0.3 eV. The distributions are normalized to the total number of scattered molecules and the bin size is 9°. (A) for NO scattered molecules with initial normal and oriented incidences; (B) for desorbed NO and NO_2 molecules for initial NO with oriented incidence. 99

3.22 Time evolution of the average NO center-of-mass distance to the surface (solid lines, left Y axis) and fraction of molecules that remain within the scattering plane (dashed lines, right Y axis) during the first 1000 fs for oriented and normal incidence at a collision energy of 0.3 eV. 100

3.23 Translational energy distribution of NO molecules scattered with an initial kinetic energy of 0.3 eV at normal and oriented incidences. The distributions are normalized to the total number of scattered NO molecules and the bin size is 0.03 eV. 101

3.24 Rotational distribution of scattered NO molecules with 0.3 eV of initial kinetic energy at normal and oriented incidence. The distributions are normalized to the total number of scattered NO molecules and the bin size is 1. 103
3.25 Translational energy distribution of NO_2 molecules scattered for an NO initial kinetic energy of 0.3 eV at oriented incidence. The distribution is normalized to the total number of scattered NO_2 molecules and the bin size is 0.1 eV. 104

4.1 Schematic representation of the iterative workflow for constructing a ML-PES. (Adapted from Miksch et al. [238]) 113
4.2 Schematic illustration of the Artrith–Urban–Ceder (AUC) descriptor as implemented in ÆNET [82]. The local atomic environment is encoded through a basis expansion of the radial (a) and angular (c) distribution functions using Chebyshev polynomials $\bar{\phi}_\alpha$. The corresponding expansion coefficients (b) and (d) are weighted by atom-type-specific factors ω_{t_j} and ω_{t_k}, which encode the local chemical composition. (Source Miksch et al. [238]) 116
4.3 Schematic representation of a feedforward neural network with a 2–3–3–1 architecture, where the numbers indicate how many nodes are in each layer (i.e., two nodes in the input layer, two hidden layers with three nodes each, and a single output node). The input layer receives a vector of symmetry functions describing the local environment of atom I, and the output layer returns its atomic energy contribution E_I. Only a few weights $w_{i,j}^{k,l}$ are shown for clarity. 121
4.4 Schematic architecture of a ML-PES, each atom is represented by a vector of ACSF, which encode its local environment within a cutoff radius. These descriptors are used as input to atomic neural networks that predict atomic energies E_I^α for each atom $I = 1...N$ of type α. The total potential energy E_{tot} is obtained by summing the contributions from all atoms. Identical neural network architectures and symmetry function sets are employed for all atoms of a given element. (Adapted from Tokita et al. [247]) 123
4.5 Schematic evolution of errors on the training and test sets in cases of underfitting (a), ideal fitting (b), and overfitting (c) during neural network training. 125

5.1 Illustration of the representative regions along the reaction pathway for NO oxidation on O@HOPG surface. Based on typical AIMD trajectories, three distinct stages are identified, (green) weak interaction, (purple) reaction or scattering, (orange) products. 134

5.2 Schematic representation of the selection criteria for configurations along the AIMD trajectories. The left panel illustrates reactive trajectories, where configurations were sampled from three regions: weak interaction (20 configurations), reaction (50 configurations), and products (30 configurations), defined by the initial state time (t_{IS}) and final state time (t_{FS}). The right panel shows non-reactive trajectories, where configurations were selected based on the turning point time (t_{TP}) and a temporal window Δt, with 20 configurations from the weak interaction region (up to $t_{TP} - \Delta t$), 50 configurations from the scattering region ($t_{TP} - \Delta t$ to $t_{TP} + \Delta t$), and 30 configurations from the products region (after $t_{TP} + \Delta t$). 135

5.3 Potential energy distribution of the training set configurations, illustrating the continuous sampling of the reaction coordinate from reactants to products. Representative system geometries corresponding to different energy regions are shown alongside the histogram. 137

5.4 Evolution of the training and testing errors during the optimization of the ML-PES over 10 000 epochs. Panel A shows the total loss function, while Panels B and C present the separate contributions from the energy and force errors, respectively. In each case, the model achieves smooth convergence with testing errors remaining consistently low, indicating robust generalization to unseen data. Energy errors are expressed in meV per atom and force errors in eV/Å. 140

5.5 Distribution of DFT reference energies for the configurations in the validation set (orange line), compared to the energy distribution of the full AIMD dataset (blue bars). 141

5.6 Comparison between DFT reference energies and neural network predicted energies for the validation set. The parity plot illustrates the strong correlation across the energy range, while the error histogram shows that the majority of configurations exhibit small deviations, with errors well below 0.1 eV. 142

5.7 Force predictions on the independent validation set. The left panel shows the joint distribution of the force magnitude ratio $\|\vec{F}_{DFT}\|/\|\vec{F}_{ML-PES}\|$ and the angle ϕ between the predicted and reference force vectors, highlighting the strong correlation in both magnitude and direction. The right panel displays the distribution of absolute errors in force magnitudes ($\|\vec{F}_{DFT}\| - \|\vec{F}_{ML-PES}\|$), with the majority of configurations showing deviations close to zero. 143

5.8 Comparison between the NO dissociation profile predicted by the neural network potential and the DFT reference at a height of 3.5 Å above the O@HOPG surface. The neural network accurately reproduces the DFT dissociation curve across the full range of N–O bond distances. 144

5.9 Comparison of reaction probabilities as a function of collision energy for NO impinging on the O@HOPG surface, obtained from AIMD simulations (solid lines) and ML-PES driven molecular dynamics (dashed lines). Results are shown for both normal (blue) and oriented (orange) incidence conditions. The probability errors were calculated by the usual formula for binomial distributions. 147
5.10 Comparison of reaction mechanism observables between AIMD (colored bars) and ML-PES (black triangles): (A) potential energy change, (B) delay times between IS and FS. 149
5.11 Scattering angle distribution of NO at an initial translational energy of 0.3 eV. The bars correspond to the reference AIMD results (blue for normal incidence, orange for oriented incidence), while black triangles and circles indicate the ML-PES predictions for normal and oriented incidence, respectively. 150
5.12 Translational energy distribution of scattered NO at an initial translational energy of 0.3 eV. The bars correspond to the reference AIMD results (blue for normal incidence, orange for oriented incidence), while black triangles and circles indicate the ML-PES predictions for normal and oriented incidence, respectively. 151

List of Tables

3.1 Energy–volume dataset for graphite obtained from DFT calculations with PBE-D3(BJ) dispersion correction. The lattice parameters $a = b$, c, unit cell volume V, total energy E, and calculated pressure p are reported for each configuration. 56
3.2 Bulk modulus values for graphite obtained in this work by fitting the energy–volume relation using various EoS, compared with theoretical predictions from Lechner et al. and experimental reference values. . . . 58
3.3 DFT adsorption energies and structural parameters of oxygen adsorbed on graphite/graphene at bridge and top sites. 63
3.4 Reaction and scattering probabilities of the 300 trajectories at each incidence energy for normal incidence and oriented incidence (in parentheses) 92

Acronyms

ACs Activated Carbons
ACSF Atom-Centered Symmetry Function
ADF Angular Distribution Function
AEF Atomic Environment Function
AI Artificial Intelligence
AIMD *Ab Initio* Molecular Dynamics
AIMDef *Ab Initio* Molecular Dynamics with Electronic Friction
AO Atomic Oxygen
ASE Atomic Simulation Environment
AUC Artrith–Urban–Ceder Descriptor

B2PLYP Becke 2006 Double-Hybrid Functional Combining B88 Exchange and LYP Correlation
B3LYP Becke Three-Parameter Lee–Yang–Parr Hybrid Functional
B86 Becke 1986 Exchange Functional
B88 Becke 1988 Exchange Functional
BDE Bond Dissociation Energy
BLYP Becke 1988 Exchange and Lee–Yang–Parr Correlation Functional
BP86 Becke 1988 Exchange and Perdew 1986 Correlation Functional
BPW91 Becke 1988 Exchange and Perdew–Wang 1991 Correlation Functional

CAM Corrected Asymptotic Model Functional
CBS-APNO Complete Basis Set–Atomic Pair Natural Orbital Method
CBS-QB3 Complete Basis Set–Quadratic Configuration Interaction with Perturbative Triples
CC Coupled Cluster
CI Configuration Interaction
CI-NEB Climbing-Image Nudged Elastic Band Method
CNTs Carbon Nanotubes

CO Carbon Monoxide
CO_2 Carbon Dioxide
CPU Central Processing Unit
CRP Corrugation-Reducing Procedure
CV Collective Variable

DCACP Dispersion-Correcting Atom-Centered Potential
DCM Dynamic Corrugation Model
DeePMD-kit Deep Potential Molecular Dynamics Software Package
DFT Density Functional Theory
DFT-D Density Functional Theory with Empirical Dispersion Corrections
DFT-D3 Density Functional Theory with Grimme's D3 Dispersion Correction
DFT-D3(BJ) Density Functional Theory with Grimme's D3 (Becke–Johnson damping) Dispersion Correction
DFTB Density Functional Tight Binding
DHDFs Double-Hybrid Density Functional
DOCs Diesel Oxidation Catalysts
DPFs Diesel Particulate Filters

E-R Eley–Rideal mechanism
EAM Embedded Atom Method
EoS Equation of State

FFNN Feedforward Neural Network
FS Final State
FT97 Filatov–Thiel 1997 Functional

G4 Gaussian-4 Composite Method
GAP Gaussian Approximation Potential
GGA Generalized Gradient Approximation
GNN Graph Neural Network
GO Graphene Oxide
GPR Gaussian Process Regression
GPU Graphics Processing Unit
GTF Gaussian-Type Function Basis Set

H_2 Molecular Hydrogen
H-A Hot Atom mechanism
HDNNP High-Dimensional Neural Network Potential
HF Hartree–Fock
HK Hohenberg–Kohn
HNO_2 Nitrous acid

HNO_3 Nitric acid
HO Hydroxyl radical
HO_2 Hydroperoxyl radical
HOPG Highly Oriented Pyrolytic Graphite
HSE Heyd–Scuseria–Ernzerhof Screened Hybrid Functional

IS Initial State

KS Kohn–Sham

L-H Langmuir–Hinshelwood mechanism
LAMMPS Large-scale Atomic/Molecular Massively Parallel Simulator
LAP Laplacian-Based Correlation Functional (Proynov–Vela–Salahub)
LAP/DCP Laplacian-Based or Dispersion-Correcting Potential Approach for Effective One-Electron Dispersion Corrections
LDA Local Density Approximation
LG Lee–Ghosh Functional
LNT Lean NO_x Trap
LSDA Local Spin Density Approximation
LYP Lee–Yang–Parr Correlation Functional

M06-2X Minnesota 2006 Double Hybrid Meta-GGA Functional with 54% Hartree–Fock Exchange
M3GNet Materials 3-Body Graph Network for Universal Interatomic Potentials
MAE Mean Absolute Error
MD Molecular Dynamics
MD-DFT Molecular Dynamics–Density Functional Theory
meta-GGA Meta-Generalized Gradient Approximation
ML Machine Learning
ML-PES Machine Learning Potential Energy Surface
MOFs Metal–Organic Frameworks
MP2 Second-Order Møller–Plesset Perturbation Theory
MPn Møller–Plesset Perturbation Theory of order n
MRCI Multireference Configuration Interaction
MS Modified Shepard
MTP Moment Tensor Potential

N_2 Molecular Nitrogen
NEB Nudged Elastic Band
NN Neural Network
NO Nitric Oxide
NO AS Nitric Oxide Adsorbed on the O@HOPG Surface

NO_2 Nitrogen Dioxide
NO_2 AS Nitrogen Dioxide Weakly Adsorbed Post-Reaction
NO_x Nitrogen Oxides
NO_{CM} Center of Mass of the NO Molecule
NSP Non-Spin-Polarized
NVE Microcanonical Ensemble with Constant Number of Particles, Volume, and Energy
NVT Canonical Ensemble with Constant Number of Particles, Volume, and Temperature

O_2 Molecular Oxygen
O_3 Ozone
O@HOPG Oxygen-functionalized Highly Oriented Pyrolytic Graphite

P86 Perdew 1986 Correlation Functional
PAHs Polycyclic Aromatic Hydrocarbons
PANs Peroxyacetyl Nitrates
PAW Projector Augmented-Wave
PBE Perdew–Burke–Ernzerhof Functional
PBE-D2 Perdew–Burke–Ernzerhof Functional with Grimme's D2 Dispersion Correction
PBE-D3(BJ) Perdew–Burke–Ernzerhof Functional with Grimme's D3 (Becke–Johnson damping) Dispersion Correction
PBE0 Perdew–Burke–Ernzerhof Hybrid Functional
PCA Principal Component Analysis
PES Potential Energy Surface
PM Particulate Matter
$PM_{2.5}$ Particulate Matter with an aerodynamic diameter smaller than 2.5 μm
PtP Point-to-Point
PW91 Perdew–Wang 1991 Functional
PW91+vdW Perdew–Wang 1991 Functional with Empirical van der Waals Correction

QM/MM Quantum Mechanics/Molecular Mechanics

RDF Radial Distribution Function
RMSE Root Mean Square Error

SCF Self-Consistent Field
SCM Static Corrugation Model
SCR Selective Catalytic Reduction
SNCR Selective Non-Catalytic Reduction
SO_2 Sulfur Dioxide
SOAP Smooth Overlap of Atomic Positions
SP Spin-Polarized
STM Scanning Tunneling Microscopy

TGCC Très Grand Centre de Calcul du CEA
TPD Temperature-Programmed Desorption
TS Transition State

UMAP Uniform Manifold Approximation and Projection for Dimension Reduction

V-Const Constrained Molecular Dynamics at Uniform Velocity
VASP Vienna *Ab Initio* Simulation Package
vdW-DF van der Waals Density Functional Incorporating Nonlocal Correlation Effects
VOC Volatile Organic Compound

XPS X-ray Photoelectron Spectroscopy

ZPE Zero-Point Energy

ænet Atomic Energy Network Package

CHAPTER 1

Introduction and State of the Art

Outline of the current chapter

1.1 The role of NOx oxidation in atmospheric pollution **5**

1.2 NO Removal and Control Strategies **10**

1.3 Carbon-Based catalysts for NO oxidation: Mechanisms and Challenges **14**

1.4 Advanced computational methods for NO oxidation on carbon-based catalysts **18**

1.5 Scientific challenges and thesis objectives **20**

1.6 Thesis Structure **21**

In our everyday environment, molecular interactions with solid surfaces sustain a wide array of natural and anthropogenic processes. For instance, the degradation of building materials exposed to urban air pollution is often driven by reactive gases interacting with mineral or metallic surfaces, while the formation and transformation of atmospheric aerosols are intimately linked to heterogeneous surface chemistry. In the atmosphere,

such processes influence the lifetimes and reactivity of trace gases like ozone (O_3), nitrogen oxides (NO_x), and volatile organic compounds (VOC), thereby modulating air quality, oxidative capacity, and climate forcing.

Understanding the fundamental principles that govern molecule–surface interactions is therefore essential not only for environmental and atmospheric chemistry but also for the rational design of materials for pollution control. In particular, reactions occurring at gas–solid interfaces are central to heterogeneous catalysis, a field of enormous industrial and environmental importance. In the context of atmospheric pollution, catalytic processes at particulate surfaces can alter reaction pathways and rates compared to the gas phase, often enabling or enhancing reactions that would otherwise be negligible. A detailed atomistic understanding of how molecules interact with, scatter from, or react on surfaces is thus key to modeling and controlling such phenomena.

From a theoretical perspective, gas–surface interactions can be categorized into three primary classes: *adsorption*, *desorption*, and *scattering* [1], where each interaction plays a distinct role in determining the fate of molecules at interfaces. These processes may be reactive or non-reactive, depending on the molecular species, the surface composition, and the energy redistribution among translational, rotational, vibrational, electronic degrees of freedom and phonons. Their detailed characterization at the atomic scale forms the basis for predicting macroscopic observables such as sticking coefficients, scattering distributions, and catalytic conversion rates.

Adsorption refers to the process by which a molecule from the gas phase collides with a surface and becomes temporarily or permanently bound to it. As illustrated in Figure 1.1, this process may proceed through different mechanistic pathways. In *dissociative chemisorption* [Figure 1.1(a)], the molecule breaks apart upon adsorption, resulting in separate atomic fragments anchored to distinct surface sites. In *physisorption* [Figure 1.1(b)], the interaction is mediated by weak van der Waals forces, typically associated with low adsorption energies (below 0.4 eV) and minimal perturbation of the electronic structure [2]. This reversible process is particularly relevant in atmospheric scenarios involving non-reactive gases interacting with soot or aerosol surfaces. In contrast, *chemisorption* involves the formation of covalent or ionic bonds between the adsorbate and the substrate [Figure 1.1(b)], often accompanied by significant energy release and electronic rearrangement. The binding energies involved range from 0.4 eV to over 8 eV, and the process is usually irreversible under ambient conditions. Finally,

in the *abstraction mechanism* [Figure 1.1(c)], one fragment remains adsorbed while the other is ejected back into the gas phase, an important pathway in surface-mediated radical reactions and selective functionalization processes.

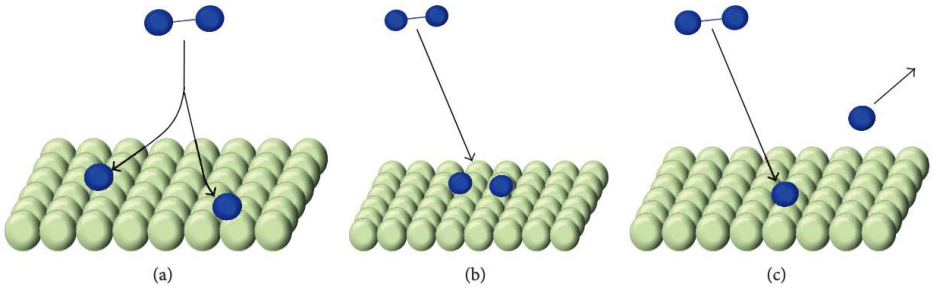

Figure 1.1: Representative adsorption mechanisms: (a) dissociative adsorption; (b) molecular adsorption (physisorption and chemisorption); (c) abstraction. (Source C. Díaz [1])

When adsorption does not occur, the incident molecule may instead scatter off the surface, as illustrated in Figure 1.2. *Elastic scattering* [Figure 1.2(a)] preserves both the translational energy and the internal state of the molecule, whereas in *inelastic scattering*, part of the incident energy is exchanged with the surface [3]. This energy transfer can manifest as *rotational* [Figure 1.2(b)] or *vibrational* [Figure 1.2(c)] excitation of the molecule but also by a variation in the translational energy of the scattered molecule, or by a combination of these different processes. Under certain conditions, coherent interference effects arising from the periodic structure of the surface may also produce discrete *diffraction patterns* [Figure 1.2(d)], which provide valuable information about surface corrugation, symmetry, and energy exchange mechanisms [4].

Finally, desorption, the release of products or unreacted species from the surface, frequently occurs as the final step of a reactive sequence. As illustrated in Figure 1.3, three primary mechanisms have been identified for surface reactions leading to desorption. The *Langmuir–Hinshelwood* (L-H) mechanism [5, 6] [Figure 1.3(a)] involves two adsorbed species diffusing on the surface and reacting. In the *Eley–Rideal* (E-R) mechanism [7, 8] [Figure 1.3(b)], a gas-phase molecule reacts directly with a surface-adsorbed species. Lastly, the *hot-atom* (H-A) mechanism [9] [Figure 1.3(c)] describes a scenario in which a recently adsorbed atom retains excess kinetic energy and reacts before reaching thermal equilibrium with the surface.

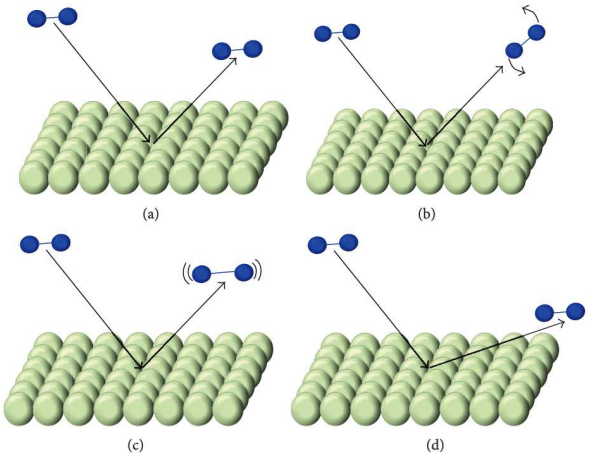

Figure 1.2: Representative scattering mechanisms: (a) elastic scattering; (b) rotational excitation; (c) vibrational excitation; (d) diffraction. (Source C. Díaz [1])

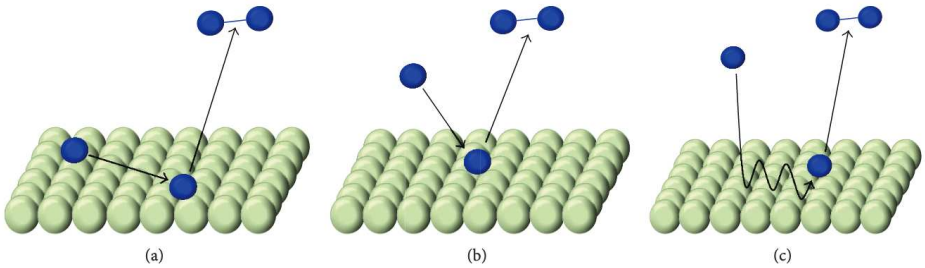

Figure 1.3: Surface-mediated reaction mechanisms: (a) Langmuir–Hinshelwood; (b) Eley–Rideal; (c) hot-atom mechanism. (Source C. Díaz [1])

While the elementary mechanisms of adsorption, scattering, and desorption are universally relevant in gas–surface dynamics, their manifestation in atmospheric environments carries particular importance. Processes such as reactive uptake, molecular trapping, and product desorption at aerosol and soot surfaces can significantly alter the lifetimes, transport, and transformation pathways of trace gases [10]. In this context, surface-mediated interactions modulate the oxidative capacity of the atmosphere and contribute to the formation of secondary pollutants. Among these mechanisms, heterogeneous oxidation processes play a particularly critical role. In this thesis, we investigate in detail one such process of particular environmental relevance: the surface-mediated oxidation of nitric oxide (NO) to nitrogen dioxide (NO_2) on oxidized soot or graphite

substrates. This transformation, which has important implications for both pollution chemistry and atmospheric modeling, will be examined in detail in the following sections.

1.1 The role of NOx oxidation in atmospheric pollution

Before delving into the specifics of NO oxidation, it is essential to ask: why study this reaction in the context of the atmosphere? Atmospheric pollution is one of the most pressing environmental challenges of the modern era, with far-reaching consequences for air quality, human health, and climate (Figure 1.4). The continuous growth in emissions from industrial activities, transportation, and energy production has led to elevated concentrations of key atmospheric pollutants, including NO_x, sulfur dioxide (SO_2), volatile organic compounds (VOC), and particulate matter (PM). These pollutants undergo complex transformations in the atmosphere, contributing to the formation of harmful secondary species such as ground-level ozone and fine particulate matter, which are strongly associated with adverse health effects and environmental degradation. Among these pollutants, NO_x plays a particularly influential role in atmospheric chemistry, acting as precursors in photochemical reactions that shape the dynamics of pollution and air quality on regional and global scales.

For example, NO plays a central role in the formation of photochemical smog. In this sunlight-driven phenomenon, NO_2 photolyzes to produce atomic oxygen, which rapidly combines with O_2 to form O_3. Ozone, in turn, reacts with NO, regenerating NO_2 and O_2:

$$NO + O_3 \longrightarrow NO_2 + O_2 \tag{1.1}$$

In a simplified atmospheric system containing only NO, NO_2, and O_3, and in the absence of VOC or carbon monoxide (CO), a photostationary steady state is established: NO_2 photolysis creates O_3, but NO rapidly consumes it, preventing net ozone accumulation [11].

However, the introduction of VOC and CO dramatically alters the chemistry. Under sunlight, these species react to form reactive organic radicals such as hydroperoxy (HO_2) and organic peroxy radicals, which efficiently oxidize NO to NO_2 without consuming O_3 [11]. This pathway decouples the photostationary cycle, allowing NO_2 to photolyze and generate additional ozone. As a result, ozone begins to accumulate rather than merely

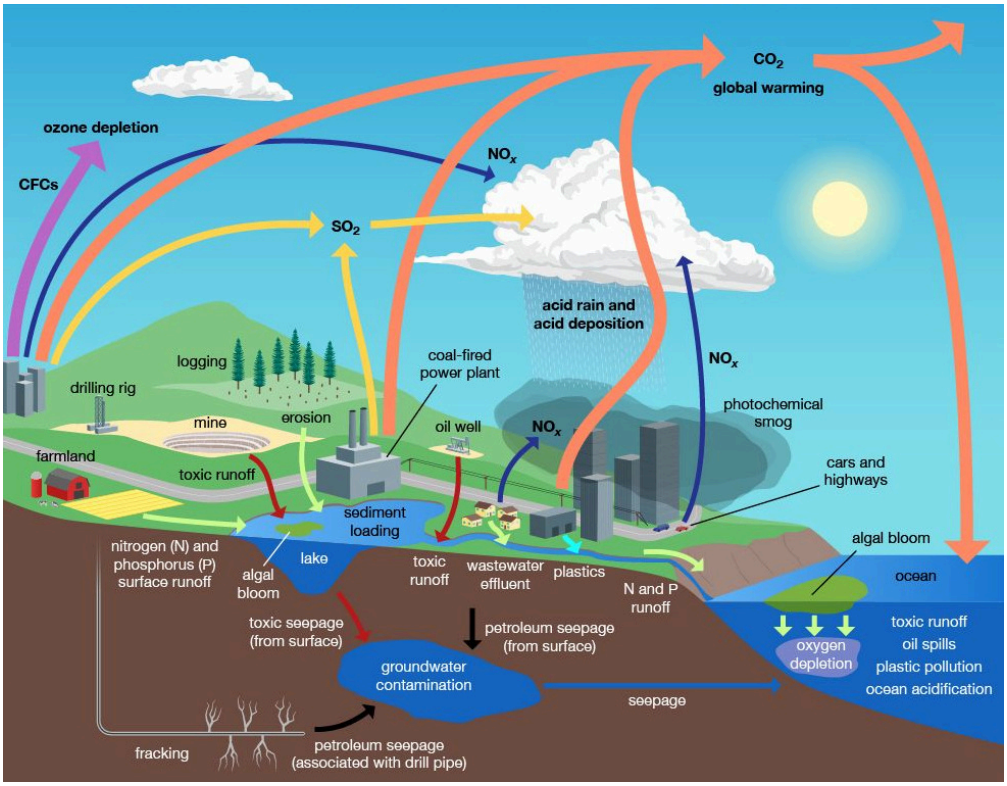

Figure 1.4: An infographic illustrating various sources and effects of environmental pollution, including NO_x emissions, acid rain, ozone depletion, and water contamination. (Source: *Britannica*)

cycle, as NO is continually converted to NO_2 by VOC-derived radicals instead of by ozone itself. Meanwhile, VOC are oxidized into a variety of secondary products (e.g., formaldehyde, acetaldehyde, peroxyacetyl nitrates (PANs), and other oxygenated organic compounds), some of which further participate in ozone formation or form secondary organic aerosols [11].

Through these chain reactions, a single NO molecule can ultimately facilitate the formation of multiple ozone molecules before the NO_x and radical chain is terminated. The resulting mixture of secondary pollutants—including ozone, NO_2, PANs, and fine particulate matter (secondary organic aerosol)—along with their precursors (NO_x, VOC, CO), constitutes photochemical smog (see Figure 1.5). Notably, soot (black carbon) is not a VOC but acts as a precursor or modulator in this context; particulate matter, including $PM_{2.5}$, can serve as a sink for free radicals such as HO_2, thereby influencing ozone

production dynamics [11]. This pollution phenomenon was first notable in Los Angeles in the mid-20th century, where strong sunlight and heavy NO_x and hydrocarbon emissions led to extreme levels of ozone. Historical smog episodes saw ozone concentrations high enough to cause acute eye and respiratory irritation, reduced lung function, and emergency hospital visits [11].

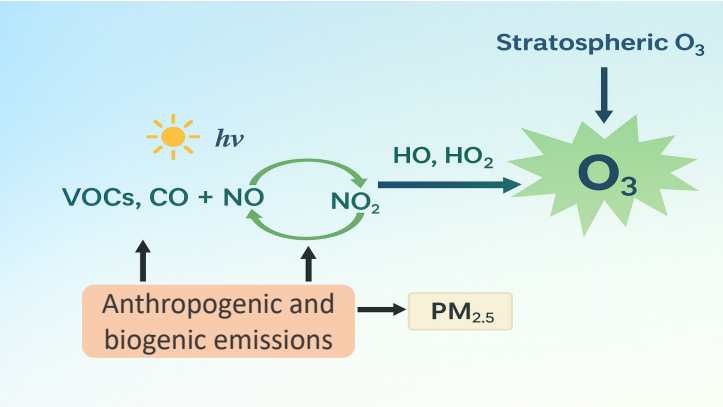

Figure 1.5: Simplified schematic representation of tropospheric ozone formation involving NO_x, VOC, and CO under solar radiation. Photolysis of NO_2 produces ozone, while reactive radicals (HO, HO_2) from VOC and CO oxidation enhance NO to NO_2 conversion without ozone consumption, leading to net ozone accumulation. $PM_{2.5}$ and stratospheric exchange also influence background ozone levels. (Adapted from *Zhang et al.* [11]).

In addition, these chemical reactions also contribute to the formation of acid rain, broadly defined as any form of precipitation (rain, snow, fog, etc.) with acidic components such as sulfuric acid or nitric acid [12]. These acid precipitation lead to soil and water acidification, damaging ecosystems. For example, it leaches essential base cations (like calcium and magnesium) from soils, altering their pH and releasing aluminum ions that can be toxic to plant roots and aquatic organisms [13]. Freshwater lakes and streams in sensitive regions (without limestone buffering) have experienced pH drops due to acid rain, resulting in the decline or disappearance of fish populations and other aquatic life [13]. In the 1970s and 1980s, many lakes in North America and northern Europe became too acidic to support fish, largely due to deposition of sulfate and nitrate from industrial emissions [13].

In this context, the contribution of NO_x to acid rain formation arises primarily from the reaction of NO_2 with atmospheric water vapor. This reaction yields nitrous

acid (HNO_2) and nitric acid (HNO_3) [14, 15], both of which are strong acids [12] characterized by high solubility. Consequently, these acids are effectively removed from the atmosphere through precipitation (wet deposition) or by direct deposition onto surfaces (dry deposition).

As previously indicated, heterogeneous processes occurring on aerosol surfaces significantly enhance the atmospheric availability of NO_2 and among these aerosol particles, soot particles are particularly important. They have the ability to catalyze the oxidation of NO into NO_2 [16] and also convert NO_2 to NO [17]. Therefore, understanding the catalytic role of soot in NO oxidation and NO_2 reduction in the atmosphere is essential.

1.1.1 Soot particles as NO oxidants and graphite as a model system

Soot particles are primarily form due to the incomplete combustion of hydrocarbons, where insufficient oxygen availability prevents complete oxidation. This is a multistep process involving a rich interplay of gas-phase chemistry, molecular growth, and particle dynamics. As illustrated schematically in Figure 1.6, the transformation begins in the lower region of the flame, where the combustion of fuel and oxidizer leads to the formation of reactive radicals. These small radicals react to form the first aromatic molecules, which further evolve into larger polycyclic aromatic hydrocarbons (PAHs) through sequential addition and ring fusion mechanisms. These species, commonly referred to as soot precursors, undergo nucleation processes that result in the formation of incipient carbonaceous particles. As the system progresses upward in the flame, these nascent particles coagulate and restructure into larger agglomerates and graphitic aggregates. Ultimately, in the outer flame zones, oxidative processes act upon the formed particles, yielding a variety of oxygenated products. Such conditions are commonly encountered in internal combustion engines, industrial burners, and biomass burning processes [18]. Structurally, soot is composed predominantly of elemental carbon arranged in disordered graphitic layers, along with variable amounts of adsorbed organic and inorganic species [19]. The physical and chemical complexity of soot reflects the wide variety of reactive intermediates and aggregation states generated under combustion conditions [20].

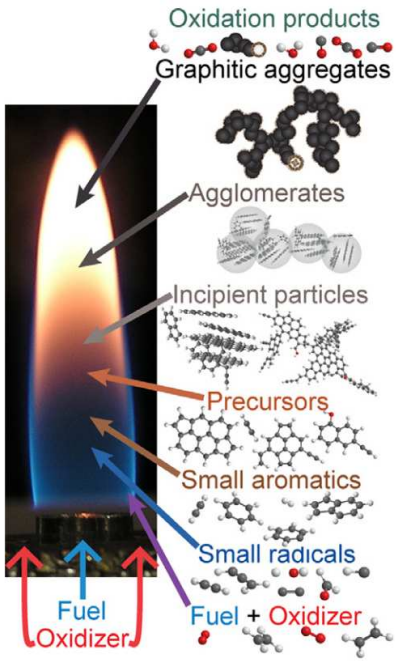

Figure 1.6: Schematic representation of soot formation and evolution in a flame. The molecular species and particles are not drawn to scale. (Source: *Michelsen et al.* [21])

The presence of soot in the atmosphere poses substantial environmental and health implications. Environmentally, soot contributes to atmospheric warming due to its strong absorption of solar radiation, directly impacting radiative forcing. Through their atmospheric aging, are also known to be effective cloud condensation or ice nuclei, promoting the formation of persistent clouds and ultimately contributing negatively to radiative forcing [22]. Additionally, soot deposition accelerates melting when settled on ice and snow surfaces, reducing surface albedo and enhancing regional climate warming [23]. From a health perspective, inhalation of fine soot particles is associated with respiratory and cardiovascular diseases. Moreover, soot particles carry adsorbed hazardous compounds, notably PAHs, which are well-established carcinogens. Furthermore, soot's interactions with other atmospheric constituents complicate its impact on air quality and climate dynamics [24]

For example, oxidized soot surfaces are recognized as active catalytic sites for NO oxidation under atmospheric conditions. Specifically, oxygen-containing functional groups, such as hydroxyl (-OH) and carboxyl (-COOH) groups, enhance surface reactivity,

facilitating heterogeneous reactions that convert atmospheric NO into NO_2 [25, 26]. This catalytic activity has significant implications for urban air pollution, as previously mentioned , an accelerated conversion of NO to NO_2 affects tropospheric ozone formation and air quality [27].

Given the pronounced structural and chemical heterogeneity of soot, well-defined carbon materials, such as functionalized graphite or graphene, are frequently adopted as model systems for fundamental studies of gas–surface interactions [28–32]. Graphite offers a crystalline and electronically homogeneous substrate that enables rigorous control of surface chemistry and atomic structure. Although soot possesses a higher degree of disorder, both materials are composed of graphitic layers and share key reactive motifs, including edge sites and oxygen-containing functional groups. Numerous studies have demonstrated that these surface sites, whether in soot or graphite, act as primary centers for the adsorption and activation of atmospheric species such as NO, NO_2, SO_2 and O_3 [33, 34].

Therefore, employing functionalized graphite or graphene as a proxy provides a computationally tractable model to dissect elementary mechanisms that would be challenging to isolate within the complex, amorphous structure of soot. For example, Morisset et al. used density functional theory (DFT) calculations to study the hydrogenation [35] and oxygenation [36] reactions on the surface of graphene and coronene, respectively. They were able to determine adsorption energies that are comparable with the results of conventional experimental methods such as Temperature-Programmed Desorption (TPD). In this context, this theoretical framework allow for a detailed atomistic description of gas–surface interactions and also yields insights into mechanistic pathways.

1.2 NO Removal and Control Strategies

Another area of significant interest concerns the control of NO emissions into the atmosphere, given that NO acts as a primary precursor to a variety of harmful atmospheric species, as discussed earlier. From natural sources, such as wildfires and lightning, exist a contribution to NO_x formation; for example, lightning accounts for approximately 10–15 % of thermal NO_x derived from atmospheric nitrogen [37]. Also wildfires significantly emit NO_x, although many such events are human-induced. However globally, anthropogenic activities are the dominant source of NO_x emissions, about 48.8 Tg of

nitrogen (as NO_x) are emitted annually, with approximately 77 % originating from human activities, predominantly through fossil fuel combustion for power generation, transportation, industrial processes, and biomass burning [38]. Consequently, substantial efforts have been made to develop effective strategies for NO emissions control. In this context, removal methods are broadly categorized into non-catalytic and catalytic methods, where is considered that NO is relatively insoluble and chemically stable, therefore necessitating specialized post-combustion treatment.

1.2.1 Non-Catalytic NOx control methods

Non-catalytic methods for the removal of NO_x from exhaust gas streams are based on physical and chemical processes that do not require a heterogeneous catalyst. These technologies are particularly relevant in applications where catalytic materials are impractical due to cost, operating conditions, or potential catalyst poisoning. Among the most widely implemented non-catalytic approaches are wet absorption (scrubbing), which remove NO_x by transferring it into liquids [39, 40], typically involving chemical reactions. Given NO's low water solubility, scrubbing alone is ineffective unless NO is oxidized to NO_2, which readily forms nitrous and nitric acids in aqueous solutions [41]. Thus, industrial scrubbing systems often incorporate oxidants, such as ozone or hydrogen peroxide, or catalysts upstream to enhance effectiveness. This approach typically achieves moderate NO_x removal efficiencies (30–50 %), suitable primarily for lower NO_x concentrations due to the corrosive nature of the reagents involved.

Other methods are based on adsorption, using solid sorbents to capture NO, frequently as NO_2 or nitrate species [42]. Activated carbon [43], zeolites [44], and metal-organic frameworks (MOFs) [45] serve as adsorbents. Emerging materials such as MOFs offer advantages in adsorption capacity and selectivity by forming stable nitrite or nitrate complexes, significantly improving NO_x uptake compared to traditional materials. However, these methods still face limitations related to adsorption capacity and regeneration processes. In this context, an alternative is the Selective Non-Catalytic Reduction (SNCR) which is a thermal reduction process in which a nitrogenous reductant, usually ammonia or urea, is injected into flue gas at elevated temperatures (approximately 850–1100 °C) to chemically reduce NO_x to nitrogen (N_2) [46]:

$$4\,NH_3 + 4\,NO + O_2 \longrightarrow 4\,N_2 + 6\,H_2O \tag{1.2}$$

This method relies heavily on precise temperature control to avoid ammonia slip or incomplete reactions, resulting in moderate practical NO removal efficiencies (around 40–60 %). SNCR is cost-effective and suitable for smaller installations or where catalytic methods are impractical.

1.2.2 Catalytic NOx control methods

Catalytic methods represent the most effective and widely applied technologies for the abatement of NO_x in exhaust streams from combustion processes. These approaches rely on the use of heterogeneous catalysts—typically based on noble metals or transition metal oxides—to promote selective chemical reactions that convert NO_x into environmentally benign products such as N_2 and water. Key catalytic processes include selective catalytic reduction (SCR) which is widely implemented in industrial and automotive exhaust systems, employing ammonia (or urea) injected into exhaust streams at temperatures ranging from 250–450 °C in the presence of a catalyst [47, 48]. This method significantly outperforms SNCR regarding efficiency and operational flexibility, achieving high NO_x conversion rates to N_2 and water. SCR efficiency further benefits from upstream oxidation processes, such as those facilitated by diesel oxidation catalysts (DOCs), primarily based on platinum-group metals (e.g., Pt/Pd), which catalyze the oxidation of NO to NO_2 in diesel engine exhaust streams [49, 50]. The generated NO_2 significantly enhances downstream SCR reaction efficiency through the fast SCR mechanism. Additionally, NO_2 facilitates soot oxidation in diesel particulate filters (DPFs) at lower temperatures (250–400 °C) compared to oxidation with oxygen alone, highlighting the critical importance of NO oxidation processes [51].

Anoter methods are Lean NO_x traps (LNT), or NO_x adsorbers, which operate by initially storing NO_x as nitrates under lean (oxygen-rich) conditions [52, 53]. The stored nitrates are periodically regenerated under rich (fuel-rich) conditions, releasing NO_x and subsequently converting it to N_2. Although LNT systems achieve notable NO_x removal efficiencies (70–90 %), they require careful periodic regeneration cycles and effective sulfur management strategies to maintain optimal performance. In general, the high efficiency, selectivity, and operational flexibility of catalytic methods have made them

indispensable for meeting stringent emission standards in both stationary and mobile sources.

1.2.3 Role of NO oxidation in emission control methods

Taking into account that NO constitute typically 90–95 % of NO_x emissions from combustion sources [41], oxidizing NO to NO_2 is a crucial step in various NO control strategies, enhancing reactivity, improving solubility, and permitting lower operating temperatures. In this context, NO_2 is thermodynamically more reactive than NO, often making its formation the rate-determining step in subsequent reactions. For example, in SCR, the presence of NO_2 facilitates the "fast SCR" reaction pathway, which requires approximately equal proportions of NO and NO_2, proceeding approximately ten (10) times faster and at lower temperatures than standard SCR [54]:

$$4\,NH_3 + 2\,NO + 2\,NO_2 \longrightarrow 4\,N_2 + 6\,H_2O \tag{1.3}$$

Additionally, in wet scrubbing processes, NO_2 readily dissolves in water to form nitric acid (HNO_3) or nitrous acid (HNO_2), markedly increasing aqueous removal efficiency. Furthermore, DPFs exploit NO oxidation to NO_2 for passive regeneration, as NO_2 oxidizes soot at substantially lower temperatures than oxygen alone, enabling efficient particulate filter cleaning without requiring high-temperature regeneration cycles.

Among various methods for converting NO into NO_2, carbon-based materials have been identified as effective oxidants, particularly under ambient conditions [55, 56]. When NO is oxidized on the carbon surface, the resulting NO_2 or nitrate species can be strongly adsorbed and subsequently processed either by catalytic reduction or chemical capture. This dual functionality, adsorption and catalytic conversion, is enhanced by appropriate surface functionalization. For example, acid treatments can increase the density of oxygen-containing groups (e.g., phenols, quinones, lactones, and epoxies), which serve as active sites for NO adsorption and oxidation [57, 58]. Similarly, nitrogen doping alters the electronic properties of the carbon matrix and facilitates the formation of reactive oxygen species, thereby promoting NO oxidation at lower temperatures compared to many metal oxide catalysts [59].

1.3 Carbon-Based catalysts for NO oxidation: Mechanisms and Challenges

Various carbonaceous materials, including activated carbons (ACs) [60–62], carbon fiber cloths [63, 64], carbon nanotubes (CNTs) [65] and carbon xerogels [66], have been extensively investigated as room-temperature catalysts for NO oxidation. Experimental studies have demonstrated that the NO oxidation activity of these materials strongly depends on their surface chemistry. In particular, the introduction of surface oxygen complexes is critical because these groups not only enhance NO adsorption but also provide the necessary oxygen for its conversion to NO_2 [67]. For example, experiments have shown that activated carbons with increased oxygen functionalities, achieved through treatments like nitric acid oxidation, exhibit enhanced NO to NO_2 conversion [55]. Complementary theoretical analyses using DFT have supported these findings by revealing that epoxy and other oxygenated groups can lower the activation energy for NO oxidation [68–70].

1.3.1 Insights from experimental investigations

Experimental investigations into the adsorption and oxidation of NO on carbonaceous surfaces have provided valuable insights; however, they often fall short in elucidating detailed reaction pathways and dynamic processes. Factors such as the complexity of carbon based surfaces, variations in surface morphology, and the presence of multiple functional groups hinder the ability of experimental techniques to capture transient intermediates and energy transfer mechanisms [71]. Additionally, discrepancies in reported kinetic parameters highlight the limitations of purely experimental approaches in fully characterizing the NO oxidation process.

In seminal studies in the 1990s, Mochida et al. [60] suggested that NO is first adsorbs on the carbon surface and then associates with an oxygen atom, from either the surface or O_2, to form a surface complex denoted [NO–O–NO_2] which after will decompose, releasing NO_2. A crucial aspect of Mochida's mechanism is that NO adsorption is rate-limiting, and neither NO dimers N_2O_2 nor N_2O_3/N_2O_4 need to form in the process. Their kinetic data support this scheme, implying a relatively direct oxidation pathway, however, this mechanism doesn't explicitly clarify the source of the oxygen in the [NO–O–NO_2] mechanism and very little NO can be physically adsorbed, since it is

considered a supercritical gas at ambient temperature. Indeed, the presence of functional oxygen groups in graphite significantly increases the adsorption probability of NO on its surface [72] and its subsequent oxidation [59].

Building on Mochida's work, Adapa et al. [62] proposed two possible routes: a L-H mechanism where NO and O_2 both adsorb on carbon and then react, and an E-R mechanism where gaseous NO reacts with pre-adsorbed oxygen on carbon. In the L-H scenario, they envisioned O_2 dissociating on carbon to provide atomic O that then oxidizes an adsorbed NO. In the E-R scenario, NO from the gas directly attacks an O_2 molecule that is transiently bounded or an oxygen functional group on carbon. Both mechanisms involve similar surface intermediates (C–NO, C–NO_2, C–NO_3, etc.), but differ in whether O_2 needs to stick to the surface first.

Rathore et al. [73] offered a multi-step mechanism: first, NO and O_2 co-adsorb on adjacent active sites of carbon, then being oxidized to form an adsorbed NO_2 (bounded to the carbon) which will reacts further with another NO (or O) to form surface nitrates like –NO_3 or a NO–NO_3 complex. Finally, the NO–NO_3 (essentially a surface N_2O_5 analogue) decomposes to release an NO_2 molecule. This mechanism suggests multiple surface intermediates (NO_2^*, NO_3^*, $N_2O_5^*$ type species) and aligns with observations of nitrate species on carbon after exposure to NO/O_2. It also rationalizes a slow step for NO_2 desorption—if NO_2 is stuck as nitrate until a second NO comes to knock it off as NO_2 gas, that could be rate-limiting.

In another mechanistic analysis, Sousa et al. [66] emphasize that the catalytic oxidation of NO on activated carbon proceeds predominantly through the participation of chemisorbed oxygen species. Their experiments show that NO from the gas phase reacts with oxygen adsorbed on the carbon surface, leading to the formation of chemisorbed NO_2, which subsequently desorbs as NO_2 gas. This sequence is supported by TPD analyses, which reveal distinct NO and NO_2 desorption profiles corresponding to different types of surface-bound intermediates. The presence of molecular oxygen in the reaction environment is crucial, as it replenishes the surface oxygen species required for sustained NO oxidation. Kinetic data indicate that while NO adsorption is the rate-limiting step, the desorption of NO_2 is essential for regenerating active sites and maintaining catalytic activity.

In a complementary study, Sousa et al. [59] demonstrated that doping activated carbon with nitrogen significantly enhances its catalytic performance in NO oxidation. Their

results highlight the key roles of specific nitrogen functionalities—particularly *pyridinic* and *pyrrolic* groups—in modulating the surface chemistry. Pyridinic nitrogen atoms, located at the edges of graphene layers, act as strong Lewis bases and increase the basicity of the carbon surface. Pyrrolic nitrogen, incorporated within five-membered rings, contributes to the π-electron system. The incorporation of these nitrogen groups increases the electron density and basic character of the carbon surface, thereby promoting the activation of oxygen species and facilitating the oxidation of NO to NO_2.

In contrast, Ahmed et al. [74] investigated NO reduction over activated carbons and observed that surface oxygen plays a pivotal role in facilitating NO conversion. Their work showed that treating activated carbons with sulphuric acid increased the concentration of acidic oxygen-containing groups such as carboxyl and carbonyl functionalities on the surface. These groups, acting as adsorbed oxygen reservoirs, directly participate in the oxidation of NO to NO_2. They found that an increase in the gas-phase oxygen concentration significantly enhances NO conversion, which indicates that the oxygen adsorbed on the carbon surface is the active oxidant. Their experimental results underline a mechanism where the adsorbed oxygen directly transfers to NO, a process that becomes more efficient with higher oxygen content on the catalyst surface, thereby confirming the critical role of surface-bound oxygen in the oxidation reaction.

Overall, these contrasting experimental studies highlight unresolved debates about NO adsorption mechanisms (physisorption versus chemisorption) and the nature and role of surface functionalities. These ambiguities necessitate complementary computational investigations.

1.3.2 Theoretical Perspectives and Computational Approaches

Computational approaches, particularly those based on DFT, have emerged as indispensable tools for probing the atomic-level interactions of NO with carbonaceous surfaces. These methods enable the calculation of adsorption energies, reaction barriers, and electron transfer mechanisms that are challenging to access experimentally [68–70].

Early DFT work by Tang and Cao [69] examined NO_x (NO, NO_2, N_2O_3/ N_2O_4) adsorption on pristine graphene vs. graphene oxide. They found that graphene by itself binds NO only weakly, whereas oxygen-functionalized graphene (graphene oxide) can bind NO or NO_2 much more strongly due to the presence of active sites (epoxy, hydroxyl,

defect sites). This suggested that perfect graphitic basal planes are relatively inert to NO, but the common functional defects in real carbons provide adsorption sites necessary for catalysis.

Building on this, later DFT studies from Hou et al. [70] incorporated van der Waals corrections for better accuracy, they performed DFT-D3 calculations on graphene oxide (GO) models to investigate the adsorption and oxidation of NO on this surface. They showed that NO can adsorb at epoxy or carbonyl groups and that an NO + O_{epoxy} reaction to form an NO_2 group on the surface is thermodynamically favorable. They also identified a low energy barrier of approximately 0.1 eV for the direct reaction of the NO with the epoxy oxygen on the graphene oxide, however, this barrier corresponds to roughly 1160 K, yet this barrier is too high to obtain oxidation at room temperature, as has been shown by experimentalists [60, 62, 74]. Similarly, Cen et al. [68] investigated the oxidation of NO by epoxy groups on graphene oxide, with particular emphasis on the effect of neighboring hydroxyl functionalities. Their DFT calculations reveal that the presence of adjacent hydroxyl groups significantly enhances the adsorption and activation of NO on the GO surface. This promotion arises from the formation of a strong N–C covalent bond between the NO molecule and the graphene oxide, which leads to a marked reduction in the oxidation barrier.

1.3.3 Unresolved challenges and knowledge gaps

Numerous experimental studies have presented differing and sometimes conflicting mechanisms for NO adsorption and oxidation on carbon-based materials. Some investigations emphasize that physical properties—particularly micropore structures—dominate NO oxidation kinetics [75, 76], while others underscore the critical role of surface chemistry, especially oxygen-containing functional groups [59, 62, 74]. Moreover, significant uncertainty remains regarding the nature of NO oxidation mechanism and the role of the oxygen atom; some studies suggest predominant physisorption on basal planes, while others highlight significant chemisorption at defect sites. Clarifying these conflicting interpretations requires complementary computational analyses capable of resolving these mechanistic ambiguities.

Despite valuable theoretical insights, such as the identification of epoxy groups as crucial for lowering activation energies for NO oxidation, current computational ap-

proaches have notable limitations. Most computational studies have not thoroughly addressed the kinetic mechanism of catalytic NO oxidation, often neglecting critical factors such as surface temperature and dynamic reaction conditions. These simplifications lead to discrepancies between theoretically predicted activation barriers and experimentally observed room-temperature activities, indicating a gap in existing computational frameworks.

Therefore, further theoretical research is essential. Specifically, advanced computational techniques, such as ab initio molecular dynamics simulations, are needed to accurately represent more realistic surface conditions, temperature effects, and dynamic interactions between NO and functionalized carbon surfaces. Addressing these challenges will facilitate a deeper mechanistic understanding and enable the rational design of more efficient carbon-based NO oxidation catalysts.

1.4 Advanced computational methods for NO oxidation on carbon-based catalysts

To address the limitations and unresolved challenges highlighted previously, advanced computational methods combining accuracy and efficiency are crucial. The theoretical study of gas–surface processes can be approached using a variety of computational strategies, which broadly fall into two categories: static calculations and molecular dynamics simulations. Static approaches are typically used to compute optimized structures, adsorption energies, or reaction barriers, while dynamic methods allow for the explicit simulation of atomic motion over time. These include electronic structure methods—such as molecular and periodic implementations of DFT, including variants like Density Functional Tight Binding (DFTB) [77]—as well as hybrid multiscale approaches like quantum mechanics/molecular mechanics (QM/MM) [78]. Other strategies rely on empirical or reactive force fields [79]. When well-established force fields are available, they can be readily applied; otherwise, new parametrization must be developed, typically based on high-level electronic structure calculations. In recent years, machine learning techniques have also emerged as promising tools for modeling gas–surface interactions, particularly when combined with DFT-based *ab initio* molecular dynamics (AIMD) simulations [80].

Among the available computational methodologies, AIMD constitutes a particularly powerful approach for investigating gas–surface interactions. In AIMD, energies and forces are determined *on the fly* at each simulation step through electronic structure calculations, most commonly based on DFT. This framework enables a consistent and simultaneous treatment of all atomic degrees of freedom within the system, encompassing both the adsorbate and the surface atoms. By explicitly accounting for the motion of the substrate, AIMD captures essential physical phenomena such as energy exchange, thermal fluctuations, and the dynamic evolution of chemical bonds and charge distribution during the interaction process. In the context of NO oxidation on oxidized graphite surfaces, this atomistic and electronic resolution provides direct access to the underlying mechanisms and dynamical pathways, offering detailed insight into the reactivity and catalytic properties of the interface.

However, despite its powerful capabilities, AIMD remains computationally demanding because it requires performing detailed electronic structure calculations at every simulation step, typically on the order of femtoseconds or less. This high computational cost drastically limits the achievable simulation length and the number of feasible trajectories, which in turn affects the statistical accuracy and reliability of important dynamical properties. Properties such as reaction probabilities, angular and energy distributions, surface diffusion events, and detailed energy transfer mechanisms can therefore suffer from limited sampling, reducing the overall precision and interpretability of AIMD-derived data. Consequently, these computational constraints have prompted researchers to explore alternative methodologies capable of providing AIMD-level accuracy at significantly reduced computational expense.

To overcome these computational limitations, recent advancements in artificial intelligence (AI) methods [81–85] have led to the development of machine learning potential energy surfaces (ML-PES). These ML-PES approaches utilize AI algorithms to accurately capture complex, multidimensional energy and force landscapes derived from electronic structure calculations, typically at the DFT level. By drastically reducing computational demands, ML-PESs enable extensive molecular dynamics simulations with significantly enhanced statistical sampling. This improvement allows for a greater number of trajectory calculations, ensuring robust statistical accuracy while preserving near-ab initio precision. Consequently, ML-PES methods effectively bridge the gap between the accuracy of AIMD and the computational efficiency of classical force fields,

establishing themselves as an essential tool for comprehensive studies of gas-surface interactions [86–91].

The development of a ML-PES for NO oxidation on oxidized graphite surfaces involves training AI models using extensive datasets generated from electronic structure calculations, typically at the DFT level. For this system, utilizing AIMD-generated DFT trajectories provides distinct advantages. Unlike traditional approaches that construct potential energy surfaces using predefined grids, AIMD naturally samples configurations most relevant to realistic catalytic conditions, including varied NO-surface interaction geometries, dynamic surface reconstructions, and temperature fluctuations. Consequently, integrating AIMD simulations with ML-PES construction represents a highly effective and state-of-the-art methodology for capturing complex reaction dynamics accurately and efficiently, making it an essential tool for studying NO oxidation mechanisms at functionalized carbon surfaces.

In summary, integrating AIMD with ML-PES represents a transformative step forward in computational modeling of NO oxidation on functionalized graphite surfaces.

1.5 Scientific challenges and thesis objectives

Within this scientific context, discrepancies between previous theoretical predictions [68, 70] and experimental observations [59, 60, 62] represent the primary motivation for undertaking new theoretical investigations using advanced simulation methodologies. These discrepancies particularly pertain to the role and influence of dynamics and surface temperature effects on NO oxidation processes over oxygen-functionalized carbon-based materials. Addressing these knowledge gaps requires sophisticated computational approaches capable of capturing dynamic reaction pathways and accurately describing molecule-surface energy transfer mechanisms at the atomic level.

Thus, the general aim of this thesis is to provide a refined theoretical description of the oxidation mechanisms of NO on oxygen-functionalized carbon-based surfaces, notably oxidized graphite, through advanced molecular dynamics computational methods. The primary objective is to use AIMD and ML-PES to:

- Characterize the role of surface thermalization on the reaction probabilities for NO oxidation on oxidized highly oriented pyrolytic graphite (O@HOPG) surface.

- Elucidate detailed reaction pathways by identifying transient intermediates and characterizing structural requirements for the oxidation reaction under realistic conditions.
- Analyze the mechanisms of energy transfer between incoming NO molecules and surface functionalities, emphasizing translational, rotational, and vibrational energy redistributions.
- Generate and utilize ML-PES to extend the scale of molecular dynamics simulations, enhancing the statistical reliability and computational efficiency of the theoretical investigations.

To achieve these objectives, the approach will initially involve precise AIMD simulations that naturally sample the relevant configurational space, providing accurate electronic structure data. Subsequently, these AIMD-generated data sets will train advanced machine learning models to construct accurate ML-PES. This combined methodology will enable larger-scale simulations of gas-surface interactions at reduced computational costs, without sacrificing the accuracy of ab initio methods.

1.6 Thesis Structure

The present thesis is organized into five main chapters, each of which addresses a specific aspect of the theoretical and computational investigation of gas–surface interactions and NO oxidation on oxidized graphite. The structure reflects the progressive development of both the scientific questions and the methodologies employed throughout this work.

- **Chapter 1: Introduction and State of the Art**
  This chapter introduces the scientific context and motivation of the study, with particular emphasis on the role of nitrogen oxides in atmospheric chemistry and the relevance of it oxidation processes on carbonaceous materials. It presents an overview of experimental and theoretical advances in the field, identifies unresolved challenges, and defines the main objectives of the thesis.
- **Chapter 2: Modeling Gas–Surface Interactions with *Ab Initio* Molecular Dynamics**

The second chapter details the theoretical foundations and practical implementation of AIMD methods for modeling gas–surface interactions. It discusses the quantum mechanical underpinnings, the role of DFT in describing potential energy surfaces, and the specific computational strategies adopted to study NO–graphite systems.

- **Chapter 3: NO Oxidation on O@HOPG from AIMD Approach**
  This chapter presents the application of AIMD simulations to the study of NO oxidation on O@HOPG. It describes the construction and validation of the surface models, the sampling of dynamical initial conditions, and a comprehensive analysis of the resulting reaction mechanisms and energy redistribution processes.

- **Chapter 4: Machine Learning-Based Potential Energy Surfaces for Studying Gas–Surface Interactions**
  The fourth chapter introduces the development and validation of ML-PES, focusing on neural network approaches tailored to high-dimensional systems. Methodologies for training, assessing, and deploying ML-PESs are detailed, and their integration into molecular dynamics simulations is discussed.

- **Chapter 5: NO Oxidation on O@HOPG: MD Simulations with a ML-PES**
  In the final scientific chapter, the ML-PES developed in Chapter 4 is employed to perform large-scale molecular dynamics simulations of the NO oxidation reaction, mirroring the initial conditions of the AIMD trajectories. The comparative analysis highlights both the strengths and current limitations of ML-driven simulations relative to first-principles dynamics.

- **Conclusions and Perspectives**
  The thesis concludes with a synthesis of the principal findings and their broader implications for the field of gas–surface reactivity. Future directions are outlined, emphasizing the methodological innovations and potential applications to related systems in atmospheric and materials science.

Overall, this manuscript aims to advance the fundamental understanding of heterogeneous NO oxidation at carbonaceous interfaces, bridging *ab initio* accuracy and machine learning efficiency, and providing insight into mechanisms relevant for atmospheric chemistry and environmental technologies.

CHAPTER 2

Modeling Gas–Surface Interactions with Ab Initio Molecular Dynamics

Outline of the current chapter

2.1 Why Ab Initio Molecular Dynamics? 24

2.2 Quantum foundations and principles of AIMD 27

2.3 DFT as the workhorse of AIMD 30

2.4 Molecular dynamics on a quantum PES 48

This chapter is dedicated to the theoretical foundations of AIMD, the central method employed in this thesis to investigate gas–surface interactions at the atomic scale. It begins with a brief overview of the main computational approaches used to study molecular dynamics in such systems, outlining their respective advantages and limitations. The focus then shifts to AIMD, starting with key quantum mechanical concepts such as the time-independent Schrödinger equation and the Born–Oppenheimer approximation. DFT is introduced as the framework for computing electronic energies and forces, with emphasis on periodic implementations suited for extended surfaces. The chapter concludes

with the theoretical and practical elements of AIMD that underpin the simulation strategy adopted in this work.

2.1 Why Ab Initio Molecular Dynamics?

Several theoretical and computational methodologies have been developed to study gas-surface interactions [1, 2, 92, 93], broadly classified into methods employing preconstructed Potential Energy Surfaces (PES) [94–98] and on-the-fly molecular dynamics approaches.

Methods based on preconstructed PES involve first computing discrete points in the configuration space using quantum electronic structure calculations, typically DFT, and then interpolating these data points to obtain a continuous potential energy landscape. Within these interpolation methods, the Corrugation-Reducing Procedure (CRP) [99] and Modified Shepard (MS) interpolation [94, 95] stand out. CRP simplifies interpolation by subtracting atomic-scale corrugation, which predominantly arises from atom-surface interactions, thereby yielding a smoother potential suitable for efficient trajectory simulations. However, its extension to polyatomic molecules and highly reactive surfaces is challenging due to the rapidly increasing dimensionality of the PES. Modified Shepard interpolation partially addresses this issue by focusing computational effort on dynamically relevant regions, identified through classical trajectory simulations, allowing a reduction in required computational resources. Nonetheless, the limited treatment of surface degrees of freedom in these interpolative methods reduces their reliability when surface atom mobility plays an important role in the gas–surface interaction.

To address these limitations, various strategies have been developed to incorporate surface dynamics and temperature effects into interpolated PES models. One such approach is the Static Corrugation Model (SCM) [100], which introduces thermal disorder by displacing surface atoms from their ideal lattice positions and applying coupling potentials derived from DFT. More recently, this has been extended by generating thermally representative surface slabs using Embedded Atom Method (EAM) potentials that accurately reproduce surface properties such as phonon dispersion and lattice expansion [101]. Building upon this, the Dynamic Corrugation Model (DCM) couples CRP-derived PES with EAM-based surface dynamics, allowing real-time atomic motion of the surface during trajectory simulations and improving the realism of energy exchange processes.

In parallel, the CRP [99] and MS [102] methods have also been adapted to corrugated surfaces by introducing position-dependent parameters, enhancing its flexibility in capturing site-specific interactions. These hybrid approaches represent significant progress toward overcoming the static surface assumption of traditional interpolation schemes, enabling more accurate modeling of complex gas-surface systems where surface mobility plays a crucial role.

However, they also present notable limitation. For instance, in SCM, although surface slabs can be constructed to mimic thermal disorder, the atoms remain frozen during the simulation, limiting the ability to capture true dynamical energy exchange. More, the use of randomly displaced surfaces based on the Debye-Waller factor, may not always yield physically realistic configurations, further complicating validation [101]. DCM, which combines CRP with EAM potentials to introduce explicit surface atom motion, represents a significant advancement, but its overall accuracy has yet to be systematically validated across a wide range of systems [103]. Moreover, the CRP, even in its extended form with position-dependent parameters, remains difficult to apply to polyatomic adsorbates due to the exponential increase in required ab initio data and the complexity of the interpolation procedure [98]. The construction and parametrization of these surfaces can be labor-intensive and require a homogeneous grid of DFT points, which is computationally expensive.

In this context, on-the-fly molecular dynamics approaches, notably AIMD, overcome many of these limitations [104]. This simulation approach marries molecular dynamics with quantum mechanical calculations to track atomic motion with on-the-fly electronic structure evaluation. In AIMD, forces on atoms are computed directly from solving the electronic Schrödinger equation (often via DFT) at each time step, rather than using predetermined empirical force fields. This first-principles force calculation means that AIMD can naturally account for chemical bond breaking and formation, electronic polarization, and charge transfer during the dynamics. Car and Parrinello's seminal 1985 work [105] introduced this unified MD-DFT scheme, demonstrating that it extends molecular dynamics beyond simple pair potentials (enabling simulation of covalently bonded or metallic systems) and allows DFT to treat larger, finite-temperature systems than previously feasible. Recently, AIMD has also been extended to account for dissipation by electron hole pair excitations via an effective electronic friction model (AIMDef) for metallic surfaces[106].

Despite these strengths, AIMD also presents some practical limitations that should be considered when designing simulations. The most significant constraint lies in its high computational cost: since interatomic forces are calculated from first principles at every simulation time step, AIMD is far more resource-intensive than classical MD using empirical or interpolated potentials. Consequently, AIMD simulations are generally limited to system sizes of a few hundred atoms and timescales of only tens of picoseconds, making it difficult to capture slow processes such as molecular diffusion, long-timescale relaxation, or rare reaction events. This computational expense also restricts the number of trajectories that can be feasibly simulated, which in turn can reduce the statistical reliability of thermodynamic averages or reaction probabilities, particularly for low-probability events. In addition, the accuracy of AIMD results depends heavily on the choice of electronic structure method, which itself can introduce approximation, especially in treating dispersion forces or long-range interactions. Including van der Waals corrections can improve reliability but further increases computational cost. Overall, AIMD remains a powerful tool, but its application requires a thoughtful balance between accuracy, computational feasibility, and the specific demands of the system under investigation.

Given the structural and reactive complexity often encountered in gas-surface systems, such as surface reconstruction, bond formation and dissociation, charge transfer, and polarization effects AIMD offers a particularly suitable framework to investigate their dynamical behavior. By computing interatomic forces directly from the evolving electronic structure, AIMD inherently captures reactive events and preserves all nuclear degrees of freedom without relying on predefined reaction coordinates. This is especially advantageous in the presence of adsorbates and chemically active surfaces, where constructing a reliable high-dimensional potential energy surface can be prohibitively difficult. AIMD enables a rigorous, on-the-fly exploration of the system's evolution under thermal conditions, making it a powerful tool for studying gas–surface interactions. Nevertheless, due to its high computational cost and methodological complexity, careful attention must be given to the choice of electronic structure approximations, the treatment of dispersion and van der Waals forces, and the possible role of spin polarization, all of which are addressed in the following sections.

2.2 Quantum foundations and principles of AIMD

To understand gas–surface dynamics based on first principal calculations, *i.e.*, calculations that do not depend on experimental observations, we start with a general (non-relativistic) quantum description of a (molecule-surface) system that contains N electrons and M nuclei. We denote by $\mathbf{x} = \{\mathbf{x}_i = (\vec{r}_i, \sigma_i) \; \forall \; i = 1, 2, \ldots, N\}$ the set of electronic coordinates, which includes the $3N$ spatial coordinates $\vec{r}_i \in \mathbb{R}^3$ and the N spin coordinates $\sigma_i \in \{\uparrow, \downarrow\}$ of the electrons and the set of nuclear coordinates denoted $\vec{R} = \{\vec{R}_I \; \forall \; I = 1, 2, \ldots, M\}$ which includes the $3M$ spatial coordinates $\vec{R}_I \in \mathbb{R}^3$ of the nuclei. The quantum state k of this system is defined by its wave function $\Psi_k(\mathbf{x}, \vec{R})$, and its energy E_k obtained by solving the Schrödinger equation [107], which, in its time-independent and non-relativistic form reads:

$$\hat{H}\Psi_k(\vec{r}, \vec{R}) = E_k\Psi_k(\vec{r}, \vec{R}) \tag{2.1}$$

where $\hat{H}$ is the Hamiltonian operator of the system. In this equation and in all subsequent equations, we omit the nuclear spin and atomic units (a.u.) [108] will be used unless specified otherwise:

$$\hbar = 1; \frac{e^2}{4\pi\epsilon_0} = 1; \quad m_e = 1 \tag{2.2}$$

where $\hbar$ is the reduced Planck's constant ($\frac{h}{2\pi} \sim 1{,}05457 \times 10^{-34}\,\mathrm{J \cdot s}$), e is the elementary charge ($1.60211{\times}10^{-19}\,C$), ϵ_0 is the vacuum permittivity $8.85419{\times}10^{-12}\,m^{-3}kg^{-1}s^4A^2$) and m_e is the electron mass ($9.10953 \times 10^{-31}\,kg$).

In the equation 2.1, the Hamiltonian operator $\hat{H}$ act on the many-body wave-function Ψ that depends on the electron (r_i) and nuclear coordinates (R_A). Since the only fundamental interaction that is relevant to our system is the electrostatic interaction, the Hamiltonian is composed of kinetic energy terms (T) and electrostatic potentials (V):

$$\hat{H} = \hat{T}_n(\vec{R}) + \hat{T}_e(\vec{r}) + \hat{V}_{nn}(\vec{R}) + \hat{V}_{ee}(\vec{r}) + \hat{V}_{en}(\vec{r}, \vec{R}) \tag{2.3}$$

where $\hat{T}_n$ and $\hat{T}_e$ denote the nuclear and electronic kinetic energy operators, respectively, and the terms $\hat{V}_{nn}$, $\hat{V}_{ee}$, and $\hat{V}_{en}$ represent the nucleus-nucleus, electron-electron, and electron-nucleus Coulomb interactions. These components are explicitly given by:

$$\hat{T}_n = \sum_{I=1}^{M} -\frac{1}{2M_I}\nabla_I^2 \tag{2.4}$$

$$\hat{T}_e = \sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 \tag{2.5}$$

$$\hat{V}_{nn} = \sum_{I=1}^{M}\sum_{J>I}^{M} \frac{Z_I Z_J}{|\vec{R}_I - \vec{R}_J|} \tag{2.6}$$

$$\hat{V}_{ee} = \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{|\vec{r}_i - \vec{r}_j|} \tag{2.7}$$

$$\hat{V}_{en} = -\sum_{I=1}^{M}\sum_{i=1}^{N} \frac{Z_I}{|\vec{r}_i - \vec{R}_I|} \tag{2.8}$$

where M_I and Z_I are the mass and charge of nucleus I, and $\vec{R}_I$, $\vec{r}_i$ denote the positions of the nuclei and electrons, respectively. This Hamiltonian does not take into account the spin explicitly and it should be added as an additional degree of freedom, this is discussed in section 2.3.2.

Going back to the equation (2.1) to be solved, it involves $3N + 3M$ degrees of freedom, corresponding to the spatial coordinates of all electrons and nuclei. This high dimensionality renders its exact solution practically unfeasible for systems beyond the case of the hydrogen atom with only one electron. In fact, even storing the many-body wavefunction becomes intractable for most realistic gas–surface systems and as a result, any first-principles simulation must rely on systematic approximations. The most fundamental of these is the separation of electronic and nuclear motion, known as the Born–Oppenheimer approximation [109], which exploits the large mass disparity between nuclei and electrons ($M_I \geq 2\times 10^3 m_e$) to treat the nuclei as effectively stationary during electronic motion. In this approximation, the nuclear kinetic energy operator is neglected ($\hat{T}_n = 0$) and the nuclear–nuclear interaction term $\hat{V}_{nn}$ is constant for a given configuration of nuclei:

$$\hat{H} = \cancelto{\hat{0}}{\hat{T}_n(\vec{R})} + \hat{T}_e(\vec{r}) + \cancelto{cte}{\hat{V}_{nn}(\vec{R})} + \hat{V}_{ee}(\vec{r}) + \hat{V}_{en}(\vec{r},\vec{R}) \tag{2.9}$$

then

$$\hat{H}_e = \hat{T}_e(\vec{r}) + \hat{V}_{ee}(\vec{r}) + \hat{V}_{en}(\vec{r}, \vec{R}) \tag{2.10}$$

Within the Born–Oppenheimer framework, the electronic wavefunction $\Psi_e(\vec{r}; \vec{R})$ and the associated energy $E_e(\vec{R})$ are obtained by solving the time-independent electronic Schrödinger equation:

$$\hat{H}_e \Psi_e(\vec{r}, \vec{R}) = E_e(\vec{R}) \Psi_e(\vec{r}, \vec{R}) \tag{2.11}$$

Here, the nuclear coordinates $\vec{R}$ are treated as fixed parameters, and the resulting eigenvalue $E_e(\vec{R})$ defines the PES on which the nuclei evolve. This electronic energy, when combined with the electrostatic repulsion between nuclei, yields the total energy of the system for a given nuclear configuration:

$$E_{\text{tot}}(\vec{R}) = E_e(\vec{R}) + \sum_{I=1}^{M} \sum_{J<I}^{M} \frac{Z_I Z_J}{|\vec{R}_I - \vec{R}_J|} = E_e(\vec{R}) + \hat{V}_{nn}(\vec{R}) \tag{2.12}$$

This total energy forms the basis for the subsequent propagation of nuclear motion, typically assumed to occur on the electronic ground-state PES. Although conceptually straightforward, solving Eq. (2.11) for realistic gas–surface systems remains computationally challenging due to the many-electron nature of the problem. Consequently, a variety of approximate methods have been developed to treat the electronic problem efficiently. Among them, many rely on the variational principle, which states that the expectation value of the energy computed with any trial wavefunction will always be greater than or equal to the true ground-state energy. Consequently, by minimizing the energy with respect to the parameters of the trial wavefunction, one can obtain an approximation to the ground-state wavefunction and its corresponding energy. For this method, a foundational example is the Hartree–Fock (HF) approximation [110]. In this approach, the many-electron wavefunction is approximated by a single Slater determinant constructed from a set of orthonormal spin orbitals. The variational principle is then applied to determine the Slater determinant that minimizes the total electronic energy, leading to the Hartree–Fock equations, solved iteratively using a self-consistent field (SCF) procedure. The HF method captures the average interaction of each electron with the mean field generated by all other electrons, but it does not account for electron correlation effects beyond this mean-field description.

In order to overcome this limitation and incorporate electron correlation more accurately, Post-Hartree–Fock methods have been developed. These methods take the HF solution as a starting point and apply more sophisticated treatments, such as Configuration Interaction (CI), Møller–Plesset Perturbation Theory (MPn), and Coupled Cluster (CC) theory. Although these approaches typically offer improved accuracy compared to HF, they also entail a significantly higher computational cost. However, in systems where the ground-state wavefunction cannot be well described by a single determinant, such as in cases involving near-degenerate electronic states, multireference methods become necessary. I such cases, techniques such as Multireference Configuration Interaction (MRCI) rely on a reference wavefunction composed of multiple Slater determinants in order to account for both, static and dynamic correlation, but this calculations are computationally demanding and are generally limited to small molecular systems.

This limitation underscores the need of more computationally efficient approaches that still retain a quantum-mechanical description of the electronic structure, and in this context, DFT has emerged as the method of choice, or workhorse, for AIMD simulations, offering a practical compromise between accuracy and computational cost.

2.3 DFT as the workhorse of AIMD

In the framework of DFT, the focus shifts from the many-body wavefunction to the electron density $\rho(\vec{r})$, a quantity that offers both conceptual and computational advantages. Unlike the electronic wavefunction, which depends on all electronic coordinates and is not directly observable, the electron density is a measurable, physically meaningful function defined over three-dimensional space. Its represents the probability of finding any of the N electrons in an infinitesimal volume element $d\vec{r}$ around position $\vec{r}$, and is formally expressed as:

$$\rho(\vec{r_1}) = N \int \cdots \int |\Psi_e(\vec{x_1}, \vec{x_2}, \ldots, \vec{x_N})|^2 d\sigma_1 d\vec{x_2} \ldots d\vec{x_N} \tag{2.13}$$

where the integral is over the spin coordinates of all electrons and the spatial coordinates of all but one electron. The electron density is also a non-negative function of three spatial variables, x, y and z, which vanishes at infinity for molecular systems, but not for periodic ones, and integrates to the total number of electrons:

$$\lim_{|\vec{r}|\to+\infty} \rho(\vec{r}) = 0; \quad \rho(\vec{r}) \in \mathbb{R}_+^* \tag{2.14}$$

$$\int \rho(\vec{r})\, d\vec{r} = N \tag{2.15}$$

From a molecular perspective, nuclei appear as localized positive charges that coincide with local maxima in the electron density, these maxima, often associated with cusps in $\rho(\vec{r})$, carry essential information about the system. Kato's theorem [111, 112] provides a quantitative relation connecting the cusp behavior to the nuclear charge Z_I of each nucleus I, via the spherical average of the electron density $\tilde{\rho}(r_I)$ around that nucleus:

$$\left.\frac{\partial\tilde{\rho}(r_I)}{\partial r_I}\right|_{r_I=0} = -2Z_I\tilde{\rho}(r_I) \tag{2.16}$$

This result shows that nuclear properties are intrinsically encoded in the electronic density, more generally, as suggested by equations (2.15) and (2.16), the full information required to define the molecular Hamiltonian including the number of electrons, their distribution, the positions and charges of nuclei, appears to be embedded in $\rho(\vec{r})$. Therefore, it is reasonable to consider the electronic density as a fundamental variable from which the Hamiltonian, and subsequently the wavefunction and total energy, can in principle be determined. This insight establishes the conceptual foundation for DFT, which reformulates the electronic structure problem in terms of the electron density, rather than the full many-body wavefunction. However, the rigorous foundation of DFT lies in the two Hohenberg-Kohn (HK) theorems.

The first Hohenberg–Kohn theorem [113] establishes that the electron density $\rho_0(\vec{r})$ can be used as the fundamental variable instead of the number of electrons N and the external potential v_{ext} due to the nuclei. Since ρ determines the number of electrons, it follows that $\rho_0(\vec{r})$ also determines the ground-state wave function Ψ_0 and all other electronic properties of the system. Consequently, the ground-state energy is a unique functional of the ground-state electron density:

$$E_0(\rho_0) = T_e(\rho_0) + V_{ee}(\rho_0) + \int \rho_0(\vec{r})\, v_{\text{ext}}(\vec{r})\, d\vec{r} \tag{2.17}$$

In this equation, $T_e(\rho_0)$ and $V_{ee}(\rho_0)$ are the exact kinetic and electron–electron interaction energy functionals of the density, respectively. Their sum defines the Hohenberg–Kohn functional, $F_{\mathrm{HK}}(\rho) = T_e(\rho) + V_{ee}(\rho)$, which is universal, i.e., it is the same for all electronic systems and does not depend on the external potential.

The importance of the Hohenberg–Kohn energy functional becomes evident through the second HK theorem [113], which states that the ground-state electron density minimizes the total energy of the system. This result enables the application of the variational principle to determine both, the ground-state energy and the corresponding electron density. Crucially, this variational procedure involves only a three-dimensional function—the electron density—rather than the full high-dimensional many-body wavefunction.

It is important to emphasize that, in principle, no approximations have been introduced at this stage: the HK theorems provide the same exact ground state as that obtained from the full electronic Schrödinger equation. However, these theorems do not offer an explicit expression for the universal functional $F_{\mathrm{HK}}(\rho)$, which must therefore be approximated in practical calculations. To facilitate the construction of such approximations, it is useful to decompose the universal functional as:

$$F_{\mathrm{HK}}(\rho) = T_e(\rho) + V_{ee}(\rho) = T_e(\rho) + E_H(\rho) + E_{\mathrm{xc}}(\rho) \tag{2.18}$$

where $T_e(\rho)$ is the exact kinetic energy which is unknown, $E_H(\rho)$ is the classical electrostatic (Hartree) energy given by:

$$E_H(\rho) = \frac{1}{2} \int \frac{\rho(\vec{r})\rho(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|}\, d\vec{r}\, d\vec{r}\,' \tag{2.19}$$

and $E_{\mathrm{xc}}(\rho)$ is the exchange–correlation functional, which encapsulates all remaining quantum mechanical many-body effects, including exchange, correlation, and the correction to the kinetic energy beyond the non-interacting reference system.

In this context, to overcome the difficulty of expressing the exact kinetic energy as a functional of the electron density, Kohn and Sham [114] introduced a practical and widely adopted reformulation of DFT. Their approach is based on the construction of an auxiliary system of N non-interacting electrons that, by design, reproduces the same ground-state electron density $\rho_0(\vec{r})$ as the true, fully interacting system. This fictitious system does not include explicit electron–electron repulsion but instead introduces an effective one-electron potential $v_s(\vec{r})$ such that the electrons move independently. The key idea is that,

for this system, a large portion of the total energy—most notably the kinetic energy—can be computed exactly, leaving only a smaller residual to be treated approximately. This strategy enables DFT to capture much of the accuracy of wavefunction-based methods at significantly lower computational cost.

The Hamiltonian of the non-interacting Kohn–Sham (KS) reference system is separable into a sum of one-electron operators:

$$\hat{H}_s = \sum_{i=1}^{N} \left[-\frac{1}{2}\nabla_i^2 + v_s(\vec{r}_i) \right] \tag{2.20}$$

where $v_s(\mathbf{r})$ is the Kohn–Sham effective potential. The ground-state wavefunction of this system is represented by a single Slater determinant Φ_{KS}^0, constructed from a set of orthonormal single-particle orbitals $\varphi_i^{\mathrm{KS}}(\mathbf{r})$, called the Kohn–Sham orbitals. These orbitals are the lowest-energy solutions of the Kohn–Sham one-electron equation:

$$\left[-\frac{1}{2}\nabla^2 + v_s(\vec{r}) \right] \varphi_i^{\mathrm{KS}}(\vec{r}\,) = \varepsilon_i \varphi_i^{\mathrm{KS}}(\vec{r}\,) \tag{2.21}$$

where ε_i are the associated orbital energies. The electron density is then computed directly from these orbitals by summing over the occupied states:

$$\rho(\vec{r}) = \sum_{i=1}^{N} \left| \varphi_i^{\mathrm{KS}}(\vec{r}\,) \right|^2 \tag{2.22}$$

A crucial advantage of the Kohn–Sham approach is that the kinetic energy of the non-interacting reference system, denoted $T_s(\rho)$, can be evaluated exactly using the Kohn–Sham orbitals:

$$T_s(\rho) = -\frac{1}{2} \sum_{i=1}^{N} \int \varphi_i^*(\vec{r}) \nabla^2 \varphi_i(\vec{r})\, d\vec{r} \tag{2.23}$$

This term accounts for the majority of the true kinetic energy of the system. The remaining, unknown contributions, namely, the correction to the kinetic energy due to electron–electron interactions and the non-classical parts of the electron–electron repulsion, are collected into the exchange–correlation functional:

$$E_{\mathrm{xc}}(\rho) = T_c(\rho) + E_{\mathrm{xc}}^{\mathrm{G}}(\rho) \tag{2.24}$$

where $T_c(\rho) = T_e(\rho) - T_s(\rho)$ is the kinetic correlation energy, and $E_{\text{xc}}^{\text{G}}(\rho)$ includes exchange and dynamic correlation effects. Together, they represent everything that remains unknown in the total energy expression.

However, a crucial question remains: how can we define the effective potential $v_s(\mathbf{r})$ such that the resulting Slater determinant yields exactly the same electron density as the real, interacting system? To address this challenge, we express the total energy of the interacting system using the decomposition introduced earlier:

$$E[\rho(\vec{r})] = T_s[\rho(\vec{r})] + E_H[\rho(\vec{r})] + E_{\text{xc}}[\rho(\vec{r})] + V_{\text{ext}}[\rho(\vec{r})] \tag{2.25}$$

which can be written explicitly as:

$$E[\rho(\vec{r})] = T_s[\rho(\vec{r})] + \frac{1}{2}\int\int \frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|}\, d\vec{r}_1\, d\vec{r}_2 + E_{\text{xc}}[\rho(\vec{r})] + \int v_{\text{ext}}(\vec{r})\rho(\vec{r})\, d\vec{r} \tag{2.26}$$

The only term in this expression that lacks an explicit form is the exchange–correlation energy $E_{\text{xc}}(\rho)$, which encapsulates the many-body effects not captured by the non-interacting kinetic energy or the classical Coulomb repulsion. Applying the variational principle to determine the set of orbitals $\{\varphi_i\}$ that minimize the total energy functional in Eq. (2.26), subject to the orthonormality constraint $\langle\varphi_i|\varphi_j\rangle = \delta_{ij}$, yields the Kohn–Sham equations:

$$\left[-\frac{1}{2}\nabla^2 + \int \frac{\rho(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|}\, d\vec{r}_2 + v_{\text{xc}}(\vec{r}_1) + v_{\text{ext}}(\vec{r}_1)\right]\varphi_i(\vec{r}_1) = \varepsilon_i\varphi_i(\vec{r}_1) \tag{2.27}$$

The quantity in brackets is the *Kohn–Sham effective potential* $v_s(\vec{r})$, defined as:

$$v_s(\vec{r}) \equiv v_{\text{eff}}(\vec{r}) = \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}\,'|}\, d\vec{r}\,' + v_{\text{xc}}(\vec{r}) + v_{\text{ext}}(\vec{r}) \tag{2.28}$$

Here, $v_{\text{xc}}(\vec{r})$ is the exchange–correlation potential, which arises from the exchange–correlation energy functional $E_{\text{xc}}(\rho)$. Since $E_{\text{xc}}(\rho)$ is not known exactly, its functional derivative, and thus $v_{\text{xc}}(\mathbf{r})$, must also be approximated:

$$v_{\text{xc}}(\vec{r}) = \frac{\delta E_{\text{xc}}(\rho)}{\delta\rho(\vec{r})} \tag{2.29}$$

If the exact functional form of the exchange–correlation energy $E_{\mathrm{xc}}(\rho)$, and consequently its functional derivative $v_{\mathrm{xc}}(\vec{r})$, were known, the Kohn–Sham formalism would yield the exact ground-state energy of the system. In such a case, the total energy obtained from solving the Kohn–Sham equations would coincide with the true eigenvalue of the electronic Hamiltonian defined in the full Schrödinger equation. This is why the Kohn–Sham approach is often regarded as exact in principle, but in practice, the need for approximation arises precisely because the exact form of $E_{\mathrm{xc}}(\rho)$ remains unknown. This challenge is particularly critical in AIMD studies of gas–surface systems, where electronic density inhomogeneities, long-range interactions, and localized states at the interface place stringent demands on the accuracy of the exchange–correlation functional. Therefore, a critical step in this simulations is the correct selection of the functional to describe the exchange–correlation energy, $E_{xc}(\rho)$, which captures the complex quantum mechanical effects arising from electron exchange and correlation.

2.3.0.1 Choice of the exchange-correlation functional

Over the past decades, numerous exchange–correlation functionals have been proposed, differing in their theoretical foundations, degree of empirical parameterization, and target properties. Some are optimized for broad applicability, while others are specifically designed for surface chemistry or catalytic systems. To organize this landscape, Perdew's "Jacob's ladder" [115] provides a systematic framework that ranks functionals by the increasing sophistication of the ingredients used to approximate $E_{xc}(\rho)$. As illustrated in Figure 2.1, each rung of the ladder incorporates additional physical information, ranging from the local electron density to orbital-dependent terms.

At the lowest rung of this hierarchy lies the Local Density Approximation (LDA), and its spin-polarized counterpart, the Local Spin Density Approximation (LSDA). These approaches are based on the concept of a uniform electron gas, in which electrons move in a homogeneous background of positive charge, leading to a constant electron density. In this model system, the exchange–correlation energy per particle is known exactly, making it a valuable reference in the development of DFT approximations.

Within the LDA framework, it is assumed that the exchange–correlation energy density at a given point in space depends solely on the local electron density $\rho(\vec{r})$. The total exchange–correlation energy is then expressed as:

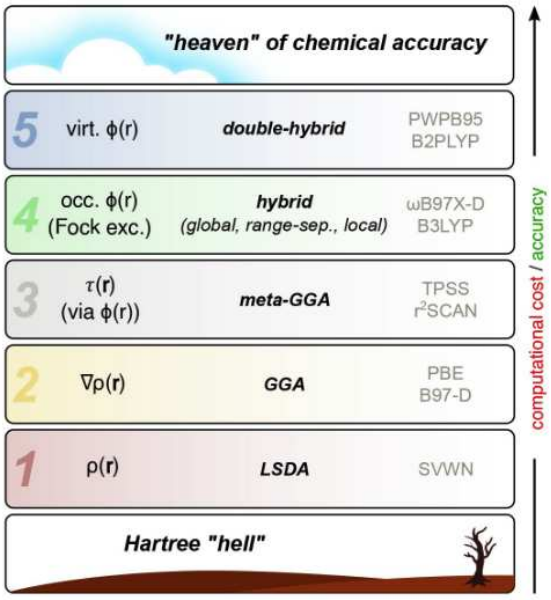

Figure 2.1: Functional categorization according to Perdew's "Jacob's ladder". ρ = electron density, τ = kinetic energy density, ϕ = molecular orbital, Fock exc. = Fock exchange. (Source Bursch et al. [116])

$$E_{\mathrm{xc}}^{\mathrm{LDA}}(\rho) = \int \rho(\vec{r})\, \varepsilon_{\mathrm{xc}}[\rho(\vec{r})]\, d\vec{r} \tag{2.30}$$

where $\varepsilon_{\mathrm{xc}}(\rho)$ is the exchange–correlation energy per particle of a uniform electron gas at density ρ. This quantity is generally separated into exchange and correlation components:

$$\varepsilon_{\mathrm{xc}}(\rho) = \varepsilon_x(\rho) + \varepsilon_c(\rho) \tag{2.31}$$

The exchange part, $\varepsilon_x(\rho)$, is known analytically from Dirac's formula for a uniform electron gas:

$$\varepsilon_x(\rho) = -C_x\, \rho^{1/3}, \quad \text{with} \quad C_x = \frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \approx 0.739 \tag{2.32}$$

The correlation part, $\varepsilon_c(\rho)$, does not have a closed-form expression but can be obtained numerically from quantum Monte Carlo calculations performed by Ceperley and Alder [117], and subsequently interpolated into analytic forms for practical use [118].

Despite its simplicity, LDA (and LSDA) often yields surprisingly accurate results in solid-state systems, particularly for bulk materials and metallic surfaces. However, its performance significantly deteriorates in systems exhibiting strong inhomogeneity in the electronic density, such as molecular complexes, weakly bound systems, and gas–surface interactions involving dispersion forces.

The second rung of Jacob's ladder introduces the Generalized Gradient Approximation (GGA) [119, 120], which improves upon LDA by incorporating not only the local electron density $\rho(\vec{r})$, but also its spatial gradient $\nabla\rho(\vec{r})$. This allows GGA functionals to account more accurately for inhomogeneities in the electronic distribution, especially in regions where the density varies rapidly, such as in molecular bonding or near surfaces. In general, the GGA exchange–correlation functional is written as:

$$E_{\text{xc}}^{\text{GGA}}(\rho_\alpha, \rho_\beta) = \int f(\rho_\alpha, \rho_\beta, \nabla\rho_\alpha, \nabla\rho_\beta)\, d\vec{r} \tag{2.33}$$

where $f(\rho_\alpha, \rho_\beta, \nabla\rho_\alpha, \nabla\rho_\beta)$ is a semi-local function designed to satisfy known physical constraints. This expression is usually split into exchange and correlation components:

$$E_{\text{xc}}^{\text{GGA}} = E_x^{\text{GGA}} + E_c^{\text{GGA}} \tag{2.34}$$

For the exchange part (E_x^{GGA}), a widely used formulation is based on the enhancement of the Dirac exchange energy through a multiplicative factor $F(s)$, which depends on the so-called reduced density gradient:

$$s = \frac{|\nabla\rho(\vec{r})|}{\rho(\vec{r})^{4/3}} \tag{2.35}$$

leading to the general form:

$$E_x^{\text{GGA}}[\rho] = -C_x \int \rho^{4/3}(\vec{r}) F(s)\, d\vec{r} \tag{2.36}$$

where $C_x = \frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \approx 0.74$, and $F(s)$ modifies the local exchange contribution depending on the degree of density inhomogeneity. Different choices of $F(s)$ define different GGA functionals like the FT97 [121], the PW91 [122, 123], the CAM(A) and CAM(B) [124], B86 [125], the P86 [126], the LG [127] and the PBE [128]. For the latter functional (PBE), the enhancement factor is given by:

$$F^{\mathrm{PBE}}(s) = 1 + \kappa - \frac{\kappa}{1 + \frac{\mu s^2}{\kappa}} \tag{2.37}$$

with empirically optimized parameters $\mu = 0.22$ and $\kappa = 0.80$.

To complement these exchange formulations, several gradient-corrected correlation functionals (E_c^{GGA}) have also been developed, although their analytical forms are often more intricate. Notable examples include the correlation component of P86 [126], which incorporates empirical fitting to the correlation energy of the neon atom, and its improved parameter-free successor PW91 [129]. Another widely adopted functional is LYP [130], derived from the correlation energy of the helium atom and independent of the uniform electron gas model. In practice, specific exchange and correlation components are typically combined to form functional pairs such as BP86, BLYP, and BPW91, each offering distinct trade-offs in accuracy and transferability. These combinations are widely used in simulations involving gas–surface interactions and materials chemistry.

The third rung of Jacob's ladder is occupied by the meta-generalized gradient approximation (meta-GGA) [131], which extends the GGA framework by incorporating additional semi-local ingredients. In addition to the local electron density $\rho(\vec{r})$ and its gradient $\nabla\rho(\vec{r})$, meta-GGA functionals depend on the kinetic energy density $\tau(\vec{r})$, typically defined as:

$$\tau(\vec{r}) = \frac{1}{2}\sum_i |\nabla\varphi_i^{\mathrm{KS}}(\vec{r})|^2 \tag{2.38}$$

where φ_i^{KS} are the occupied Kohn–Sham orbitals. In some formulations, the Laplacian of the density $\nabla^2\rho(\vec{r})$ is also included like in the LAP correlation functional [132], although the general form of the exchange–correlation energy without consider such terms is expressed as:

$$E_{\mathrm{xc}}^{\mathrm{MGGA}}[\rho, \nabla\rho, \tau] = \int d\vec{r}\, \varepsilon_{\mathrm{xc}}^{\mathrm{MGGA}}(\rho, \nabla\rho, \tau) \tag{2.39}$$

The inclusion of $\tau(\vec{r})$ enables meta-GGA to capture more detailed information about electron localization, making it possible to distinguish between covalent, metallic, and weakly bound systems. Moreover, these functionals can be constructed to satisfy additional exact constraints, such as the self-correlation-free condition for one-electron

systems $E_{\mathrm{xc}}[n, 0] = 0$. The meta-GGA maintain a semi-local character and do not require computationally demanding non-local integrals. As such, they offer improved accuracy over GGA functionals without the high computational cost of hybrid or fully non-local approaches.

The fourth rung of Jacob's ladder corresponds to hybrid functionals, which combine a fraction of exact exchange from Hartree–Fock theory with a semi-local exchange–correlation functional, typically of GGA type. This hybridization aims to reduce self-interaction errors inherent in standard DFT approximations and improves the accuracy of computed reaction barriers, atomization energies, and structural properties. The exact exchange energy, as derived from Hartree–Fock theory, is given by:

$$E_x^{\mathrm{exact}} = -\frac{1}{2}\sum_{i,j}\int \frac{\varphi_i^*(\vec{r})\varphi_j^*(\vec{r}\,')\varphi_i(\vec{r}\,')\varphi_j(\vec{r})}{|\vec{r}-\vec{r}\,'|}\, d\vec{r}\, d\vec{r}\,' \tag{2.40}$$

A hybrid functional introduces this term into the exchange–correlation energy expression by mixing it with a GGA exchange component. The general form of a hybrid exchange–correlation functional is:

$$E_{\mathrm{xc}}^{\mathrm{hybrid}} = aE_x^{\mathrm{exact}} + (1-a)E_x^{\mathrm{GGA}} + E_c^{\mathrm{GGA}} \tag{2.41}$$

where the parameter $a \in [0, 1]$ defines the fraction of exact exchange included in the hybrid. Although this mixing parameter can be theoretically motivated, it is often empirically tuned to improve agreement with experimental reference data.

One of the most widely used hybrid functionals is B3LYP [133], which has been the workhorse of quantum chemistry due to its consistent accuracy in predicting molecular geometries and thermochemical properties. Another example is PBE0 [134], which incorporates exact exchange into the PBE functional. More recently, exact exchange has also been incorporated into meta-GGA functionals to create so-called hybrid meta-GGA, further improving performance in thermochemistry and kinetics.

Despite their success in molecular systems, the use of hybrid functionals in solid-state and surface simulations is less common. The computational cost associated with evaluating the exact exchange term in Eq. (2.40) becomes significant for extended systems, particularly metals, where delocalized electrons and the long-range nature of the Coulomb interaction lead to slow convergence. To mitigate this, screened hybrid

functionals have been developed, such as HSE [135], where the long-range part of the exact exchange is replaced by its GGA analogue, and only the short-range part is treated explicitly. Screened hybrids like HSE offer a promising balance between accuracy and computational efficiency and have shown good performance in both solid-state systems and small molecules. However, the computational overhead remains a challenge in dynamic simulations of gas–surface interactions, where large supercells and dense k-point meshes are often required.

Finally, the fifth and highest rung of Jacob's ladder is occupied by double-hybrid density functionals (DHDFs) [136], which combine the strengths of both density-based and wavefunction-based approaches. These functionals incorporate a fraction of exact exchange from Hartree–Fock theory, as in hybrid functionals, together with a non-local, perturbative correlation correction derived from second-order wavefunction theory. The general form of the exchange–correlation energy in a double-hybrid functional is expressed as:

$$E_{\mathrm{xc}}^{\mathrm{DHDF}} = a_{\mathrm{x}} E_{\mathrm{x}}^{\mathrm{HF}} + (1 - a_{\mathrm{x}}) E_{\mathrm{x}}^{\mathrm{GGA}} + (1 - a_{\mathrm{c}}) E_{\mathrm{c}}^{\mathrm{GGA}} + a_{\mathrm{c}} E_{\mathrm{c}}^{\mathrm{PT2}} \tag{2.42}$$

where $E_{\mathrm{x}}^{\mathrm{HF}}$ is the exact Hartree–Fock exchange energy, and $E_{\mathrm{x}}^{\mathrm{GGA}}$ and $E_{\mathrm{c}}^{\mathrm{GGA}}$ are semi-local GGA exchange and correlation components, respectively. The terms a_{x} and a_{c} are empirical mixing parameters that control the relative contribution of exact exchange and perturbative correlation. This perturbative correlation energy is denoted $E_{\mathrm{c}}^{\mathrm{PT2}}$ and is often computed using second-order Møller–Plesset theory (MP2) [137], which is evaluated post-self-consistently using Kohn–Sham orbitals and eigenvalues. In the spin-restricted formalism, it is given by:

$$E_{\mathrm{c}}^{\mathrm{MP2}} = -\frac{1}{4} \sum_{i,j}^{\mathrm{occ}} \sum_{a,b}^{\mathrm{virt}} \frac{|\langle ij||ab\rangle|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} \tag{2.43}$$

where i, j run over the occupied orbitals, a, b over virtual (unoccupied) orbitals, and ε are the corresponding orbital energies. The antisymmetrized two-electron integrals $\langle ij||ab\rangle$ are defined as:

$$\langle ij||ab\rangle = \langle ij|ab\rangle - \langle ij|ba\rangle \tag{2.44}$$

with the conventional Coulomb integral given by:

$$\langle ij|ab\rangle = \int\int \varphi_i^*(\vec{r}_1)\varphi_j^*(\vec{r}_2)\frac{1}{|\vec{r}_1-\vec{r}_2|}\varphi_a(\vec{r}_1)\varphi_b(\vec{r}_2)\,d\vec{r}_1\,d\vec{r}_2 \tag{2.45}$$

The addition of a perturbative correlation term allows DHDFs to capture non-local correlation effects that are otherwise inaccessible to semi-local density functionals. As a result, they achieve high accuracy for thermochemical properties, reaction barrier heights, and non-covalent interactions, rivaling more computationally expensive wavefunction methods. A well-known example of DHDFs is the B2PLYP functional [138], which uses $a_x = 0.53$ and $a_c = 0.27$, combining Becke exchange (B88) [139], LYP correlation [130], HF exchange, and an MP2-like correlation correction.

However, while MP2 provides a rigorous and physically motivated correction to the Kohn–Sham correlation energy, its practical implementation is computationally demanding. The need to evaluate two-electron integrals over virtual orbitals leads to steep scaling with system size, making DHDFs challenging to apply in large-scale AIMD simulations or periodic surface models, particularly when using plane-wave basis sets. Nevertheless, double-hybrid functionals represent a significant advance in the pursuit of chemical accuracy within the DFT framework.

2.3.1 Dispersion Corrections in DFT: The DFT-D Method

As previously showed, DFT successfully describe a wide range of molecular and condensed-phase systems with its well established conventional exchange–correlation functionals, ranging from LDA and GGA to the hybrid functionals. However, they fail to account for long-range dispersion (van der Waals) interactions [140]. These interactions arise from dynamic electron correlation effects between distant regions of electron density, and are crucial for accurately modeling non-covalent interactions, particularly in systems involving molecular adsorption, layered materials, or weakly bound clusters. The underlying reason for this failure lies in the construction of most standard functionals, which are based solely on local or semi-local properties of the electron density and do not explicitly include virtual orbitals or nonlocal correlation terms. To address this deficiency, a variety of dispersion-corrected DFT methods have been developed such as vdW-DF and parameterized functionals that incorporate nonlocal correlation terms directly into the energy or potential; C_6-based corrections like the DFT-D family, which add an explicit pairwise dispersion energy term to the total energy; and effective one-electron potential

schemes such as DCACP and LAP/DCP, which introduce dispersion effects through modifications to the KS potential as illustrated in Figure 2.2.

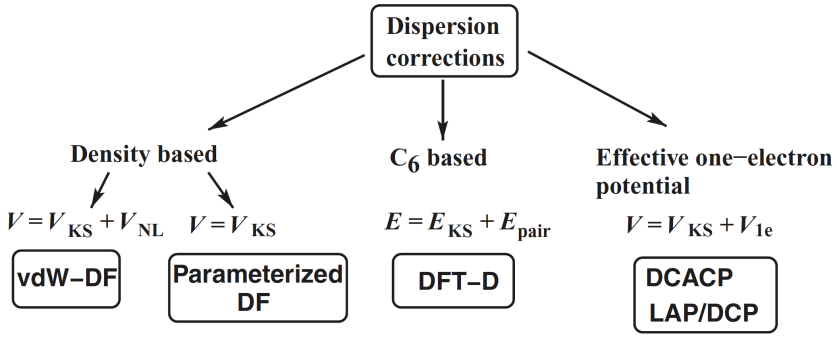

Figure 2.2: Overview of commonly used dispersion correction methods in DFT where E_{KS} and V_{KS} correspond to the bare Kohn–Sham total energies and potentials, respectively. These scheme include density-based approaches (vdW-DF, parameterized functionals), C_6-based pairwise correction schemes (DFT-D), and effective one-electron potential modifications (DCACP, LAP/DCP). (Source Grimme [140])

Among these, the semiempirical DFT-D methods introduced and systematically developed by Grimme [141, 142] have emerged as some of the most widely adopted due to their simplicity, accuracy, and ease of implementation across a broad range of systems. In the DFT-D framework, the total energy is augmented by an explicit pairwise dispersion correction term:

$$E_{\text{total}} = E_{\text{KS}} + E_{\text{disp}}^{\text{DFT-D}} \tag{2.46}$$

where E_{KS} is the standard Kohn–Sham DFT energy and $E_{\text{disp}}^{\text{DFT-D}}$ accounts for dispersion. The correction term is computed as:

$$E_{\text{disp}}^{\text{DFT-D}} = -\sum_{A<B} \sum_{n=6,8,10,\ldots} s_n \frac{C_{AB}^{(n)}}{R_{AB}^n} f_{\text{damp}}(R_{AB}) \tag{2.47}$$

where the first summation runs over all unique atomic pairs A, B in the system, and the second over the dispersion orders $n = 6, 8, 10, \ldots$, with $n = 6$ being the dominant term. The coefficients $C_{AB}^{(n)}$ are dispersion constants specific to each atom pair, typically derived from atomic polarizabilities and ionization potentials or fitted to reference data. The quantity R_{AB} is the interatomic distance between atoms A and B, s_n is a global scaling

factor depending on the exchange–correlation functional, and $f_{\text{damp}}(R_{AB})$ is a damping function that limits the correction at short interatomic distances to avoid double-counting correlation already captured by the DFT functional.

Among the various generations of the DFT-D method (D1 [141], D2 [143], D3 [142], D4 [144]), the DFT-D3(BJ) variant, which employs a rational Becke–Johnson damping function [145], is particularly notable for its robust performance and transferability across molecular and periodic systems. It uses atom-specific coordination numbers to adjust dispersion coefficients in a geometry-dependent manner, thus introducing a degree of system sensitivity lacking in earlier models.

In the context of AIMD, the inclusion of dispersion corrections modifies the nuclear forces, which are derived from the total energy as:

$$\vec{F}_I = -\nabla_I \left(E_{\text{KS}} + E_{\text{disp}}^{\text{DFT-D}} \right) \tag{2.48}$$

This correction allows AIMD simulations to account for long-range correlation effects essential to accurately modeling gas–surface interactions, molecular interfaces, and non-covalent interactions in extended systems.

2.3.2 Treatment of spin in the Kohn–Sham formalism

In the previous discussion, the spin wasn't take into consideration in the equations, however, in gas–surface interaction studies, it is common to encounter open-shell species, such as radicals or transition-metal-containing adsorbates, for which the consideration of electronic spin is essential to accurately describe the electronic structure, adsorption energetics, and reactivity of the system. Within the KS formulation of DFT, spin effects can be explicitly incorporated, giving rise to *spin-dependent DFT*, also referred to as *spin-polarized* or *spin-unrestricted* DFT.

In standard *spin-restricted* DFT, it is assumed that electrons with spin α (up) and spin β (down) occupy the same spatial orbitals. However, for systems where the spin densities differ, it becomes necessary to introduce not only the total electron density $\rho(\vec{r})$, but also the individual spin densities $\rho_\alpha(\vec{r})$ and $\rho_\beta(\vec{r})$, such that:

$$\rho(\vec{r}) = \rho_\alpha(\vec{r}) + \rho_\beta(\vec{r}) \tag{2.49}$$

Analogously to the spin-independent case, a fictitious system of non-interacting electrons is introduced, reproducing the same spin densities as the interacting system. This gives rise to *spin-dependent Kohn-Sham orbitals*, $\phi_i^{\alpha,KS}(\vec{r})$ and $\phi_i^{\beta,KS}(\vec{r})$, from which the spin densities are constructed as:

$$\rho_\alpha(\vec{r}) = \sum_i \left|\varphi_i^{\alpha,KS}(\vec{r})\right|^2 \tag{2.50}$$

$$\rho_\beta(\vec{r}) = \sum_i \left|\varphi_i^{\beta,KS}(\vec{r})\right|^2 \tag{2.51}$$

where the sums run over the occupied orbitals of each spin type. The Kohn-Sham energy functional becomes dependent on both spin densities, written as $E_{KS}(\rho_\alpha, \rho_\beta)$. Application of the variational principle yields the spin-dependent Kohn-Sham equations:

$$\left(-\frac{1}{2}\nabla^2 + V_{KS}^{\alpha}(\vec{r})\right)\varphi_i^{\alpha,KS}(\vec{r}) = \epsilon_i^{\alpha}\varphi_i^{\alpha,KS}(\vec{r}) \tag{2.52}$$

$$\left(-\frac{1}{2}\nabla^2 + V_{KS}^{\beta}(\vec{r})\right)\varphi_i^{\beta,KS}(\vec{r}) = \epsilon_i^{\beta}\varphi_i^{\beta,KS}(\vec{r}) \tag{2.53}$$

where ϵ_i^{α} and ϵ_i^{β} are the Kohn-Sham orbital energies for spin-up and spin-down electrons, respectively. The spin-dependent Kohn-Sham potentials $V_{KS}^{\alpha}(\vec{r})$ and $V_{KS}^{\beta}(\vec{r})$ are composed of three contributions: the external potential from the nuclei $V_{ext}(\vec{r})$, the classical Hartree potential $V_H(\vec{r})$, and the spin-dependent exchange-correlation potential $V_{xc}^{\alpha}(\vec{r})$, $V_{xc}^{\beta}(\vec{r})$:

$$V_{KS}^{\alpha}(\vec{r}) = V_{ext}(\vec{r}) + V_H(\vec{r}) + V_{xc}^{\alpha}(\vec{r}) \tag{2.54}$$

$$V_{KS}^{\beta}(\vec{r}) = V_{ext}(\vec{r}) + V_H(\vec{r}) + V_{xc}^{\beta}(\vec{r}) \tag{2.55}$$

The Hartree potential is determined from the total electronic density:

$$V_H(\vec{r}) = \int \frac{\rho(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|}\, d\vec{r}\,' \tag{2.56}$$

The exchange-correlation potentials are obtained as functional derivatives of the exchange-correlation functional $E_{xc}(\rho_\alpha, \rho_\beta)$ with respect to the spin densities:

$$V_{xc}^{\alpha}(\vec{r}) = \frac{\delta E_{xc}(\rho_\alpha, \rho_\beta)}{\delta \rho_\alpha(\vec{r})} \tag{2.57}$$

$$V_{xc}^{\beta}(\vec{r}) = \frac{\delta E_{xc}(\rho_\alpha, \rho_\beta)}{\delta \rho_\beta(\vec{r})} \tag{2.58}$$

As in the *spin-restricted* case, these equations must be solved self-consistently, since the Kohn-Sham potentials depend on the spin densities, which in turn are determined by the orbitals being solved for. In practical applications, the exchange-correlation functional $E_{xc}(\rho_\alpha, \rho_\beta)$ is usually approximated using either LSDA or GGA, which are generalized to depend explicitly on both spin densities and their gradients. Spin-unrestricted versions of widely used functionals such as PBE are routinely employed for systems where spin effects are relevant.

2.3.3 DFT for Periodic Systems: Plane-Wave Expansion and Bloch Formalism

DFT provides a powerful framework for studying the electronic structure of many-electron systems, including crystalline solids. In periodic systems, the infinite repetition of a fundamental structural motif—the unit cell—imposes translational symmetry that can be leveraged to simplify the quantum mechanical problem. Rather than explicitly treating an infinite array of atoms, one can focus on a single unit cell and impose *periodic boundary conditions* to recover the essential physics of the extended solid.

The mathematical exploitation of this symmetry leads naturally to the concept of the *reciprocal lattice*, which underpins much of the formalism in solid-state DFT. The reciprocal lattice vectors $\vec{b}_1, \vec{b}_2, \vec{b}_3$ are defined in terms of the direct lattice vectors $\vec{a}_1, \vec{a}_2, \vec{a}_3$ and the volume of the unit cell Ω as:

$$\vec{b}_1 = 2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\Omega} \tag{2.59}$$

$$\vec{b}_2 = 2\pi \frac{\vec{a}_3 \times \vec{a}_1}{\Omega} \tag{2.60}$$

$$\vec{b}_3 = 2\pi \frac{\vec{a}_1 \times \vec{a}_2}{\Omega} \tag{2.61}$$

These vectors define a momentum-space lattice that is reciprocal to the real-space crystal and are essential in the construction of the basis sets and Brillouin zone sampling schemes used in periodic DFT calculations.

Building upon this framework, Bloch's theorem provides the foundation for describing the quantum states of electrons in a periodic potential. According to this theorem, the wavefunction of an electron in a periodic potential can be written as:

$$\varphi_{j\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{j\vec{k}}(\vec{r}) \tag{2.62}$$

where $\vec{k}$ is a wavevector located within the first Brillouin zone, and $u_{j\vec{k}}(\vec{r})$ is a function with the same periodicity as the underlying lattice, satisfying $u_{j\vec{k}}(\vec{r} \pm n\vec{T}) = u_{j\vec{k}}(\vec{r})$ for any lattice vector $\vec{T}$. This decomposition allows the complex behavior of electrons in a solid to be described in terms of band structures and crystal momentum.

To represent the periodic part of the Bloch functions $u_{j\vec{k}}(\vec{r})$, plane waves offer a natural and efficient basis set. Their delocalized nature and compatibility with periodic boundary conditions make them ideal for use in solids. In this representation, the Kohn–Sham orbitals are expanded as:

$$\varphi^{KS}_{j\vec{k}}(\vec{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\vec{G}} C_{j\vec{k}}(\vec{G}) e^{i(\vec{k}+\vec{G})\cdot\vec{r}} \tag{2.63}$$

where $\vec{G}$ are reciprocal lattice vectors, and $C_{j\vec{k}}(\vec{G})$ are the expansion coefficients to be determined. The use of plane waves brings computational advantages: derivatives become simple multiplicative operations in reciprocal space, and Pulay forces vanish due to the fixed basis, simplifying the evaluation of forces during geometry optimization or molecular dynamics.

However, since the number of plane waves in the expansion must be finite in practice, a kinetic energy cutoff E_{cut} is introduced to limit the basis to those plane waves satisfying:

$$\frac{1}{2}|\vec{k}+\vec{G}|^2 \leq E_{\text{cut}} \tag{2.64}$$

This cutoff defines the spatial resolution of the wavefunction and is a crucial parameter governing both the accuracy and computational cost of the simulation. The number of plane waves N_{PW} required for a given E_{cut} grows approximately as:

$$N_{PW} \propto \Omega \cdot E_{\text{cut}}^{3/2} \tag{2.65}$$

illustrating the importance of balancing precision and efficiency, particularly for large unit cells. The use of *pseudopotentials* becomes essential in this context, as they replace the all-electron potential near atomic nuclei with a smoother effective potential. This simplification eliminates the need to explicitly treat tightly bound core electrons, allowing the potential in the core region to be smoothed and significantly reducing the number of plane waves required to accurately represent the valence states.

The Kohn–Sham equations, when expressed in the plane-wave basis, transform into a matrix eigenvalue problem for each $\vec{k}$-point in the Brillouin zone:

$$\sum_{\vec{G}'} H_{\vec{k}}(\vec{G},\vec{G}')C_{j\vec{k}}(\vec{G}') = \varepsilon_{j\vec{k}}C_{j\vec{k}}(\vec{G}) \tag{2.66}$$

where $H_{\vec{k}}(\vec{G},\vec{G}')$ is the Hamiltonian matrix, composed of kinetic energy terms $|\vec{k}+\vec{G}|^2$ and the effective potential—including the ionic (pseudo)potential, the Hartree potential, and the exchange-correlation potential, all of which depend on the electron density.

Solving this eigenvalue problem for a sufficiently dense set of $\vec{k}$-points yields access to the electronic structure of the periodic system, including its energy bands, total energy, and derived properties such as charge densities, densities of states, and forces. Thus, by exploiting translational symmetry and employing a plane-wave basis, periodic DFT allows for a tractable and systematically convergent description of crystalline materials.

2.4 Molecular dynamics on a quantum PES

In the Born–Oppenheimer approximation the nuclei are treated as classical particles moving on the PES generated by the electrons, which are treated quantum mechanically using DFT. Thus, the nuclear motion on this potential can be described classically using Newton's second law [104, 146–148]:

$$\mathbf{F}_I = m_I \ddot{\mathbf{R}}_I(t) \tag{2.67}$$

At each time step, for a given nuclear configuration $\{R(t)\}$, the electronic ground state is calculated using DFT by solving the KS equations self-consistently. Once the converged electron density is obtained, the forces acting on the nuclei F_A are calculated using the Hellman-Feynman theorem [149, 150]:

$$F_I = -\nabla_I \langle \Psi^0_{KS} | \hat{H}_e + \hat{V}_{nn} | \Psi^0_{KS} \rangle = -\nabla_I [E^0_{KS}[\rho_0] + V_{nn}(\{R(t)\})] \tag{2.68}$$

where $\hat{H}_e$ is the electronic Hamiltonian, $\hat{V}_{nn}$ is the internuclear repulsion potential, and $E^0_{KS}[\rho_0]$ is the ground state electronic energy obtained from DFT.

Once the forces are calculated, the equations of motion must be numerically integrated to determine the positions $\mathbf{R}(t)$ and velocities $\mathbf{v}(t)$ of the nuclei. Numerical integration involves discretizing time into small intervals Δt that are sufficiently small to ensure energy conservation throughout the simulation, enabling the calculation of positions and velocities at successive timesteps. This discretization transforms the continuous differential equations (2.67) into discrete forms, for which various algorithms have been developed [151].

A widely used, direct approach is to expand the nuclear coordinates $\mathbf{R}(t \pm \Delta_t)$ in a Taylor series, leading to the Verlet integration scheme [152]:

$$\begin{cases} \mathbf{R}(t+\Delta t) = \mathbf{R}(t) + \mathbf{v}(t)\Delta t + \frac{\mathbf{F}_I}{2m_I}\Delta t^2 + \frac{\Delta t^3}{3!}\dddot{\mathbf{R}}(t) + o(\Delta t^4) \\ \\ \mathbf{R}(t-\Delta t) = \mathbf{R}(t) - \mathbf{v}(t)\Delta t + \frac{\mathbf{F}_I}{2m_I}\Delta t^2 - \frac{\Delta t^3}{3!}\dddot{\mathbf{R}}(t) + o(\Delta t^4) \end{cases} \tag{2.69}$$

Adding these two equations (2.69), we obtain an expression for the position $\mathbf{R}(t+\Delta t)$ in terms of positions at the times t and $t-\Delta t$, and the forces $\mathbf{F}_I$:

$$\mathbf{R}(t+\Delta t) \approx 2\mathbf{R}(t) - \mathbf{R}(t-\Delta t) + \frac{\mathbf{F}_I}{m_I}\Delta t^2 + O(\Delta_t^4) \tag{2.70}$$

where m_I is the mass of atom I. This process is repeated at each time step, allowing for the simulation of the dynamical evolution of the system.

It is important to note, however, that most modern molecular dynamics codes, including VASP, employ the so-called velocity Verlet algorithm, a modification of the original Verlet scheme. The velocity Verlet formulation provides direct access to both positions and velocities at each time step, which is advantageous for calculating kinetic properties and for use in thermostats and barostats. The velocity Verlet scheme updates positions and velocities as follows:

$$\mathbf{R}(t+\Delta_t) = \mathbf{R}(t) + \Delta_t\,\mathbf{V}(t) + \frac{\Delta_t^2}{2m_I}\mathbf{F_I}(t) \tag{2.71}$$

$$\mathbf{V}(t+\Delta_t) = \mathbf{V}(t) + \frac{\Delta_t}{2m_I}\left[\mathbf{F_I}(t) + \mathbf{F_I}(t+\Delta_t)\right] \tag{2.72}$$

where $\mathbf{V}(t)$ denotes the velocities at time t.

A typical AIMD workflow for gas–surface simulations begins with the preparation of initial conditions. This usually involves setting up the gas–surface system with an appropriate DFT exchange–correlation functional. The preparation of the surface may include structural optimization, relaxation of surface atoms, and—in most cases—thermalization to the desired temperature through preliminary MD steps. In parallel, the initial position and velocity of the incoming gas-phase molecule are defined independently to control the collision parameters. The dynamics are then propagated by integrating Newton's equations of motion, where the interatomic forces are computed *on the fly* from the electronic structure at each time step, typically using DFT. Throughout the simulation, atomic positions, velocities, total and kinetic energies, and selected electronic properties are recorded. The resulting trajectory is subsequently analyzed to extract information about the dynamical evolution of the system, including energy transfer, reaction pathways, and statistical distributions derived from the atomic positions and velocities. In the next chapter, the specific physical systems and simulation protocols used in this work will be presented in detail.

CHAPTER 3

NO oxidation on O@HOPG from AIMD approach

Outline of the current chapter

3.1 Computational Setup	**52**
3.2 Computational Modeling of the O@HOPG surface	**53**
3.3 Validation of the NO–O@HOPG Surface Interaction Model	**66**
3.4 AIMD's initial conditions	**78**
3.5 Results and discussion	**91**
3.6 Conclusion of the chapter	**104**

This chapter provides an in-depth theoretical and computational discussion of our publication in *The Journal of Physical Chemistry C*, entitled *"Ab Initio Molecular Dynamics Calculations on NO Oxidation over Oxygen-Functionalized Highly Oriented Pyrolytic Graphite"* [153], which is accessible via the QR code shown in Figure 3.1. Building upon

Figure 3.1: QR code linking to the publication by Alou et al. [153]

the methodological framework established previously, we first detail the computational setup, emphasizing the choice of DFT, dispersion corrections, and plane-wave basis sets essential for accurately capturing the subtle interactions governing gas–surface processes. Subsequently, we systematically develop a reliable structural model for the O@HOPG surface, validating its energetic and geometric characteristics against both theoretical benchmarks and experimental findings. The chapter proceeds to rigorously validate the interaction model between NO molecules and the O@HOPG surface, with a particular focus on molecular adsorption and dissociation behaviors. Finally, we present and discuss the results of comprehensive AIMD simulations, analyzing reaction probabilities, mechanisms, and product distributions in detail.

3.1 Computational Setup

In this work, all calculations based on DFT were carried out using the *Vienna Ab Initio Simulation Package* (VASP) [154, 155], a plane-wave-based electronic structure code widely used for periodic systems and surface chemistry due to its efficiency and reliability.

The exchange–correlation energy was treated within the generalized gradient approximation (GGA) using the *Perdew–Burke–Ernzerhof* (PBE) functional [128], which offers a favorable balance between accuracy and computational cost for a broad class of materials. The choice of PBE is supported by previous studies that have demonstrated its suitability for modeling gas-surface interactions. [156–162] In the context of the present study, which focuses on the interaction of a molecule with a graphite surface, the use of the PBE functional has proven to be a reliable choice. In particular, for HOPG surfaces, Lechner et al. [163] showed that PBE yields structural and energetic properties comparable to those obtained with hybrid functionals, especially when long-range dispersion interactions are properly included.

To account for these interactions more accurately in the present study, the DFT-D3(BJ) dispersion correction developed by Grimme and co-workers was employed [142, 164]. This scheme introduces an empirical long-range dispersion term to the DFT total energy, improving the accuracy of adsorption energies and interlayer distances in weakly bound systems. Its effectiveness in capturing interplanar binding in graphite has been demonstrated in prior extensive benchmark studies [163]. In addition to these benchmarks, the PBE-D3(BJ) combination has also been successfully employed in

AIMD studies [165–169] involving graphite surfaces, further supporting its suitability for describing the structural and dynamical behavior of such systems. Moreover, its effectiveness in modeling NO adsorption on functionalized carbon surfaces has been highlighted in works by Hou et al. [70], Cen et al. [68] and Tang and Cao [69], among others.

The interaction between valence electrons and ionic cores was described using the *projector augmented-wave* (PAW) method [170, 171], which combines the computational efficiency of a pseudopotential approach with the accuracy of an all-electron description in the valence region. A plane-wave energy cutoff of 400 eV was applied and the electronic occupancies were determined using Gaussian smearing with a width of 0.05 eV.

The combination of the PBE functional, DFT-D3(BJ) dispersion correction, and PAW formalism, implemented in VASP with the aforementioned parameters, provides a robust and well-established computational framework for studying the reactivity of NO on oxygenated HOPG surfaces.

3.2 Computational Modeling of the O@HOPG surface

To perform AIMD simulations of the reaction between NO and O@HOPG, a computational model representative of the substrate and the reactive species must be constructed. While a real experimental sample of highly oriented pyrolytic graphite (HOPG) is a macroscopic polished crystal comprising on the order of 10^{18} atoms, the periodic arrangement of carbon atoms within each graphene layer allows for its representation by a smaller, computationally tractable periodic unit. Thus, the starting point is the definition of the crystalline unit cell of graphite.

3.2.1 Unit cell lattice geometry

Graphite crystallizes in a hexagonal lattice with space group $P6_3/mmc$, characterized by a layered stacking of carbon atoms along the out-of-plane c-axis. Each individual layer consists of a graphene sheet, where carbon atoms are arranged in a two-dimensional honeycomb network formed by two interpenetrating triangular sublattices. In the bulk crystal, these layers stack in an AB (Bernal) configuration, giving rise to the conventional unit cell of graphite as showed in figure 3.2.

The hexagonal unit cell is defined by two lattice vectors in the basal plane and one along the vertical direction. In Cartesian coordinates, the conventional lattice vectors can be expressed as:

$$\vec{a}_1 = a\,(1, 0, 0), \tag{3.1}$$

$$\vec{a}_2 = a\left(-\frac{1}{2}, \frac{\sqrt{3}}{2}, 0\right), \tag{3.2}$$

$$\vec{a}_3 = c\,(0, 0, 1), \tag{3.3}$$

where a and c are the in-plane and out-of-plane lattice constants, respectively.

Once the conventional lattice vectors were defined, a full structural relaxation of the graphite unit cell was performed using VASP with the previously described setup and sampling the Brillouin zone with a $16 \times 16 \times 8$ Monkhorst–Pack k-point mesh [172]. The optimized lattice parameters for graphite were found to be $a = b = 2.47$ Å and $c = 6.70$ Å, which are in very good agreement with both theoretical and experimental values previously reported in the literature. Specifically, our in-plane lattice parameter a and interlayer distance c closely match theoretical calculations performed using various dispersion-corrected DFT methods, such as $a = 2.46$ Å, $c = 6.67$ Å reported by Lechner et al. [163] using the PBE-D3(BJ) functional, $a = 2.46$ Å, $c = 6.69$ Å reported by Teobaldi et al. [173] employing the PW91+vdW functional, and $a = 2.46$ Å, $c = 6.45$ Å reported by Bucko et al. [174] using the PBE-D2 functional. Additionally, our calculated lattice parameters align closely with experimentally measured values, including those reported by Baskin and Meyer at low temperature ($a = 2.46$ Å, $c = 6.67$ Å) [175] and by Trucano and Chen via neutron diffraction at 300 K ($a = 2.46$ Å, $c = 6.71$ Å) [176].

3.2.1.1 Bulk modulus

As a further means of validating the choice of exchange–correlation functional and the inclusion of van der Waals corrections, the evaluation of the bulk modulus provides an important benchmark. The bulk modulus B_0 is a fundamental mechanical property that characterizes a material's resistance to uniform compression. In graphite, it serves as a sensitive probe of the weak interlayer van der Waals interactions, which are a defining feature of its layered structure. As such, an accurate determination of B_0 is crucial

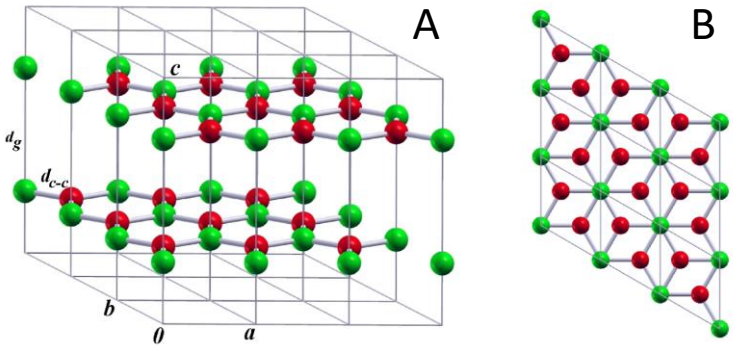

Figure 3.2: Crystal structure of highly oriented pyrolytic graphite. (A) Side view of the primitive hexagonal unit cell with lattice vectors a, b, and c; red and green spheres represent the two inequivalent carbon atoms occupying Wyckoff positions with fractional coordinates (0, 0, 1/4) and (1/3, 2/3, 1/4), respectively. The labels d_g and d_{C-C} indicate the interlayer distance and in-plane C–C bond length. (B) Top view along the (0001) direction, illustrating the hexagonal arrangement of the carbon atoms. (Source Lechner et al. [163])

for assessing the performance of dispersion-corrected exchange–correlation functionals within DFT. Comparing the computed values of B_0 with experimental data thus provides a robust test of the reliability of the theoretical approach used in this study.

To compute the bulk modulus, an energy–volume dataset was constructed and subsequently fitted to analytical equations of state (EoS). This dataset was generated starting from the previous fully relaxed graphite unit cell by systematically scaling all three lattice vectors relative to the equilibrium structure. Each new configuration corresponds to a volume $V_i = V_0 \cdot \delta_i^3$, where V_0 is the equilibrium volume and $\delta_i \in [0.9, 1.1]$ is a dimensionless scaling factor. The calculated total energies E_i, corresponding volumes V_i, and external pressures p_i are summarized in Table 3.1.

In this study, we examined and compared four commonly adopted EoS formulations, each derived from different assumptions about the compressibility of matter under pressure. The models considered include the Murnaghan, Birch, Birch–Murnaghan, and Vinet equations. Their analytical expressions are given below:

Murnaghan equation [177]

$$E_{(v)} = E_0 + \frac{B_0 V}{B'} \left[\frac{\left(\frac{V_0}{V}\right)^{B'}}{B' - 1} + 1 \right] - \frac{V_0 B_0}{B' - 1} \tag{3.4}$$

Table 3.1: Energy–volume dataset for graphite obtained from DFT calculations with PBE-D3(BJ) dispersion correction. The lattice parameters $a = b$, c, unit cell volume V, total energy E, and calculated pressure p are reported for each configuration.

a [Å]	b [Å]	c [Å]	V [Å^3]	E [eV]	p [kbar]
2.4162	2.4162	5.2002	26.29	-36.7826	354.86
2.4377	2.4377	5.5625	28.63	-37.1953	168.83
2.4455	2.4455	5.7532	29.80	-37.3187	110.39
2.4549	2.4549	6.0443	31.55	-37.4309	44.39
2.4610	2.4610	6.3112	33.10	-37.4802	7.78
2.4664	2.4664	6.6537	35.05	-37.5000	-20.41
2.4692	2.4692	6.9708	36.81	-37.4973	-32.76
2.4703	2.4703	7.1828	37.96	-37.4900	-36.47
2.4716	2.4716	7.6198	40.31	-37.4667	-39.40
2.4719	2.4719	7.8147	41.35	-37.4551	-39.41
2.4721	2.4721	8.2793	43.82	-37.4279	-37.27

Birch equation [178]

$$E_{(v)} = E_0 + \frac{9V_0B_0}{8}\left[\left[\left(\frac{V_0}{V}\right)^{\frac{2}{3}} - 1\right]^2 + \frac{9B_0V_0}{16}(B' - 4)\left[\left(\frac{V_0}{V}\right)^{\frac{2}{3}} - 1\right]^3\right] \tag{3.5}$$

Birch-Murnaghan equation [179]

$$E_{(v)} = E_0 + \frac{9V_0B_0}{16}\left[\left[\left(\frac{V_0}{V}\right)^{2/3} - 1\right]^3 B' + \left[\left(\frac{V_0}{V}\right)^{2/3} - 1\right]^2\left[6 - 4\left(\frac{V_0}{V}\right)^{2/3}\right]\right] \tag{3.6}$$

Vinet equation [180]

$$E_{(v)} = E_0 + \frac{2B_0V_0}{(B' - 1)^2}\left[2 - \left[5 + 3B'\left[\left(\frac{V_0}{V}\right)^{\frac{1}{3}} - 1\right] - 3\left(\frac{V_0}{V}\right)^{\frac{1}{3}}\right]e^{\frac{-3}{2}(B'-1)\left[\left(\frac{V_0}{V}\right)^{\frac{1}{3}}-1\right]}\right] \tag{3.7}$$

In all expressions, E_0, V_0, and B_0 denote the equilibrium energy, equilibrium volume, and bulk modulus, respectively, while B' represents the first pressure derivative of the bulk modulus. The parameters were optimized via nonlinear least-squares regression to best reproduce the energy–volume curve obtained from DFT calculations (Figure 3.3).

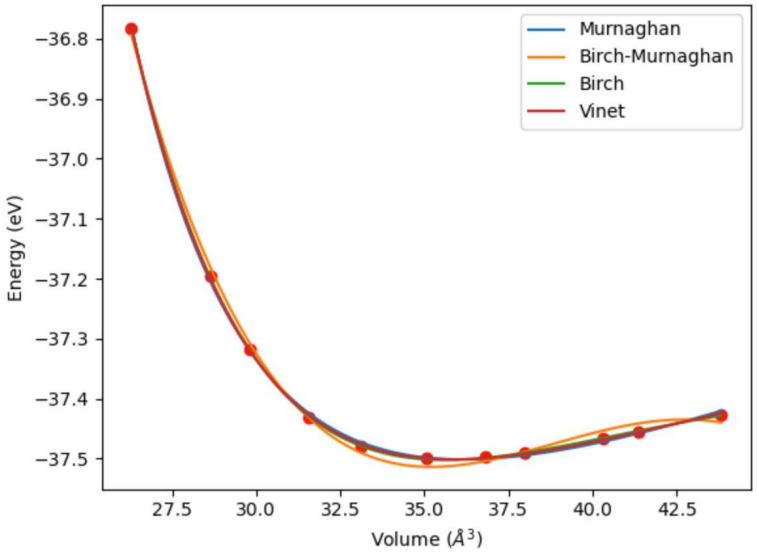

Figure 3.3: Energy versus volume curve for graphite obtained from first-principles calculations using the PBE-D3(BJ) functional. The red dots represent the computed total energies for different volumes of the fully relaxed unit cell. The solid lines correspond to fits using different equations of state: Murnaghan, Birch-Murnaghan, Birch, and Vinet. The equilibrium volume and bulk modulus were extracted from each fit and are reported in Table 3.2.

To assess the accuracy of our calculated bulk modulus for graphite, we compared the values obtained from EoS fitting using PBE-D3(BJ) with theoretical [163] and experimental benchmarks [181] (see Table 3.2). The fitted values for B_0 range from 27.2 GPa (Murnaghan) to 50.3 GPa (Birch-Murnaghan), with equilibrium volumes in the range of 35.2–36.1 Å^3. Notably, the values of B_0 obtained using the Vinet and Murnaghan equations (29.1 GPa and 27.2 GPa, respectively) are in good agreement with the theoretical value of 27.1 GPa reported by Lechner et al. [163] using a PAW basis. However, the value obtained using Birch equation (32.18 GPa) present a better approximation to the experimental interlayer bulk modulus of 33.8–36.5 GPa. The Birch-Murnaghan fit yields a significantly higher bulk modulus (50.3 GPa), which is likely due to sensitivity to the curvature of the energy–volume curve near equilibrium. The fitted values for

V_0 range from 35.20 Å^3 (Birch-Murnaghan) to 36.14 Å^3 (Murnaghan), with the Birch and Vinet models yielding intermediate values of 35.61 Å^3 and 35.84 Å^3, respectively. These results show excellent agreement with theoretical references from Lechner et al., who reported V_0 values of 35.06 Å^3 and 35.07 Å^3 using PBE-D3(BJ) with GTF and PAW basis sets, respectively. Our values also lie within the experimental range of 34.94–35.29 Å^3, though the Murnaghan fit slightly overestimates the equilibrium volume. Among the tested models, the Birch equation offers the best compromise between theoretical and experimental consistency, particularly for capturing the experimentally reported interlayer compressibility of graphite. While no single equation of state universally outperforms the others across all materials, in the case of graphite, the Birch fit provides a reliable and physically meaningful estimate of the bulk modulus. More broadly, the overall agreement of our fitted parameters—across multiple EoS formulations—with both theoretical predictions and experimental values strongly supports the adequacy of the PBE-D3(BJ) functional in accurately describing the elastic properties of graphite. This level of agreement reinforces the validity of the computational framework employed here and justifies its use in subsequent surface and dynamical studies involving van der Waals-dominated systems such as HOPG.

Table 3.2: Bulk modulus values for graphite obtained in this work by fitting the energy–volume relation using various EoS, compared with theoretical predictions from Lechner et al. and experimental reference values.

Source / EoS Model	V_0 (Å^3)	B_0 (GPa)	B'
This work (PBE-D3(BJ) (PAW basis)			
Murnaghan	36.14	27.23	8.20
Birch-Murnaghan	35.20	50.32	−5.66
Birch	35.61	32.18	10.97
Vinet	35.84	29.11	10.08
Lechner et al.			
PBE-D3 (GTF basis)	35.06	37.2	–
PBE-D3 (PAW basis)	35.07	27.1	–
Experiment	34.94–35.29	33.8–36.5	–

3.2.2 HOPG surface

Once the optimal lattice parameters were determined, the graphite surface was modeled using a slab geometry derived from the bulk structure discussed in Section 3.2.1. Considering graphite's layered hexagonal structure, the slab was naturally cleaved along the basal (0001) plane, where weak interlayer van der Waals forces ensure a well-defined and stable surface termination. The crystallographic planes of graphite are presented in Figure 3.4.

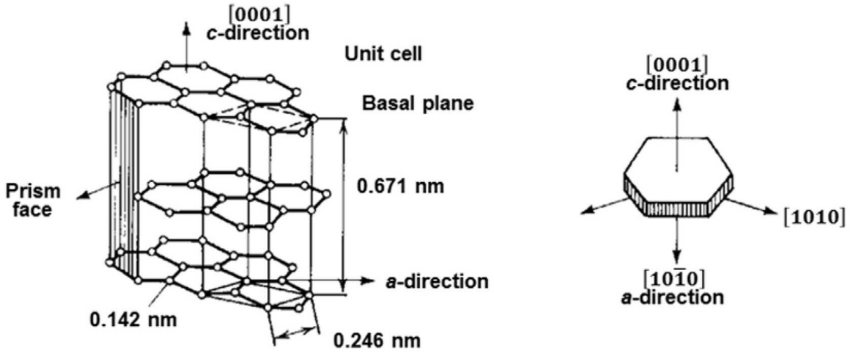

Figure 3.4: Crystallographic structure of graphite.

The slab was constructed as a $p(4 \times 4)$ supercell by repeating the in-plane lattice vectors $\vec{a}_1$ and $\vec{a}_2$ four times each. This resulted in lateral dimensions of approximately 9.87 Å along both in-plane directions where each graphene layer within the slab contained 32 carbon atoms. This provides sufficient lateral area to minimize interactions between periodic images of molecules adsorbed or scattered on the surface. The out-of-plane direction was defined by three graphene layers in the characteristic AB (Bernal) stacking, capturing the essential interlayer structure while keeping the model computationally efficient. To prevent artificial interactions along the z-axis, a vacuum layer of 20 Å was added above the slab. The resulting simulation cell, illustrated in Figure 3.5, consists of a $p(4 \times 4)$ graphite slab with three atomic layers and an explicit vacuum region.

A subsequent structural relaxation of the $p(4 \times 4)$ graphite slab was performed using a $5 \times 5 \times 1$ Monkhorst–Pack k-point mesh. During the relaxation, the bottommost layer of the slab was kept fixed to mimic bulk-like constraints and preserve the structural integrity of the substrate. The final geometry was analyzed by measuring two critical parameters:

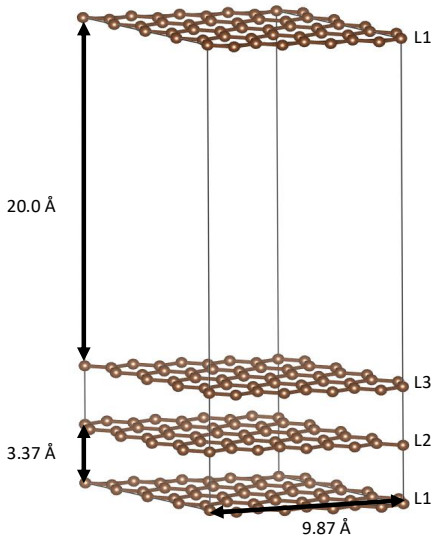

Figure 3.5: Graphite slab model used in surface simulations. The supercell consists of a $p(4 \times 4)$ surface unit, with three graphene layers (labeled L1, L2, and L3) stacked in the AB (Bernal) sequence. Each layer contains 32 carbon atoms, totaling 96 atoms in the slab. The interlayer spacing is 3.37 Å, and a vacuum region of 20 Å is added along the z-direction to eliminate spurious interactions between periodic images. The in-plane dimensions are 9.87 Å along both $\vec{a}_1$ and $\vec{a}_2$.

the average in-plane carbon–carbon distance d_{C-C} and the interlayer spacing d_g. These parameters were determined as $d_{C-C} = 1.42$ Å and $d_g = 3.37$ Å, both closely matching reference theoretical values and experimental results reported for graphite [163, 175, 181]. This confirms that the relaxed slab structure accurately reproduces the intrinsic structural characteristics of bulk graphite.

To validate the graphite slab model energetically, the cohesive energy E_{coh} was computed, which is a quantity that directly measures the slab's thermodynamic stability and is defined as:

$$E_{\text{coh}} = \frac{NE_{\text{atom}} - E_{\text{slab}}}{N}, \tag{3.8}$$

where E_{atom} is the energy of an isolated carbon atom, E_{slab} is the total energy of the slab, and N is the number of carbon atoms in the slab. The calculated cohesive energy was found to be 7.84 eV/atom, closely matching theoretical reference values reported in the literature (approximately 7.9 eV/atom) [163]. This confirms that the constructed

$p(4\times4)$ graphite slab accurately captures the intrinsic energetic characteristics of graphite, further establishing its suitability for surface interaction studies.

3.2.3 O@HOPG surface

Following the validation of the clean graphite surface model, the investigation proceeded with the adsorption of an oxygen atom to assess whether the computational setup could reliably capture the structural characteristics of the O@HOPG surface. In this sense, accurately locating the adsorption site of the chemisorbed oxygen atom is essential for establishing the correct initial configuration in the study of the surface reaction NO + O@HOPG → HOPG + NO_2. To this end, a systematic investigation was carried out by placing the oxygen atom at different adsorption sites on HOPG and computing the corresponding binding energies after structural relaxation. The binding energy (E_b) was computed using the following expression:

$$E_b = E_{HOPG} + E_O - E_{O@HOPG} \tag{3.9}$$

where E_{HOPG} is the total energy of the relaxed graphite slab, E_O is the energy of an isolated oxygen atom, and E_{HOPG+O} is the total energy of the slab with the adsorbed oxygen atom.

Three distinct adsorption sites were investigated: the bridge site (above the midpoint of a C–C bond), the hollow or center site (above the center of a carbon hexagon), and the top site (directly above a surface carbon atom) as showed in Figure 3.6. For each configuration, the oxygen atom was initially positioned at different vertical distances above the surface to explore the energy landscape and ensure convergence to a local minimum upon relaxation. The final results at each adsorption site as well as a comparison with a selection of previously published theoretical studies are summarized in the table 3.3.

Spin polarization was considered in all cases when determining the relaxed functional group on O@HOPG and the ground state always resulted to be non-magnetic (total spin equal to zero) for bridge and top adsorption sites. This behavior aligns with the theoretical understanding reported in the literature for similar systems, in particular, Incze *et al.* [182, 183], who reported that all calculations for the graphite-oxygen system were spin-unpolarized and that tests conducted with spin-polarized calculations produced

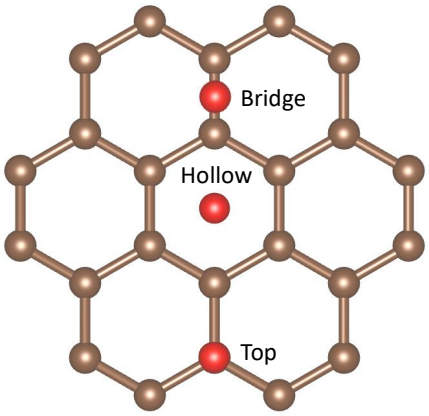

Figure 3.6: Top view of a graphene sheet illustrating the three high-symmetry adsorption sites commonly considered for O adsorption: *top* (above a carbon atom), *bridge* (above the midpoint of a C–C bond), and *hollow* (above the center of a hexagonal ring).

negligible or no changes in total energy for the stable adsorbed structures. Further confirmation is provided by Barinov *et al.*, [184] who performed spin-polarized DFT calculations on atomic oxygen adatoms on both perfect and defective graphite, explicitly stating that relaxed oxygen-related structures, such as the epoxide group, naturally adopt a singlet, non-magnetic ground state. In summary, our results align very well with theoretical studies consistently showing that while the isolated atomic oxygen in the gas phase exhibits a triplet ground state (^3P), its stable adsorption on graphite or graphene surfaces induces a transition to a singlet configuration [185–187], which likely facilitated by intersystem crossing during the adsorption process, results in a final relaxed structure that is inherently non-magnetic [188].

Specifically, in our study about the adsorption at the top site, the atomic oxygen adsorbs directly above a surface carbon atom in the basal plane after the electronic relaxation, typically forming a covalent bond with the underlying carbon. In terms of chemisorption strength, this site is commonly considered to exhibit intermediate stability, less favorable than the bridge site, yet more stable than the hollow site. For this site, our calculations yield an adsorption energy of −1.46 eV, along with a C–O bond length of 1.42 Å [Table 3.3], being this values aligns with previous theoretical studies. Incze et al. [182] reported a value of −1.28 eV for the top site at a coverage of 5.5%, with an equilibrium adsorption height of 1.61 Å, obtained using DFT calculations within the Perdew–Wang GGA functional and a fixed graphite substrate. In their work, the

Table 3.3: DFT adsorption energies and structural parameters of oxygen adsorbed on graphite/graphene at bridge and top sites.

Site	Reference	E_b (eV)	d(C–O) (Å)	d(C–C) (Å)	Z (Å)
Bridge	This work	-2.11	1.47	1.51	–
	Tian−Tian et al. [189]	-2.07	1.46	1.52	–
	Incze et al. [182]	-1.75	1.51	1.47	1.32
	Incze et al. [183]	-1.73	–	–	–
	Polynskya et al. [185]	-2.80	1.47	1.51	–
	Vasic Anicijevic et al. [190]	-1.94	–	–	1.69
	Mehmood et al. [191]	-2.40	–	–	–
	Šljivančanin et al. [187]	-2.03	–	1.51	–
Top	This work	-1.46	1.42	–	–
	Incze et al. [182]	-1.28	1.61	–	–
	Incze et al. [183]	-1.26	1.60	–	–

surface coverage was defined as the percentage of occupied sites regarding the total number of available adsorption sites, with their 3×3 graphite supercell comprising 18 possible adsorption sites, leading to a 5.5% coverage for a single adsorbed species. In the present study, our 4×4 graphite slab contains 32 available adsorption sites, so the adsorption of a single atom corresponds to a coverage of 3.1%. They also reported in a subsequent study [183] a slightly lower binding energy of −1.26 eV for the same top site and coverage, confirming the intermediate stability of this configuration when compared to the bridge and hollow sites, and further emphasizing the relatively weak chemisorption character of oxygen at this position.

Regarding the study of adsorption at the hollow site, corresponding to the geometric center of a hexagonal ring formed by six carbon atoms in the basal plane, unlike bridge or top sites, lacks direct proximity to a specific C–C bond or atomic nucleus, and is characterized by a lower electron density distribution in the bonding region. As a result, the interaction of atomic oxygen with the hollow site on HOPG is typically considered the

least favorable, and in some cases even repulsive, from a theoretical standpoint. Indeed, in our DFT calculations wasn't found an adsorption minimum for the atomic oxygen at the hollow site. This observation is consistent with the literature, for instance, Incze et al. (2001, 2003) [182, 183] explicitly reported that they did not find any significant energy minimum associated with adsorption at the hollow site. Similarly, Morón et al. (2012) [188] concluded that atomic oxygen does not adsorb favorably on hollow sites, and their DFT-derived potential energy maps show the hollow regions to be among the most repulsive on the graphite surface.

At the bridge position, following both electronic and ionic relaxation, our DFT calculations confirm that atomic oxygen remains stably adsorbed, indicating that this site corresponds to a true local minimum on the PES. The resulting configuration exhibits a well-defined epoxide-like geometry, in which the oxygen atom retains its coordination to two adjacent carbon atoms without spontaneous migration to neighboring sites. This structural arrangement induces a significant local distortion in the surrounding carbon lattice and is characterized by a C–O bond length of 1.47 Å and a C–C bond length of 1.51 Å within the strained three-membered ring. The associated adsorption energy obtained of −2.11 eV, further confirms the strong and exothermic nature of the interaction, reinforcing the thermodynamic favorability of the bridge site for atomic oxygen chemisorption on graphite [Table 3.3].

Theoretical studies of oxygen adsorption on graphite and graphene have reported a range of adsorption energies, depending on the computational setup and surface coverage. Incze et al. first identified the bridge site as the most stable adsorption configuration in 2001 [182], and later reported an adsorption energy of −1.75 eV at 5.5% surface coverage [183]. Morón et al. (2011) [186] obtained significantly less negative values of −1.28 eV and −1.26 eV for relaxed graphite surfaces, and in a follow-up study (2012) [188], proposed a broader adsorption energy range of 0.95–3.2 eV depending on the surface coverage and methodology used. These earlier studies did not account for long-range dispersion interactions, which are known to influence adsorption energies in layered systems. The inclusion of van der Waals corrections in the present work leads to a slightly more negative adsorption energy, consistent with this expectation. In line with this, Vasić Anićijević et al. (2016) [190] reported adsorption energies of −1.94 eV and −1.91 eV on graphene using 2×2 and 3×3 supercells, respectively. A study by Polynskaya et al. (2022) [185] also found comparable binding energies for

oxygen on graphene-based systems. These results further support the reliability of our computational approach.

From a structural perspective, our computed C–O bond length of 1.47 Å is in close agreement with values reported by Sorescu et al. [192] (1.47 Å), Morón et al. [186, 188] (1.48 Å, 1.49 Å), and Polynskaya et al. [185] (1.47 Å). Similarly, our C–C bond length of 1.51 Å closely matches those reported by Polynskaya [185] and Sorescu [192] (1.51 Å and 1.50 Å, respectively). Notably, this value is significantly larger than the typical C–C bond length in pristine graphite (~1.42 Å), reflecting the local strain imposed by the formation of the epoxide ring.

Further support for these theoretical findings comes from experimental evidence obtained through surface-sensitive techniques such as X-ray photoelectron spectroscopy (XPS), scanning tunneling microscopy (STM), and temperature-programmed desorption (TPD), all of which consistently identify the epoxide as the primary oxygen functional group under a broad range of conditions. For example, Kinoshita et al. [193] demonstrated the formation of epoxide groups on HOPG exposed to hyperthermal atomic oxygen (AO) beams. Their XPS spectra revealed a distinct chemical shift attributed to oxygen atoms bonded at bridge positions, while STM measurements further revealed isolated protrusions consistent with localized chemisorption events. Further evidence comes from the thermal reduction studies by Larciprete et al. [194, 195], who oxidized HOPG and graphene surfaces with AO and monitored the evolution of oxygen functionalities during annealing. At low coverages ($\theta \lesssim 0.1$ ML), XPS spectra showed that the surface was predominantly covered with epoxide groups, with no evidence of carbonyl or carboxyl functionalities. Upon heating, desorption of molecular O_2 commenced near 370 K and peaked around 430–450 K, with activation energies estimated between 1.1 and 1.2 eV, reflecting strong chemisorption consistent with the expected binding strength at bridge sites. In a complementary experiment, Paci et al. [196] combined beam-surface scattering experiments with direct dynamics simulations and showed that O_2 is the primary desorption product following AO exposure, consistent with epoxide recombination mechanisms. Although performed in tandem with theory, the experimental data showed that oxygen removal via CO or CO_2 only occurred after prolonged exposure or higher coverage, reinforcing the idea that epoxides formed at bridge sites are both stable and kinetically dominant under mild oxidation conditions. Collectively, these findings con-

verge on the conclusion that oxygen atoms initially and preferentially adsorb at bridge positions on graphite, forming stable epoxide structures.

In conclusion, our DFT results confirm that the bridge site is the most thermodynamically favorable configuration for atomic oxygen chemisorption on pristine graphite. The calculated adsorption energy of −2.11 eV, along with structural parameters characteristic of the epoxide species, provides strong evidence for the stability and nature of this adsorbed configuration. While quantitative differences in binding energies across studies are expected due to variations in computational approaches (e.g., exchange–correlation functionals, pseudopotentials, spin treatment, and supercell dimensions), the qualitative agreement in both energetic ordering and geometric features strongly supports the formation of epoxide-like structures at bridge sites. These findings are in excellent agreement with the theoretical consensus and contribute to a consistent understanding of oxygen chemisorption on graphitic surfaces.

3.3 Validation of the NO–O@HOPG Surface Interaction Model

Having validated the structural model of the O@HOPG surface, the next step involves assessing the ability of the computational setup to accurately describe the interaction between the NO molecule and the O@HOPG surface. This validation is essential to ensure that the chosen methodological framework provides a reliable description of both the adsorption energetics and the structural features associated with NO–surface interactions, prior to exploring the full molecule–surface reaction mechanism.

3.3.1 Dissociation of NO

A fundamental step in understanding the reactive dynamics of NO on surfaces is the precise characterization of its bond dissociation behavior. This analysis is essential not only to validate the accuracy of the employed exchange–correlation functional in describing the NO molecule, but also to identify the energetically accessible fragmentation channels that may be activated during its interaction with the surface. A reliable dissociation profile allows us to assess whether the functional can correctly reproduce experimental

bond energies and spin states, and provides a benchmark for evaluating possible reactive pathways in scattering or adsorption-induced dissociation events.

To this end, we performed detailed dissociation scans of the NO molecule in two environments: (i) above O@HOPG, with the molecule located inside a plane parallel to the topmost carbon layer at a high of 10 Å (Figure 3.7 A), and (ii) in vacuum, modeled by placing NO in a 20 Å cubic simulation box (Figure 3.7 B). These calculations provide a consistent framework for comparing gas-phase and surface-proximate dissociation, enabling us to quantify the influence of long-range surface effects on the energetic and magnetic evolution of NO.

The scans were conducted using a point-to-point (PtP) dissociation approach where the interatomic N–O distance is incremented in discrete steps, from a compressed bond length of $d_0 = 0.8$ Å to a fully dissociated configuration at $d_f = 4.0$ Å, and a single-point DFT calculation is performed at each geometry without atomic relaxation. This strategy was chosen to explicitly capture the abrupt spin transition that accompanies dissociation in open-shell systems. As the N–O bond is stretched, the delocalized molecular spin density evolves into two spatially separated spin-$\frac{1}{2}$ radicals. This crossover is sharply reflected in the system's total magnetic moment, which rises from $1\,\mu_B$ at equilibrium to approximately $5\,\mu_B$ beyond $d \approx 2.3$ Å. For comparison, we also applied a dissociation scheme in which the NO molecule was separated through a constrained molecular dynamics at uniform velocity (V-Const). In this approach, the atoms were incrementally displaced along the bond axis at fixed intervals, and the system evolved under frozen geometries with no internal degrees of freedom permitted to relax.

The obtained results are shown in Figure 3.7, where are represented the potential energy profiles (solid lines) and the magnetic moments of the NO molecule (dashed lines) at both environments: above an O@HOPG surface (Panel A) and in vacuum (Panel B). In all four studied cases, the energetic profile exhibits a well-defined minimum around $d_{\text{eq}} \approx 1.18$ Å, corresponding to the equilibrium bond length of NO and if the bond length increases, the total energy rises steeply, reaches a plateau near $d \approx 2.5$ Å, and remains constant thereafter, indicating full dissociation. However, regarding the magnetic moment of the system, differences in the results can be observed at each environment and methodology employed. The PtP approach shows a sharp transition from $1\,\mu_B$ at equilibrium to $2\,\mu_B$ beyond 2.3 Å over the O@HOPG surface and a two step transition from 1.75 Å to 2.3 Å in the cubic cell, confirming the formation of two spin-$\frac{1}{2}$ fragments

at both cases and reflecting the spin bifurcation at the open-shell dissociation. In the V-Const scheme, the magnetic moment remains at 1 μ_B during the dissociation, highlighting the limitations of this method in reproducing the same features. In both schemes, the V-Const energy curves show a smoother rise with distance and converge to slightly lower dissociation energies than those obtained with PtP. This difference stems from the lack of spin adaptation during constrained molecular dynamics, which tends to delay or smooth out the onset of spin polarization.

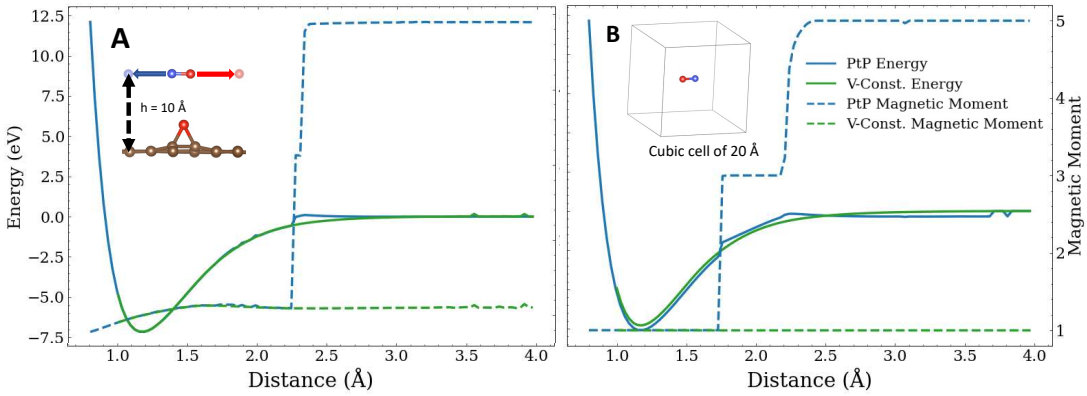

Figure 3.7: Dissociation energy profiles and corresponding magnetic moments for the NO molecule computed using two approaches: point-to-point (PtP) and constant-velocity (V-Const) dissociation. **(A)** Results for NO placed at 10 Å over the O@HOPG surface. **(B)** Results for NO in a periodic vacuum box of 20 Å; the inset illustrates the simulation cell. Solid lines represent total energies as a function of the N–O bond length, while dashed lines indicate the evolution of the system's magnetic moment. The color scheme is consistent across panels: blue for PtP and green for V-Const.

Overall, the data presented in Figure 3.7 emphasize the necessity of using spin-unrestricted point-to-point scans to resolve both the energetic and magnetic features of bond dissociation in open-shell systems such as NO. A key quantity derived from these profiles is the bond dissociation energy (BDE), which serves as a benchmark for validating the accuracy of the employed exchange–correlation functional.

$$E_{\mathrm{d}} = E(d_f) - E(d_{\mathrm{eq}}) \tag{3.10}$$

Here, $E(d_f)$ is the total energy of the system at a large internuclear separation, where the N and O atoms are effectively non-interacting, and $E(d_{\text{eq}})$ is the energy at the equilibrium bond distance. The resulting values from the PtP profiles are $E_{\text{d}} = 7.17\,\text{eV}$ over O@HOPG and 7.25 eV in vacuum. These values compare well with the experimentally determined BDE of nitric oxide, reported as 6.53 eV [197], and with high-level theoretical estimates obtained using the M06-2X [198], CBS-QB3 [199], CBS-APNO [200], and G4 [201] methods, which yields 6.51, 6.58, 6.52 and 6.55 eV respectively.

The close correspondence between the BDE computed in the vacuum and over the O@HOPG surface indicates that, at a vertical separation of 10 Å, the surface exerts negligible influence on the intrinsic bond energetics of the NO molecule. Consequently, this justifies the use of the gas-phase dissociation profile, and an associated potential energy fit for representing vibrational motion when the molecule is initialized far from the surface in dynamics molecular simulations.

Another result from our dissociation study of the NO molecule is that while spin-polarized (SP) calculations are essential to accurately describe the dissociation behavior of open-shell species such as NO, particularly the asymptotic formation of two spin-$\frac{1}{2}$ fragments, non-spin-polarized (NSP) calculations remain valid in the vicinity of the equilibrium geometry. As illustrated in Figure 3.8, both SP and NSP energy profiles exhibit nearly identical behavior in the bound region ($d \lesssim 1.65$ Å), with appreciable deviations emerging only at extended bond lengths beyond the onset of spin bifurcation. This observation confirms that NSP calculations can reliably describe the near-equilibrium energetics of NO.

3.3.1.1 Morse potential fitting and vibrational analysis

A key aspect of describing molecular vibrations in dynamics simulations is an accurate representation of the PES near the equilibrium geometry. In the case of diatomic molecules such as NO, the shape of the PES along the bond coordinate governs the vibrational spectrum and zero-point energy. For this purpose, empirical or analytical functions are often employed to reproduce the curvature and dissociative behavior of the PES in a computationally efficient manner. Among these, the Morse potential offers a simple yet physically meaningful form that captures both the harmonic behavior near

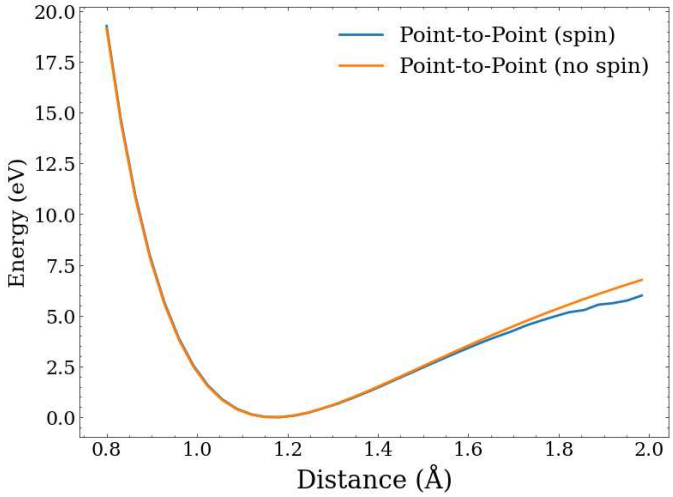

Figure 3.8: Comparison between spin-polarized (SP) and non-spin-polarized (NSP) DFT calculations of the NO dissociation profile over the O@HOPG surface. The energy curves are nearly indistinguishable in the bound region ($d \lesssim 1.65$ Å), with significant deviations arising only as the molecule approaches the dissociation limit.

equilibrium and the correct asymptotic dissociation limit. This potential is expressed in the form:

$$V(r) = D_e \left[1 - e^{-a(r-r_e)}\right]^2 \tag{3.11}$$

where D_e is the well depth, r_e the equilibrium bond length, and a a width parameter.

As previously mentioned, given the minimal influence of the substrate at a vertical distance of 10 Å, the dissociation profile computed over the O@HOPG surface was used to fit the near-equilibrium potential energy surface of NO. Then, a non-linear least squares fit to the PtP energy values using the Morse potential (Figure 3.9) yielded the following parameters: $D_e = 7.65$ eV, in good agreement with previously reported theoretical and experimental values; $r_e = 1.17$ Å, consistent with the experimental equilibrium bond length of 1.15 Å [202]; $a = 2.57$ Å^{-1}; and $R^2 = 1.0$.

From the analytical form of the Morse potential, the harmonic vibrational frequency can be derived by taking the second derivative of the potential energy with respect to the bond length at equilibrium. This yields the expression:

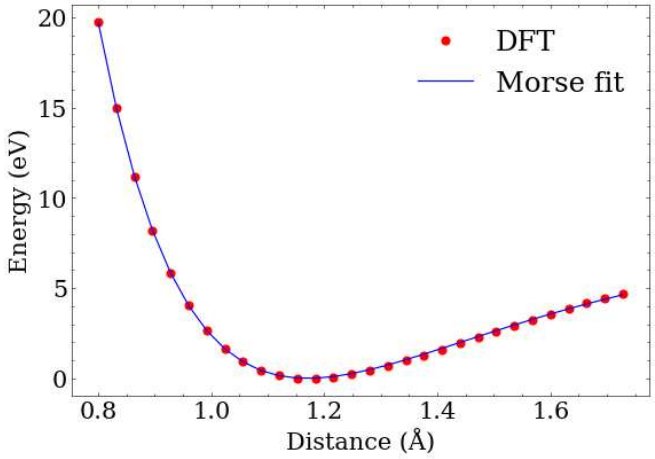

Figure 3.9: Morse potential fit to the vacuum dissociation curve of NO. Red dots: DFT energy points. Blue line: analytical Morse fit.

$$\omega_e = \frac{a}{2\pi c}\sqrt{\frac{2D_e}{\mu_{NO}}} \tag{3.12}$$

where μ_{NO} is the reduced mass of the NO molecule and c is the speed of light. The resulting frequency, $\omega_e = 1913.96\,\mathrm{cm}^{-1}$, is in excellent agreement with the experimental value $\omega_e^{\mathrm{exp}} = 1904.14\,\mathrm{cm}^{-1}$ [202], thereby confirming both the accuracy of the PBE-D3(BJ)-derived potential and the suitability of the Morse model in the vicinity of the equilibrium geometry.

3.3.2 NOx adsorption on the O@HOPG surface

Understanding the adsorption behavior of nitrogen oxides (NO and NO_2) on O@HOPG is a key step toward elucidating their reactivity and potential surface-induced transformations. In particular, the interaction of the NO molecule with the epoxy functional group and the NO_2 molecule with the HOPG surface, provides insights into their stabilization mechanisms, charge transfer characteristics, and the role of surface chemistry in mediating adsorption strength. Therefore, the aim of this analysis is to characterize the nature of NO_x adsorption on O@HOPG based on DFT calculations and establishes a theoretical framework for comparison with literature. Also, a further objective is to construct a reaction coordinate diagram for the surface-mediated oxidation process NO + O@HOPG

$\rightarrow NO_2$ + HOPG, which captures the key adsorption states and associated energetics. This graph serves as a mechanistic reference for interpreting the energetics of the AIMD simulations of the reaction, as well as a comparison with DFT-based predictions from previous theoretical studies.

3.3.2.1 NO adsorption on O@HOPG

To characterize the interaction of nitric oxide with the oxidized graphite surface, a series of spin-polarized PBE-D3(BJ) geometry optimizations were performed, exploring several initial adsorption geometries of the NO molecule on O@HOPG. The configurations varied in molecular orientation (horizontal or vertical), vertical placement above the surface, and the specific adsorption site, including N_{top}, O_{top}, and the epoxy at the center of the NO. These initial placements are illustrated in Figure 3.10.

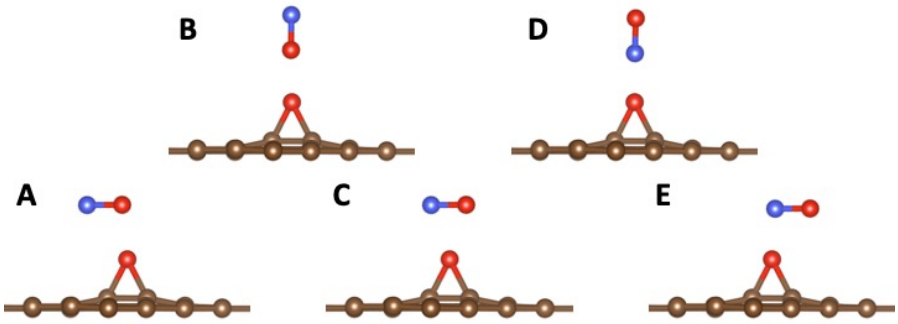

Figure 3.10: Initial configurations considered for the spin-polarized PBE-D3(BJ) geometry optimizations of NO adsorbed on the O@HOPG surface. (**A**) NO horizontal with oxygen at top, (**B**) NO vertical with oxygen at top, (**C**) NO horizontal at mid-plane, (**D**) NO vertical with N at top, and (**E**) NO horizontal with N at top.

Among all configurations examined, the most stable adsorption geometry obtained corresponds to the case where the NO molecule is initially positioned 1.0 Å above the epoxy functional group, along the direction normal to the surface, in the disposition shown in Figure 3.10 (C). This distance is defined as the vertical separation between the nitrogen atom of NO and the oxygen atom of the epoxy group. The resulting configuration

(Figure 3.11) yields an adsorption energy of −0.17 eV and retains a magnetic moment of 0.97 μ_B, indicative of a physisorbed state with minimal perturbation to the electronic structure of the NO.

This result aligns well with literature data. For example, Cen et al. [68] report a spin-unpolarized adsorption energy of −0.10 eV for NO on a graphene oxide model containing only epoxy groups, using the PBE-D2 functional. The slightly stronger interaction observed in our study can be attributed to the use of the DFT-D3(BJ) dispersion correction, which provides a more refined description of van der Waals interactions. Similarly, Tang and Cao [69] report NO adsorption energies on functionalized graphene oxide surfaces ranging from 0.19 eV on pristine graphene to higher values when hydroxyl or carbonyl functionalities are included. In both studies, the NO molecule remains at a vertical distance greater than 2.6 Å from the surface, with no significant bond distortion or magnetic quenching—features that are fully consistent with the physisorbed character observed in our results.

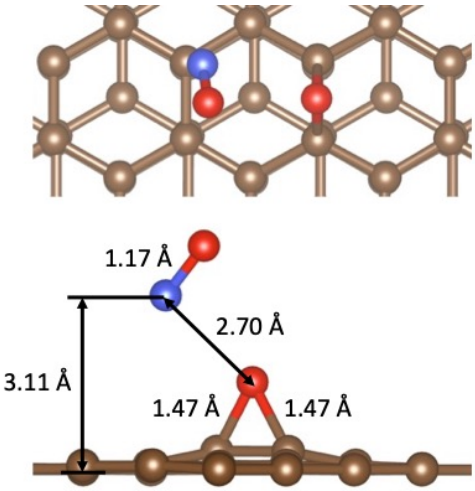

Figure 3.11: Optimized spin-polarized adsorption geometry of NO on the O@HOPG surface. The NO molecule is positioned above the epoxy group in a tilted configuration, with a N–O bond length of 1.17 Å and a N–O_{HOPG} distance of 2.70 Å. The nitrogen atom lies 3.11 Å above the surface plane, while the C–O bonds of the underlying epoxide measure 1.47 Å. This configuration corresponds to the global minimum among the adsorption structures explored.

These energetic and electronic features are further supported by the analysis of the optimized adsorption geometry. The key structural features of this configuration are

shown in Figure 3.11 where the height of the NO molecule above the surface is 3.11 Å, and the N–O bond length is 1.17 Å, consistent with the gas-phase equilibrium value reported in Section 3.3.1. The separation between the nitrogen atom and the epoxy oxygen is 2.70 Å, confirming the absence of covalent bonding and pointing to a weak interaction primarily mediated by electrostatic and dispersion forces.

The geometric parameters obtained are also in agreement with previous theoretical studies. Cen et al. [68] report an adsorption configuration where the NO molecule lies approximately 3.26 Å above the epoxy-functionalized graphene oxide surface, with no appreciable change in the intramolecular N–O bond length. Similarly, Tang and Cao [69] describe NO positioned at vertical distances between 2.6 and 3.0 Å from functionalized graphene surfaces, again observing that the N–O bond remains largely unperturbed. Both studies characterize the NO adsorption as physisorptive, which is consistent with our observed adsorption distance of 2.70 Å between NO and the epoxy oxygen.

3.3.2.2 NO_2 adsorption on O@HOPG

To characterize the adsorption behavior of nitrogen dioxide on the graphite surface, a two-step approach was employed. First, the isolated NO_2 molecule was relaxed in a vacuum box to determine its equilibrium geometry. The optimized structure corresponds to an O–N–O bond angle of 133.77° and an N–O bond length of 1.21 Å, in agreement with recent experimental reported values of 133.75° and 1.19 Å [203]. The magnetic moment of the relaxed gas-phase molecule was found to be 1.00 μ_B, confirming its open-shell character.

Following this gas-phase characterization, the optimized NO_2 molecule was positioned above the HOPG surface in several distinct initial configurations, designed to probe the influence of molecular orientation and adsorption site. Different geometries were considered, including horizontal alignments where either the oxygen or the nitrogen atom was initially oriented toward the HOPG surface (Figure 3.12 (A, B)) and a vertical alignment (Figure 3.12 (C)) where the closer oxygen atom from NO_2 to the HOPG surface was place at bridge, hollow and top positions. All structures were initialized at heights ranging from 1.0 to 4.0 Å above the surface and subjected to spin-polarized PBE-D3(BJ) geometry optimizations to assess their energetic stability and adsorption characteristics.

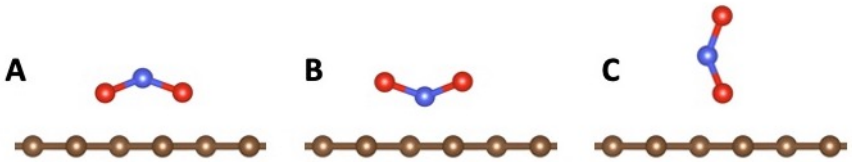

Figure 3.12: Initial configurations considered for spin-polarized PBE-D3(BJ) geometry optimizations of NO_2 on the HOPG surface. (**A**) Horizontal alignment with the NO_2 molecular plane normal to the surface and the oxygen oriented up. (**B**) Horizontal alignment with the NO_2 molecular plane normal to the surface and the oxygen oriented down. (**C**) Vertical alignment with the NO_2 molecular plane normal to the surface; the closer oxygen atom to the surface was placed above the bridge, top, and hollow sites of the underlying carbon lattice to explore the influence of the local adsorption environment.

Among all configurations examined, the most stable adsorption geometry corresponds to the case where the NO_2 molecule is horizontally aligned with the nitrogen atom oriented away from the surface. In contrast, when the nitrogen atom was initially oriented toward the surface, a spontaneous reorientation of the molecule was observed during relaxation. This behavior prevented the stabilization of a configuration in which the molecular plane remained parallel to the surface, and thus such geometries were not considered in the set of initial configurations.

The most stable disposition, illustrated in Figure 3.13, yields an adsorption energy of −0.22 eV and a magnetic moment of 0.93 μ_B, indicative of a physisorbed state with minimal perturbation to the electronic structure of the NO_2. In the optimized geometry, the N–O bond length is 1.23 Å and the O–N–O angle is 129.56°, both showing only small deviations from the gas-phase equilibrium values. The vertical distance between the nitrogen atom and the surface plane is approximately 3.60 Å, further confirming that the adsorption occurs in a weak physisorption regime dominated by van der Waals and electrostatic interactions.

This result aligns well with literature data, for example, Tang and Cao [69] report an adsorption energy of 0.19 eV for NO_2 physisorbed on pristine graphene, with an equilibrium vertical distance of 2.75 Å between the molecule and the surface. Their spin-polarized DFT calculations indicate that NO_2 retains its bent geometry and molecular character upon adsorption, with minimal charge transfer (–0.06 e), confirming a weak physisorption regime. The slightly stronger interaction observed in our study (−0.22 eV)

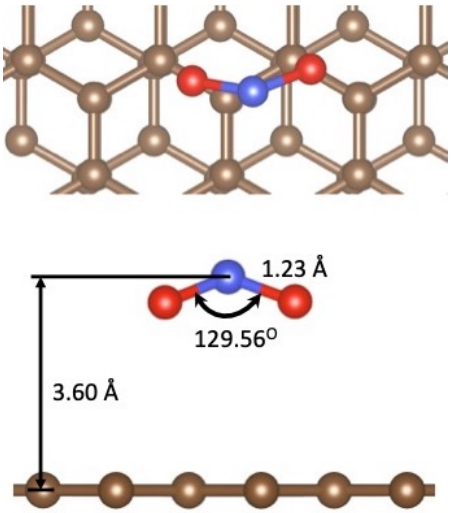

Figure 3.13: Optimized spin-polarized adsorption geometry of NO_2 on the HOPG surface. The NO_2 molecule is arranged in a triangular configuration above the HOPG surface, with the nitrogen atom tilted toward the surface. The N–O bond lengths are 1.23 Å, the O–N–O angle is 129.56°, and the vertical height of the molecule is 3.60 Å. This structure corresponds to the global minimum among the adsorption geometries explored.

can be attributed to the inclusion of Grimme's DFT-D3(BJ) dispersion correction, which provides an improved treatment of long-range van der Waals forces.

3.3.2.3 Reaction coordinates for the NO oxidation on O@HOPG

Once the adsorption geometries and energies of both NO and NO_2 on the HOPG surface were determined, a qualitative reaction pathway for the surface-mediated oxidation of NO was constructed. This pathway was designed based on insights from previous mechanistic studies, specifically the E-R mechanism, which is particularly suitable given the nature of the adsorbate and the molecular setup. In the E-R mechanism, a gas-phase NO molecule reacts directly with pre-adsorbed oxygen species on the surface, such as epoxy functional groups (O@HOPG), as proposed in studies by Adapa et al. [62] and Rathore et al. [73]. Furthermore, the direct interaction between gas-phase NO and surface-bound oxygen is consistent with experimental findings that emphasize the role of pre-adsorbed oxygen in facilitating the oxidation reaction [59]. This mechanism is particularly relevant considering the challenges associated with NO adsorption on

carbonaceous surfaces, since NO is a supercritical gas at ambient temperature, limiting its physisorption capacity [72]. Additionally, this approach aligns with reaction pathways observed in other studies where NO interacts directly with surface oxygen species [74].

To define the energetic boundaries of the reaction pathway, the NO and NO_2 molecules were initially placed 10.0 Å above the surface and relaxed in their gas-phase configurations. These configurations represent the starting and ending points of the oxidation mechanism, respectively. Between these two states, two intermediate adsorption configurations were characterized: NO adsorbed on the O@HOPG surface (NO AS) and NO_2 weakly adsorbed post-reaction (NO_2 AS). The NO AS configuration corresponds to the global minimum identified in Section 3.3.2.1, where the NO molecule is positioned 1.0 Å above the epoxy group, with a vertical height of 3.11 Å and an adsorption energy of −0.17 eV. The NO_2 AS configuration, as described in Section 3.3.2.2, features a triangular alignment of the NO_2 molecule at a distance of 3.60 Å from the graphite surface, with an adsorption energy of −0.22 eV. With these intermediate states, along with the initial and final configurations, the complete reaction pathway is depicted in Figure 3.14.

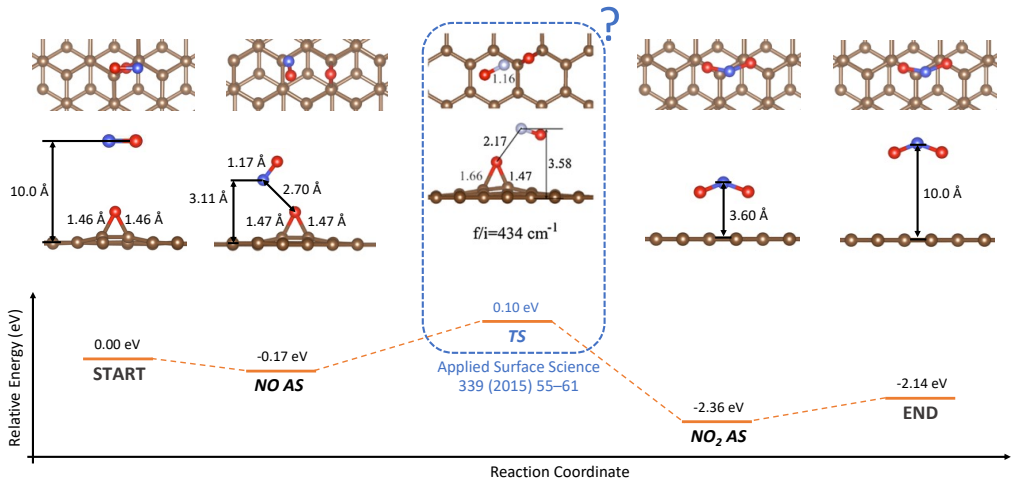

Figure 3.14: Proposed reaction coordinate for the surface-mediated oxidation of NO by an epoxy group on O@HOPG. The start and end correspond to relaxed NO and NO_2 molecules in the gas phase at 10.0 Å above the surface. Adsorbed intermediates (NO AS and NO_2 AS) represent the most stable configurations identified in Sections 3.3.2.1 and 3.3.2.2. The transition state (TS) structure and activation energy (0.10 eV) were adapted from Hou et al. [70].

In the reaction path shown in Figure 3.14, the transition state (TS) connecting NO AS and NO_2 AS is taken from the CI-NEB results reported by Hou et al. [70], who studied the oxidation of NO by an epoxy group on a graphene oxide model. Their calculations identified an activation energy of 0.10 eV, with the transition state characterized by partial N–O bond formation (2.17 Å), C–O bond elongation (1.66 and 1.47 Å), and an imaginary vibrational frequency of 434 cm^{-1}. Although this activation barrier is relatively low in the context of surface reactions, it still represents a significant threshold with respect to thermal energies at room temperature ($k_B T \approx 0.026$ eV at 300 K). Therefore, while the process may be accessible under specific conditions, the reaction is not expected to proceed efficiently at ambient temperature in the absence of additional activation mechanisms. Nevertheless, this study provides a valuable reference for constructing initial configurations in AIMD simulations of reactive trajectories.

3.4 AIMD's initial conditions

Building on the reaction coordinates established in the section 3.5.1, this section of the methodology describes the preparation of initial atomic configurations and velocities for the AIMD simulations of the reaction. The accuracy of AIMD results depends critically on initializing the system in a manner that reflects realistic thermal and structural conditions. To this end, two separate yet complementary procedures were implemented. First, the O@HOPG surface was equilibrated via a finite-temperature DFT-based molecular dynamics simulation to ensure a thermally relaxed substrate. Second, the gas-phase NO molecule was initialized through a stochastic sampling protocol designed to reproduce key statistical and dynamical properties like spatial distribution, orientation, and energy content, consistent with the collision scenarios of interest. These procedures jointly define the initial state of the NO + O@HOPG system, providing a robust and physically grounded foundation for the subsequent dynamical exploration of the reaction pathway.

3.4.1 O@HOPG surface thermalization

The initial atomic configuration of the O@HOPG surface was obtained by performing AIMD simulation in the canonical (NVT) ensemble, where the number of particles (N), the volume (V), and the temperature (T) are held constant. The objective of this

simulation was to generate atomic positions and velocities consistent with thermal equilibrium at 300 K. The choice of 300 K reflects typical ambient atmospheric conditions under which NO oxidation is known to occur on carbonaceous surfaces such as soot or oxidized graphite. Numerous experimental studies have reported efficient NO-to-NO_2 conversion at or near room temperature on oxygen-functionalized carbon materials [59, 60, 62, 74], indicating that 300 K is a relevant and representative temperature for capturing thermally activated surface dynamics.

In order to maintain the system temperature constant during the simulation, it is important to note that for a system composed of M atoms and M_{dof} active degrees of freedom, the instantaneous temperature can be computed from the total kinetic energy via the equipartition theorem:

$$T(t) = \frac{2}{k_B M_{\text{dof}}} E_M^{\text{kin}}(t) \tag{3.13}$$

where $E_{\text{kin}}(t)$ is the total instantaneous kinetic energy and k_B is the Boltzmann constant. In canonical MD simulations, the temperature is dynamically regulated through coupling to a fictitious thermal reservoir, which modifies the system's momenta to maintain a target average temperature. Algorithms that implement this coupling are referred to as thermostats, where notable examples include the Berendsen [204], Andersen [205], and Nosé–Hoover [206, 207] thermostats.

In this work, temperature control was achieved using the Nosé–Hoover thermostat (implemented in VASP), which introduces an auxiliary dynamical variable $\xi(t)$ to scale the atomic momenta $\vec{p}_I$ of each atom I. The governing equations of motion are:

$$\begin{cases} \dot{\vec{p}}_I(t) = \vec{F}_I - \xi(t)\vec{p}_I(t) \\ \dot{\xi}(t) = \frac{1}{Q}\left(\sum_{I=1}^{M} \frac{\vec{p}_I(t)^2}{m_I} - M_{\text{dof}} k_B T\right) \end{cases} \tag{3.14}$$

where $\vec{F}_I$ is the force acting on atom I, m_I is its mass, and Q is the fictitious mass associated with the thermostat, which determines the strength of coupling between the system and the thermal reservoir.

This simulation phase, known as thermalization, resulted in the relaxation of the system to thermal equilibrium. During this process, the first layer of the slab was left unmoved to facilitate stabilization. The temperature stabilized near 300 K within the first

1000 MD steps, while the kinetic and potential energy components of the Nosé–Hoover thermostat exhibited bounded, oscillatory behavior throughout the simulation. This is indicative of a properly equilibrated system and a well-tuned thermostat coupling.

Based on this well-equilibrated trajectory, 300 statistically uncorrelated configurations were randomly selected from the final 1000 MD steps. These configurations were used as thermally representative initial conditions for the subsequent AIMD simulations, ensuring structural and dynamical consistency across the generated ensemble.

3.4.2 Geometrical sampling of the NO molecule

Following the thermal equilibration of the O@HOPG substrate, the next step in preparing the initial conditions for the AIMD simulations involved defining the spatial and energetic distribution of the incoming NO molecule. In gas–surface collision studies, it is customary to assume that the adsorbate approaches the surface within a predefined interaction region. Accordingly, in the present methodology, the center-of-mass coordinates of the NO molecule (x_{CM}, y_{CM}, z_{CM}) were sampled within a circular disk of radius R oriented parallel to the XY-plane. This disk was centered above the position of the adsorbed oxygen atom on the substrate, located at $(x_{O_{\mathrm{HOPG}}}, y_{O_{\mathrm{HOPG}}}, z_{\mathrm{C.M.}})$, as illustrated in panel A of Figure 3.15. The vertical component z_{CM} was determined by adding a fixed offset $h_{\mathrm{C.M.}}$ to the average z-coordinate h_{sur} of a reference graphite layer, as expressed by:

$$z_{CM} = h_{\mathrm{C.M.}} + h_{\mathrm{sur}} \tag{3.15}$$

The x_{CM} and y_{CM} coordinates were generated using linear transformations of uniformly distributed random numbers:

$$\begin{aligned} x_{CM} &= (a_{\max} - a_{\min})\, a_{\mathrm{rand}} + a_{\min} \\ y_{CM} &= (b_{\max} - b_{\min})\, b_{\mathrm{rand}} + b_{\min} \end{aligned} \tag{3.16}$$

subject to the constraint that the sampled point lies within the circular disk:

$$(x_{O_{\mathrm{HOPG}}} - x_{CM})^2 + (y_{O_{\mathrm{HOPG}}} - y_{CM})^2 \leq R^2 \tag{3.17}$$

Here, $a_{\min} = x_{O_{\text{HOPG}}} - R/2$, $a_{\max} = x_{O_{\text{HOPG}}} + R/2$, with analogous definitions for $b_{\min}$ and $b_{\max}$. The variables a_{rand} and b_{rand} are random numbers drawn from a uniform distribution in the interval [0, 1].

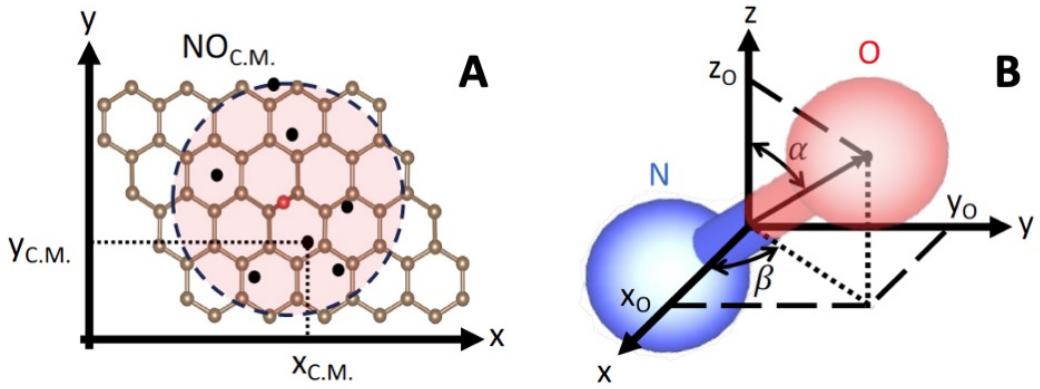

Figure 3.15: Schematic representation of the initial conditions used to sample the gas-phase NO molecule in the AIMD simulations. **(A)** Sampling region for the NO center of mass $(x_{\text{C.M.}}, y_{\text{C.M.}})$ in the XY-plane. The red disk is centered on the position of the adsorbed oxygen atom (red sphere) on the graphite surface, and black dots indicate randomly selected center-of-mass positions. **(B)** Definition of the molecular orientation in spherical coordinates. The polar angle α and azimuthal angle β determine the orientation of the NO molecular axis with respect to a framework at its center of mass.

Once the center of mass was defined, the positions of the oxygen and nitrogen atoms (x_O, y_O, z_O) and (x_N, y_N, z_N) were generated relative to the center of mass using spherical coordinates as shown in panel B of Figure 3.15. The Cartesian coordinates were then assigned according to:

$$\begin{aligned} x_{O(N)} &= x_{CM} \pm R_{\text{NO}} \sin(\alpha)\cos(\beta) \\ y_{O(N)} &= y_{CM} \pm R_{\text{NO}} \sin(\alpha)\sin(\beta) \\ z_{O(N)} &= z_{CM} \pm R_{\text{NO}} \cos(\alpha) \end{aligned} \tag{3.18}$$

where the sign determines the atom's placement relative to the center of mass along the molecular axis. The interatomic distance R_{NO} corresponds in this case to the equilibrium bond length. However, if the internal momentum of the NO molecule is initialized to reflect its zero-point energy (ZPE), R_{NO} must be recalculated, as discussed in subsequent

sections. The orientation of the molecular axis was defined by randomly sampled polar and azimuthal angles:

$$\begin{aligned} \alpha &= \arccos(1 - 2\gamma) \\ \beta &= 2\pi\gamma \end{aligned} \tag{3.19}$$

where γ is a uniformly distributed random number in the interval $[0, 1]$.

3.4.3 Initial velocity vector of the NO molecule

Following the geometrical sampling of the NO molecule described in the previous section, the next step is to compute the velocity vector of the NO center of mass. This velocity is obtained from the translational kinetic energy $E_{\mathrm{NO}}^{\mathrm{trans}}$ of to the NO molecule, which in our simulations, corresponds to the prescribed collision energy $E_{\mathrm{NO}}^{\mathrm{col}}$ used to control the dynamical regime of the gas–surface interaction. In this context, the magnitude of the center-of-mass velocity vector $\|\vec{v}_{\mathrm{NO}}\|$ was computed from the classical relation between translational kinetic energy and velocity:

$$\|\vec{v}_{\mathrm{NO}}\| = \sqrt{\frac{2E_{\mathrm{NO}}^{\mathrm{col}}}{m_{\mathrm{NO}}}} \tag{3.20}$$

where $m_{\mathrm{NO}} = m_{\mathrm{N}} + m_{\mathrm{O}}$ is the total mass of the molecule. The full center-of-mass velocity vector $\vec{v}_{\mathrm{NO}}$ was then obtained by multiplying this magnitude by a unit vector $\hat{v}$ defining the direction of incidence in a three-dimensional space:

$$\vec{v}_{\mathrm{NO}} = \|\vec{v}_{\mathrm{NO}}\| \cdot \hat{v} \tag{3.21}$$

The orientation of the velocity vector $\hat{v}$ was defined in spherical coordinates, where the polar angle φ is measured from the positive z-axis and the azimuthal angle θ is measured in the xy-plane from the positive x-axis as shown in the Figure 3.16. In this convention, the cartesian components of the unit vector $\hat{v}$ are given by:

$$\begin{aligned} \hat{v}_x &= \sin(\varphi)\sin(\theta) \\ \hat{v}_y &= \sin(\varphi)\cos(\theta) \\ \hat{v}_z &= \cos(\varphi) \end{aligned} \tag{3.22}$$

Two types of incidence conditions were considered, depending on the desired impact geometry between the NO molecule and the functionalized surface. In the case of *normal incidence*, the velocity vector was aligned perpendicularly to the graphite plane, corresponding to a polar angle of $\varphi = 0°$, such that in cartesian coordinates:

$$\hat{v}_{\text{normal}} = (0, 0, 1) \tag{3.23}$$

Alternatively, under *oriented incidence* conditions, the translational velocity vector was directed toward the surface oxygen atom in order to enhance the probability of a reactive NO–O_{HOPG} encounter. In this configuration, the polar and azimuthal angles were calculated from the relative position vector $\vec{r}_{\text{CM-O}} = \vec{r}_{O_{\text{HOPG}}} - \vec{r}_{\text{CM}}$, connecting the NO center of mass to the surface oxygen. The orientation was then defined according to:

$$\begin{aligned} \varphi &= \arcsin\left(\frac{\left|\vec{r}^{\,XY}_{\text{CM-O}}\right|}{\left|\vec{r}_{\text{CM-O}}\right|}\right) \\ \theta &= \arccos\left(\frac{\left|\vec{r}^{\,Y}_{\text{CM-O}}\right|}{\left|\vec{r}^{\,XY}_{\text{CM-O}}\right|}\right) \end{aligned} \tag{3.24}$$

where $\vec{r}^{\,XY}_{\text{CM-O}}$ and $\vec{r}^{\,Y}_{\text{CM-O}}$ denote the projections of $\vec{r}_{\text{CM-O}}$ onto the xy-plane and the y-axis, respectively.

3.4.3.1 Adapting the internal momentum of NO to its ZPE

Once the translational velocity vector of the NO was determined, a quasi-classical approach was implemented to adapt the internal momentum of the NO to its quantum mechanical ZPE. In this framework, the rovibrational levels of the NO molecule need to be determined, where first is necessary to obtain a continuous representation of the dissociation curve of the molecule ($V(R_{NO})$). In this study, it was obtained by fitting to the predefined analytical potential of Morse the potential energy at discrete values of the internuclear distance R_{NO}, as described in section 3.3.1.1.

Once the function $V(R_{NO})$ was obtained, a grid of uniformly spaced points R^i_{NO} was constructed to discretize the space between a minimal internuclear distance R^{min}_{NO} and a maximal internuclear distance R^{max}_{NO}, using a step ΔR_{NO}. Using this grid, the effective potential $V_{eff}(R^i_{NO})$ is calculated, where the analytical potential $V(R_{NO})$ is determined at

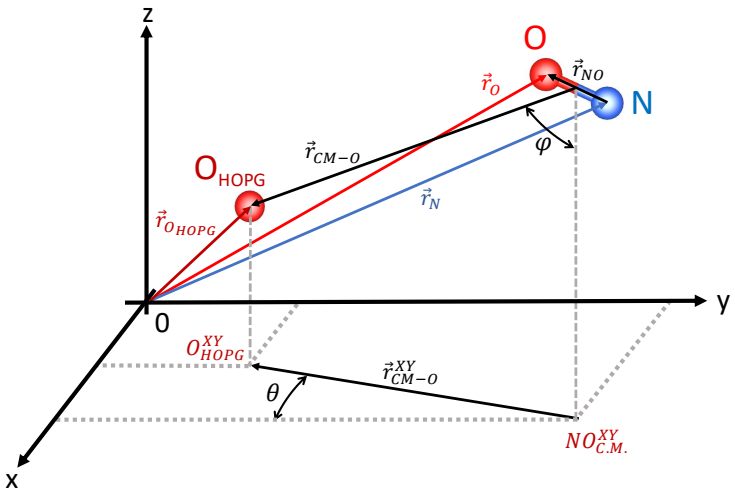

Figure 3.16: Schematic representation of the initial spatial arrangement of the NO molecule relative to the graphite surface. The vector $\vec{r}_{O_{\mathrm{HOPG}}}$ represents the position of the oxygen atom adsorbed on the surface, while $\vec{r}_{\mathrm{CM\text{-}O}}$ connects the center of mass of NO to this reactive site. The polar angle φ and azimuthal angle θ define the orientation of the velocity vector relative to $\vec{r}_{\mathrm{CM\text{-}O}}$, with dotted lines indicating the projection of vectors onto the *xy*-plane.

each grid point R^i_{NO}, and the contribution of rotational energy is added to account for the effects of rotation of the diatomic molecule on its energy levels:

$$V_{eff}(R^i_{NO}) = V(R^i_{NO}) + \frac{j(j+1)}{2\mu_{NO}(R^i_{NO})^2} \tag{3.25}$$

where j is the rotational quantum number and μ_{NO} is the reduced mass of the NO molecule.

Using the "Fourier Grid Hamiltonian" method developed by Balint-Kurti *et al.* [208, 209], the eigenvalues and eigenfunctions of the bound states are calculated. The Hamiltonian matrix **H** is obtained, with diagonal elements ($a = b$) calculated by:

$$H_{aa} = \frac{h^2}{4\mu_{NO}L^2}\left[\frac{N^2+2}{6}\right] + V_{eff}(R^i_{NO}) \tag{3.26}$$

where h is Planck's constant, L is the grid length, N is the number of points in the grid and the non-diagonal elements ($a \neq b$) are calculated by:

$$H_{ab} = \frac{(-1)^{(a-b)}}{\mu_{NO}} \left\{ \frac{h}{2L \sin[\pi(a-b)/N]} \right\}^2 \tag{3.27}$$

Finally, the Hamiltonian matrix is diagonalized to obtain eigenvalues and eigenvectors. The eigenvalues E_{vj} correspond to the rovibrational energy levels of the system, while the eigenvectors ψ_{vj} represent the associated wavefunctions. To ensure consistency with the quantum mechanical zero-point motion of NO, the vibrational and rotational ground state $v = j = 0$ were selected and this values were used as the reference rovibrational energy for the quasi-classical initialization of the intramolecular motion.

In order to initialize the intramolecular motion of the molecule, it was taken into account that the internal energy of NO is equal to the sum of its rotational kinetic energy E_{NO}^{rot} and its vibrational energy, which has both kinetic and potential terms denoted as $E_{NO}^{vib,tot} = E_{NO}^{vib,kin} + V_{NO}(R_{NO})$:

$$E_{NO}^{int} = E_{NO}^{rot} + E_{NO}^{vib,kin} + V_{NO}(R_{NO}) \tag{3.28}$$

where the rotational kinetic energy is zero since the molecule was initialized rotationally cold ($j = 0$) and the vibrational kinetic energy is defined as the kinetic energy of the molecule along the direction of the internuclear axis $\vec{r}_{NO}$ (Figure 3.16):

$$E_{NO}^{vib,kin} = \frac{1}{2} \frac{\|\vec{p}_{NO}^{\,R}\|^2}{\mu_{NO}} \tag{3.29}$$

where $\vec{p}_{NO}^{\,R}$ is the internal momentum. Then, from the equations 3.28, 3.29 and the E_{vj} corresponding to the ZPE of the NO molecule (E_{00}) obtained from the Hamiltonian matrix diagonalized, the internal momentum of the NO molecule can be computed as:

$$\|\vec{p}_{NO}^{\,R}\|^2 = \sqrt{2\mu_{\mathrm{NO}} \left(E_{00} - V_{\mathrm{eff}}(R(t))\right)} \tag{3.30}$$

In this equation, the value of the internuclear distance at time t was obtained from the integral:

$$R(t) = \int_0^t \sqrt{\frac{2}{\mu_{\mathrm{NO}}} \left(E_{00} - V_{\mathrm{eff}}(r)\right)}\, dt \tag{3.31}$$

where $t = \delta\tau$ within the rovibrational period τ, being δ a random rovibrational phase $\delta \in [0, 1]$. The period τ was evaluated as:

$$\tau = 2 \int_{R_{\min}}^{R_{\max}} \frac{dr}{\sqrt{\frac{2}{\mu_{NO}} (E_{00} - V_{\text{eff}}(r))}} \tag{3.32}$$

This procedure ensured that the NO molecule was initialized in a geometrically and dynamically consistent configuration with respect to its quantum mechanical vibrational ground state.

3.4.3.2 Assignment of atomic velocities in the NO molecule

Finally, the total energy of the NO molecule ($E^{\text{tot}}_{\text{NO}}$) is the sum of its translational kinetic energy ($E^{\text{kin}}_{\text{NO}}$) and its internal energy ($E^{\text{int}}_{\text{NO}}$):

$$E^{\text{tot}}_{\text{NO}} = E^{\text{kin}}_{\text{NO}} + E^{\text{int}}_{\text{NO}} \tag{3.33}$$

Therefore, the velocity vectors of the oxygen and nitrogen atoms in the NO molecule were constructed by combining the translational motion of the center of mass with the internal vibrational motion along the bond axis. The final atomic velocities $\vec{V}_O$ and $\vec{V}_N$ were defined as:

$$\begin{aligned} \vec{v}_N &= \vec{v}_{\text{NO}} + \vec{p}^{\,R}_{\text{NO}}/m_N \\ \vec{v}_O &= \vec{v}_{\text{NO}} - \vec{p}^{\,R}_{\text{NO}}/m_O \end{aligned} \tag{3.34}$$

where $\vec{v}_{\text{NO}}$ is the translational velocity vector of the center of mass from equation 3.21 and $\vec{p}^{\,R}_{\text{NO}}$ is the internal vibrational momentum resulting from the quasi-classical phase sampling of the vibrational motion.

3.4.4 Molecular dynamics simulations

Once the procedure to define the initial conditions had been established, the next step was to determine the collision energy of the incoming NO molecules. Four distinct center-of-mass incidence energies were considered: 25 meV (290 K), 50 meV (580 K), 100 meV (1160 K), and 300 meV (3481 K), where the temperatures in parentheses correspond to the translational energy of NO_{CM}, excluding contributions from ZPE.

These selected values span a broad energy range, from thermal conditions representative of the ambient atmosphere (25 meV), to high-energy collisions (300 meV), thereby allowing us to probe a variety of dynamical regimes. Notably, the intermediate energy of 100 meV corresponds closely to the previous estimated activation barrier for NO adsorption on oxidized graphene, as reported by Hou *et al.* [70]. Furthermore, this range encompasses typical values employed in surface scattering studies to investigate energy transfer, adsorption, and reactive events involving nitrogen oxides on carbon-based materials [210–212].

At each of the selected collision energies, 300 initial configurations for the center of mass of the colliding NO molecule (denoted NO_{CM}) were randomly sampled within a disk of radius $R = 5$ Å, centered on the surface-adsorbed oxygen atom and located at a height $H = 9$ Å above the oxidized graphite surface. The choice of $R = 5$ Å was dictated by the in-plane dimensions of the simulation cell, ensuring that the NO molecule is always placed within the lateral bounds of the periodic surface. The vertical distance of $H = 9$ Å was selected to guarantee that the NO molecule is initialized sufficiently far from the surface, such that its initial interaction with the substrate occurs in the gas phase regime. At this height, the potential energy landscape is essentially flat and the influence of the surface is negligible, thus justifying the assumption of gas-phase initial conditions. For each energy, two distinct incidence conditions were considered: one corresponding to normal incidence (velocity vector perpendicular to the surface), and the other oriented such that the trajectory of NO_{CM} was directed toward the adsorbed oxygen atom, thereby enhancing the likelihood of direct NO–O interaction (Figure 3.17). Taking into account the four collision energies and two incidence schemes, a total of 2400 AIMD initial conditions were established.

Before conducting extensive AIMD simulations, we performed tests by running trajectories with a sample of initial conditions using both SP and NSP approaches to evaluate the calculation strategy. The results from SP calculations showed no significant variation in total system magnetization for both reactive and non-reactive trajectories during all the dynamics. Specifically, the total magnetization remains almost constant at 1 μ_B throughout the NO molecule's approach to the surface, during the crucial interaction phase with the O atom, and after the scattering of the NO molecule or the formation of the NO_2 molecule. The NSP trajectories explored the same configuration space as those in SP calculations, indicating that both SP and NSP multidimensional PES exhibit

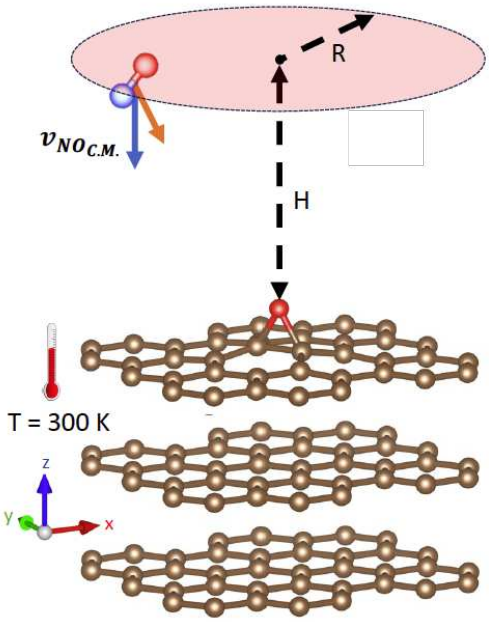

Figure 3.17: Schematic representation of the initial configuration for the quasi-classical AIMD simulations. The center of mass of the NO molecule is randomly placed inside a disk of radius $R = 5$ Å centered on the surface oxygen atom and located at a height $H = 9$ Å above the O@HOPG surface. The surface is thermalized at $T = 300$ K. The direction of the NO translational velocity vector $\vec{v}_{\mathrm{NO_{CM}}}$ defines the incidence conditions (normal (blue vector) and oriented (orange vector)).

similar characteristics. This can be attributed to the fact that in SP calculations, the total magnetization of the system remains almost constant during the trajectory, a behavior that is inherently preserved in NSP calculations, where the system's magnetization is not considered. Based on these observations, the NSP calculation strategy was chosen for extensive AIMD calculations, as it reliably describes the interactions between NO and the O@HOPG surface under the simulation conditions studied. The choice of the NSP approach also significantly reduces computational cost and improves the stability of the AIMD calculations.

Finally, under these conditions, constant-energy (NVE) non-spin polarized AIMD simulations were performed with a maximum propagation time of 4 ps and a timestep of 1 fs, making the NVE regime a reasonable approximation given the short simulation times considered. Four possible exit channels were observed: (i) two reactive pathways—NO_2 scattered or NO_2 adsorbed; and (ii) two non-reactive outcomes—NO scattered or NO

adsorbed. Since the dissociation energy of NO on the O@HOPG surface was calculated to be 4.2 eV, the dissociative adsorption channel of NO is energetically inaccessible under the present conditions. Molecules (NO or NO_2) were considered desorbed when their center of mass reached a distance of 7 Å above the surface; otherwise, if this condition was not met within the 4 ps time window, they were classified as adsorbed. The total energy was well conserved along the trajectories, with a standard deviation of approximately 6 meV.

Reflected trajectories were discriminated according to the difference between the initial NO molecule incidence angle θ_i and the final product (NO or NO_2) scattering angle θ_s, computed as $\Delta\theta = \theta_s - \theta_i$. Angular distributions of the desorbed molecules were calculated using constant intervals of 9°. The energy distributions of the NO molecules along the dynamics were obtained using the standard semi-classical determination method for the translational, vibrational and rotational energies [213]. Further details on the computational procedure and analysis will be provided in the following sections.

3.4.4.1 Translational distributions

The translational kinetic energy of the NO_x (NO or NO_2) scattered molecules was computed as follows:

$$E^{\text{kin}}_{\text{NO}_x} = \frac{1}{2}\|\vec{v}_{\text{NO}_x}\|^2 m_{\text{NO}_x} \tag{3.35}$$

where m_{NO_x} is the mass of the NO_x molecule, and $\vec{v}_{\text{NO}_x}$ is the velocity vector of its center-of-mass, which is computed from the atomic velocities obtained from the XDATCAR file generated by the AIMD trajectory in VASP, using finite difference methods.

From the obtained values of the translational kinetic energy, histograms were constructed to represent the product translational energy distribution. A specific bin width was selected, and the events within each bin were counted and then divided by the total number of scattered molecules to determine the probability of occurrence.

3.4.4.2 Rovibrational distributions

The rotational quantum number j specifying the scattered NO molecule rotational state is assigned first by applying the quantization of the classic rotational angular momentum,

$\|\vec{j}_{\mathrm{NO}}\|^2 = j(j+1)\hbar^2$, where the rotational angular momentum $\vec{j}_{\mathrm{NO}}$ is computed from $\vec{j}_{\mathrm{NO}} = \vec{R}_{\mathrm{NO}} \wedge \vec{p}^{\,rel}_{\mathrm{NO}}$, where $\vec{R}_{\mathrm{NO}} = \vec{r}_{\mathrm{O}} - \vec{r}_{\mathrm{N}}$ is the internuclear distance vector between the oxygen and nitrogen atoms, and $\vec{p}^{\,\mathrm{rel}}_{\mathrm{NO}} = \mu_{\mathrm{NO}}\vec{v}_{\mathrm{rel}}$, with $\mu_{\mathrm{NO}} = \frac{m_{\mathrm{O}}m_{\mathrm{N}}}{m_{\mathrm{O}}+m_{\mathrm{N}}}$ the reduced mass, and $\vec{v}_{\mathrm{rel}} = \vec{v}_{\mathrm{O}} - \vec{v}_{\mathrm{N}}$ the relative velocity vector. A real value of j is thus obtained from:

$$\|\vec{j}_{\mathrm{NO}}\|^2 = j(j+1)\hbar^2 \quad \Rightarrow \quad j \approx \frac{1}{2}\sqrt{1 + \frac{4\|\vec{j}_{\mathrm{NO}}\|^2}{\hbar^2}} - \frac{1}{2} \tag{3.36}$$

where the resulting value was rounded to the nearest integer following the standard binning procedure.

Finally, the vibrational quantum number v of the of the NO scattered molecule is computed by the semi-classical quantization of the vibrational action integral [214–217]:

$$\oint p^R(r')\,dr' \approx 2\pi\left(v + \frac{1}{2}\right) \tag{3.37}$$

where $p^R(R_{\mathrm{NO}})$ is the conjugate momentum of the NO bond length R_{NO} obtained as:

$$p^R(R_{\mathrm{NO}}) = \sqrt{2\mu_{NO}\left(E^{int}_{NO} - V_{NO}(R_{NO}) - E^{rot}_{NO}\right)} \tag{3.38}$$

The integral 3.37 is evaluated over a vibration period, between the two turning points R^{min}_{NO} and R^{max}_{NO} of the potential $V_{NO}(R_{NO})$ corresponding to the final internal energy E^{int}_{NO} of the NO scattered molecule. The real vibrational quantum number v is then rounded to the nearest integer following the standard binning procedure.

The histogram representations of the vibrational and rotational distributions were obtained using bin sizes equal to 1, ranging from 0 to $v_{\max}$ or $j_{\max}$. The events within each bin were counted and then divided by the total number of scattered molecules to determine the probability of occurrence.

Regarding the characterization of the rovibrational state of the triatomic NO_2 molecule, a multidimensional PES describing its internal motion is required. The construction of this PES is currently underway and will ultimately allow for the quantification of rovibrational energy distributions of NO_2 scattered molecules.

3.5 Results and discussion

This section presents the outcomes of the AIMD simulations performed to investigate the oxidation of NO on O@HOPG surfaces. The results are analyzed in terms of the observed reaction probabilities, underlying mechanistic pathways, scattering behavior of the products, and the energy redistribution among translational, vibrational, and rotational degrees of freedom. Emphasis is placed on understanding how the incidence energy and initial orientation of the NO molecule affect the probability and dynamics of NO_2 formation, and on identifying the geometrical and energetic requirements that favor reactivity. The discussion begins with an analysis of the reaction probabilities and mechanistic trends observed across the simulated trajectories.

3.5.1 Reaction probabilities and reaction mechanism

Table 3.4 and Figure 3.18 summarizes the probability of occurrence for the observed exit channels at the incidence energies studied, considering two different orientations of the initial velocity vector (normal and oriented incidences). Total probabilities are computed by adding adsorbed and scattered molecules for both reactive and non-reactive channels. The results show a dependence with respect to the initial orientation of the NO molecule velocity vector. For normal incidence, the reaction probability (NO_2 formation) is low, in the range 7.3 % - 9.3 %, and does not show dependence on the initial NO kinetic energy. On the other hand, when the NO incidence is oriented to target the O atom on the HOPG surface (in parentheses), a clear increase in the reaction probability is observed, rising from 13.6 % to 26.3 % as the initial kinetic energy of the incident NO molecule increases from 0.025 to 0.300 eV.

The distinct behavior observed can be explained by analyzing the initial position of the NO_{CM} of the impinging molecule. Our calculations demonstrate that, irrespective of the incidence energy, the probability of reaction is significantly enhanced when the NO_{CM} is initially located in the *xy* plane inside a disk of radius 2.5 Å centered on the epoxy surface group. More than 89 % and 72 % of reactive events take place in this region for normal and oriented incidences, respectively. This condition is crucial for normal incidence; if the NO_{CM} is outside of this disk, the oxidation reaction is very unlikely to occur. In fact, our calculations indicate that under these conditions, NO_2 is only formed

Table 3.4: Reaction and scattering probabilities of the 300 trajectories at each incidence energy for normal incidence and oriented incidence (in parentheses)

E_k (eV)	NO_2 (reactive channel)			NO (non-reactive channel)		
	Scattered (%)	Adsorbed (%)	Total (%)	Scattered (%)	Adsorbed (%)	Total (%)
0.025	4.33 (6.70)	5.00 (7.00)	9.33 (13.6)	5.67 (1.60)	85.0 (84.7)	90.7 (86.4)
0.050	6.33 (10.7)	1.00 (6.30)	7.33 (17.0)	11.0 (7.60)	81.7 (75.3)	92.7 (83.0)
0.100	6.33 (12.3)	2.33 (7.30)	8.67 (19.6)	23.7 (15.3)	67.7 (65.0)	91.3 (80.4)
0.300	7.33 (20.0)	1.67 (6.30)	9.00 (26.3)	57.0 (35.7)	34.0 (38.0)	91.0 (73.7)

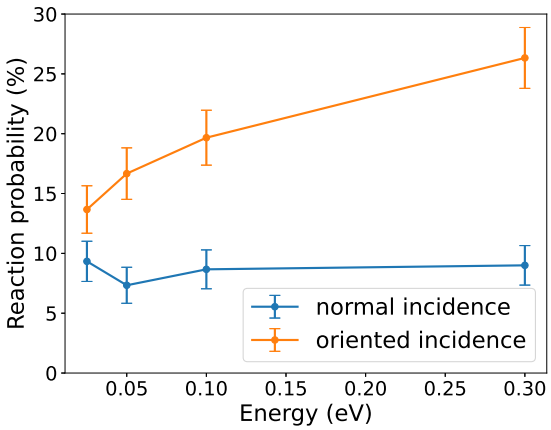

Figure 3.18: Reaction probabilities as a function of the collision energy for both normal and oriented incidences. The probability errors were calculated by the usual formula for binomial distributions.

when the NO molecule bounces at least once on the surface, i.e., in an indirect manner. For cases of oriented incidence, we also observe a decrease in the reaction probability when the position of the NO molecule c.m in the *xy* plane is far from the O atom on the surface, this decrease can be linked to an earlier reaction barrier. However, this probability reduction is compensated by the initial orientation of the NO velocity vector, which favors the direct encounter between the NO molecule and the oxygen atom of the HOPG surface. The higher the total incident energy of NO and the normal component of its velocity vector, the greater the probability of NO_2 formation. If the distance, in the *xy* plane, between the NO_{CM} and the oxygen atom on the surface is greater than 2.5 Å, the molecule must approach with a favorable orientation and sufficient energy to

trigger the reaction. In contrast, when the NO center of mass position in the *xy* plane falls within a radius of 2.5 Å around the epoxy group on the surface, the reaction becomes substantially more probable and is independent of the incidence energy. Remarkably, even when the reaction probabilities are computed using only the trajectories that start within the restricted disk radius of 2.5 Å, the oriented-incidence probability remains higher than that of normal incidence. This confirms that favoring the NO-O collision is fundamental to increasing NO oxidation on O@HOPG surfaces and suggests that the reaction predominantly follows a direct Eley-Rideal mechanism.

Interestingly, our calculations show that the reaction (NO_2 formation) can occur even at the lowest collision energy investigated (0.025 eV), roughly equivalent to room temperature (300 K). These findings qualitatively align with earlier experimental observations, which indicate a high likelihood of NO oxidation on oxidized graphite surfaces at room temperature. [60–62] Quantitatively, the reaction probabilities observed in our simulation are relatively low, amounting to less than 20 %, in contrast to the experimental probabilities [61] (~ 80 %). However, we must take into account that performing a direct comparison of the simulation results with the experimental data is challenging because of the characteristics of the simulations (simulation time, number of incident molecules, number of epoxy groups on the HOPG surfaces). In this sense, our results indicate the right trend, if an encounter between a NO molecule and an epoxy group on the surface of HOPG occurs, the possibility of the reaction is not null, it is expected that the multiplication of these encounters over a long period of time may result in a higher reaction probability as the experimental observations show.

From a theoretical perspective, no studies exist regarding the system under investigation; however, Hou et al. [70], using the nudged elastic band (NEB) method based on DFT calculations, identified an energy barrier of 0.1 eV for NO oxidation in graphene oxides, a system with characteristics similar to the one studied. Although the system is not exactly the same, given their similarities, we could expect a similar barrier in our system. If this were the case, this barrier would still be significant enough to observe reactivity at room temperature. Interestingly, our AIMD trajectories indicate that under room temperature conditions, the reaction barrier can be confidently estimated to be below 0.025 eV. This may be due to a lower reaction barrier for NO oxidation on graphite surfaces compared to graphene ones, or the influence of surface temperature, which, in some cases, could effectively lower the reaction barrier. In this regard, the

effect of surface temperature in enhancing O desorption on oxidized graphite has been previously observed. For instance, at temperatures exceeding 370 K [194, 218, 219], oxygen desorption becomes easier, primarily in the form of molecular oxygen (O_2) after O-O recombination. A similar effect could be expected when NO encounters O atoms on the surface.

The key to understanding the reactivity in this system lies in the exothermic nature of the NO oxidation reaction on the O@HOPG surface. The top panel of Figure 3.19 shows the typical evolution of the potential energy along a reactive trajectory. As can be seen, there is a sudden decrease in potential energy during the reaction. This decline occurs at a specific moment, termed the initial state (IS), marking when the system attains the necessary structural and energetic conditions for the reaction to commence. Subsequently, a notable amount of potential energy is liberated within the system, converting into kinetic energy until the potential energy stabilizes once more. At this point, henceforth identified as the final state (FS), the reaction has already occurred and the NO_2 molecule has been formed. In practical manner, the identification of the IS and the FS for each trajectory was based on specific criteria related to the potential energy profile of the system. First, in order to reduce the noise, a Savitzky-Golay filter [220] have been applied to the potential energy data, using a window size of 199 points and a polynomial order of 2, second the IS was identified as the point where the potential energy of the system start to decrease and the FS as the point where the system's potential energy stabilizes after the potential energy drop. It should be noted that this numerical procedure can occasionally lead to inaccuracies in the determination of the IS and FS states.

The availability of kinetic energy plays a pivotal role in elucidating reactivity; once the reaction conditions are met, the kinetic energy in the system is sufficient to break the bonds between the O and C atoms on the surface, providing enough translational energy for the formed NO_2 molecule to move away from the surface. Notably, the average duration of the reaction process, from the IS to the FS, is independent of both, the incident energy and the type of incidence, with a distribution that spans from 75 fs to 250 fs, with most trajectories centered around an average value of 143 fs as showed in panel B of Figure 3.19. The obtained reaction time distribution indicates a rapid oxidation process, suggesting a direct reaction pathway.

Deepening on the exothermicity of the reaction, the panel A in the Figure 3.19 represents the distribution of the potential energy differences obtained between the FS

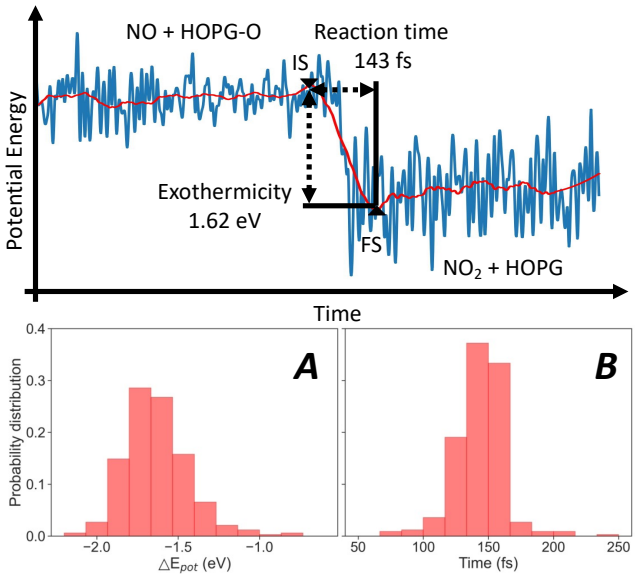

Figure 3.19: Top Panel: Time evolution of the total potential energy along a typical reactive trajectory. IS designates the initial state, marking the point when the potential energy of the system starts to decrease, while FS indicates the final state, where the potential energy of the system stabilizes and NO_2 is formed. The red line correspond to a Savitzky-Golay filter [220] applied to the potential energy. Panel A: Probability distribution of the energy released during the reaction. Panel B: Probability distribution of the reaction time computed as ($\Delta t = t_{FS} - t_{IS}$). Both distributions are normalized to all the reactive trajectories from the four initial kinetic energies (0.025, 0.05, 0.1, 0.3 eV) and the two investigated incidences (normal and oriented). The size of the bins are 0.13 eV and 16.6 fs respectively.

and the IS ($\Delta E = E_{FS} - E_{IS}$) across all reactive trajectories. This distribution spans from -2.2 eV to -0.8 eV, with a peak at -1.6 eV, suggesting the most probable potential energy difference between reactants and products. This average value is significant for understanding the energetics of the reaction and provides insight into the energy accessible to break the C-O bonds in the epoxy group and facilitate the removal of the nascent NO_2 molecule from the surface of HOPG under our simulation conditions. In a previous DFT study [70], the exothermicity of this process on a very similar system (graphene oxides) was estimated to be -2.38 eV. Although this value is larger than the one obtained in our AIMD simulations, these differences can be attributed to the fact that our calculation takes into account the surface temperature and the dynamic conditions of the reaction, which are not considered in NEB calculations. In summary, the reaction's

exothermicity, in conjunction with surface temperature, significantly aids the breaking of the C-O bonds within the epoxy group, leading to NO_2 formation even at relatively low NO incidence energies.

To understand the geometry requirements for the reaction to occur under our simulation conditions, we performed a statistical analysis of the configurations across all reactive trajectories at the moment identified as the IS. For this purpose, average values have been calculated for the following parameters: distance between the N atom and the O atom of the epoxy group (d_{N-O_e}), distance between the NO_{CM} and the surface of HOPG along the normal ($d_{NO_{cm}-HOPG}$), angle defined by the N atom, the O atom of the NO molecule and the O atom of the epoxy group (a_{O-N-O_e}), and the distance between the O atom and the carbons in the epoxy group ($d_{O_e-C_1}$, $d_{O_e-C_2}$). The corresponding distributions are shown in the panels A, B, C, D and E of Figure 3.20. It is noteworthy that the width of the distributions may be associated with a broad configurational space for the reaction to occur. Additionally, values deviating from the main distribution could arise due to inherent inaccuracies in the numerical method employed to determine the IS state. Focusing on the configuration characterized by the average values, as depicted in panel F of Figure 3.20, the NO_{CM} is located 3.49 Å from the O@HOPG surface, with the N atom facing the epoxy group at a distance of 2.34 Å. The formed O-N-O angle is 100.68° and the distance between the adsorbed oxygen and the carbon atoms are 1.51 Å and 1.52 Å respectively.

When we compare the average configuration obtained in our calculations for the IS with the one obtained by Hou et al. [70] using the NEB methodology for the TS of NO oxidation on graphene oxide, some interesting correspondences appear. In their reported values, the N atom is located at 2.17 Å (2.34 ± 0.24 Å in this work) from the adsorbed oxygen, which is at 3.58 Å (3.49 ± 0.42 Å in this work) ahead of the graphene surface. The distance between the carbon atoms bonded to the adsorbed oxygen and the oxygen atom itself are 1.66 Å (1.51 ± 0.12 Å in this work) and 1.47 Å (1.52 ± 0.12 Å in this work), respectively. These comparisons indicate that the IS averaged configuration is very close to the previously predicted TS [70], highlighting that the geometric conditions necessary for the reaction to occur are met. It confirms that our AIMD simulations follow the same reaction pathway. In summary, to efficiently form NO_2 from the NO oxidation reaction over room temperature graphitic surfaces, the NO molecules should be at a distance of ~ 3.50 Å from the surface and ~ 2.30 Å from the surface oxygen

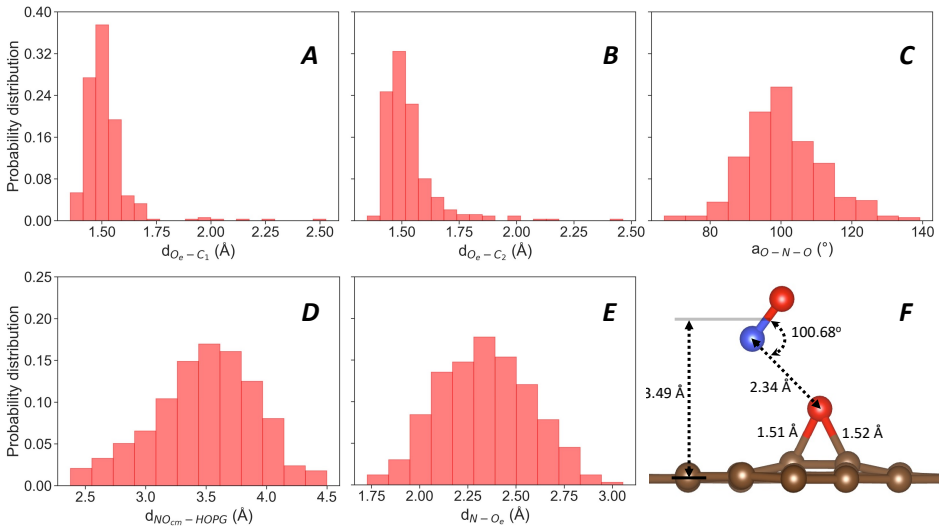

Figure 3.20: Probability distribution of the geometrical configuration regarding the NO and the epoxy group at the initial state (IS): distance between the oxygen atom of the epoxy group with respect to the first carbon atom (A) and the second carbon atom (B); angle defined by the O atom of NO, the N atom of NO, and the O atom of the epoxy group (C); distance between the center of mass of NO and the HOPG surface (D); distance between the N atom of NO and the O atom of the epoxy group (E); and a representative configuration at the average values of these geometrical parameters (F). All distributions are normalized to all the reactive trajectories from the four initial kinetic energies (0.025, 0.05, 0.1, 0.3 eV) and the two incidences (normal and oriented) of the NO. The size of the bins are 0.06 Å, 0.06 Å, 6.0°, 0.17 Å and 0.11 Å respectively.

atom, oriented with the N atom facing the epoxy group, and forming an O-N-O angle of $\sim 100°$.

Regarding the non-reactive events, an analysis of the probabilities reported in Table 3.4 shows that the NO adsorption probability decreases with increasing incident energy, from ~ 85 % at 0.025 eV to ~ 36 % at 0.3 eV, regardless of the initial orientation. This result indicates, as expected, that the increase in NO incidence energy makes the trapping of the NO molecule on the surface less probable and points out the minor role of the incidence conditions on the adsorption probability in non-reactive collisions.

3.5.2 Products angular distributions

Recognizing the complex interplay between the molecules and the surface, we conducted an examination of the scattering dynamics of NO and NO_2 molecules upon reflection from the thermally equilibrated O@HOPG surface. This analysis concentrated on trajectories featuring the highest investigated initial incidence energy of the NO molecule, set at 0.3 eV, for both normal and oriented incidences. This particular initial incidence energy was selected because it corresponds to the highest percentage of NO and NO_2 molecules desorbed from the surface in both directions of incidence, thus providing a more robust statistical data set. Considering the initial conditions used in this work, where the incidence angle is 0° for NO normal incidence and varies in a range between 0° and 18° for NO oriented incidence, we have evaluated the difference between the initial NO molecule incidence angle θ_i and the final product (NO or NO_2) scattering angle θ_s, computed as $\Delta\theta = \theta_s - \theta_i$. The probability distribution of this deviation angle is presented in Figure 3.21. Note that for normal incidence, this angular distribution corresponds directly to the NO scattering distribution since the initial angle θ_i is equal to 0.

For the non-reactive NO scattering trajectories, our results indicate some differences in the angular distributions obtained under the two different incidence conditions. For the oriented incidence, the distribution represented in the panel A of Figure 3.21 peaks at smaller angles, indicating a higher probability for specular reflection ($\theta_s = \theta_i$) than for the normal incidence conditions. On the other hand, the normal incidence angular distribution exhibits a more spread-out shape, indicating a higher probability of molecules with large scattering angles ($\theta_s >> \theta_i = 0$).

To understand the observed differences, we analyzed the point along the AIMD trajectory where the NO molecule deviates from the scattering plane, which is geometrically defined as the plane determined by three points: the position of the adsorbed oxygen atom, the center of mass of the NO molecule, and a third point obtained by projecting the center of mass of the NO molecule along the surface normal to the graphite. Figure 3.22 shows the time evolution of the fraction of molecules that remain within 5° of the initial scattering plane, alongside the average distance of the NO center of mass from the top-most layer of the O@HOPG surface. Since most scattering occurs within the first 1000 fs, the figure focuses on this interval, from which we can extract the relevant information. As can be observed, for both incidence orientations, during the initial approach phase, the

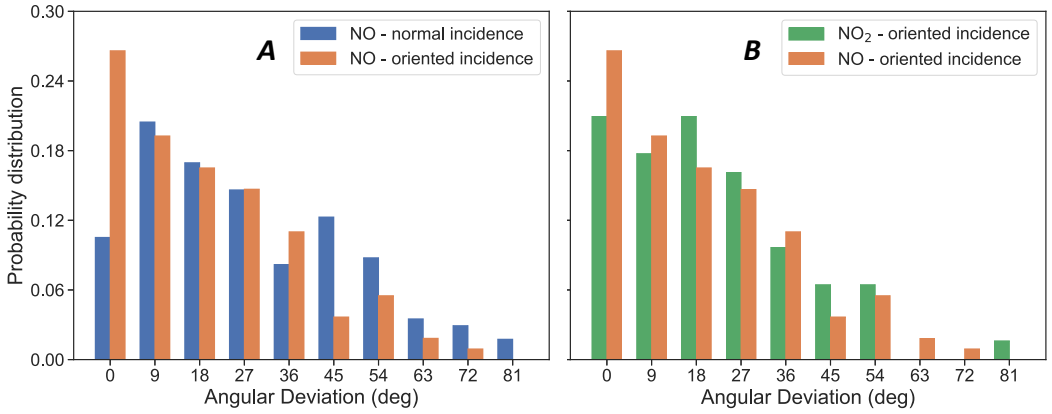

Figure 3.21: Comparative analysis of the distributions of the angular deviations $\Delta\theta$ after reflection from the oxidized graphite surface at an initial kinetic energy of 0.3 eV. The distributions are normalized to the total number of scattered molecules and the bin size is 9°. (A) for NO scattered molecules with initial normal and oriented incidences; (B) for desorbed NO and NO_2 molecules for initial NO with oriented incidence.

majority of molecules remain within the scattering plane, suggesting that the incoming NO molecules primarily follow straight paths dictated by their initial center of mass momentum before colliding with the O@HOPG surface. The figure also highlights that the deviation from the scattering plane occurs primarily at distances of around 4 Å between the NO center of mass and the surface's topmost layer. This region corresponds to where the molecule-surface interaction starts exhibiting significant corrugation. This distance is very close to the trajectory turning point at approximately 3 Å, indicating that inelastic scattering is mainly dominated by surface corrugation rather than earlier deviations caused by the interaction between the NO molecule and the epoxy group on the surface. Overall, the observed differences might be attributed to two main factors. First, in the oriented incidence cases, the direct interaction between the incident NO molecule and the epoxy group slightly favors elastic scattering, increasing the number of molecules scattered close to the specular reflection. Second, in normal incidence cases, the z-component of the molecule's velocity vector tends to undergo the most significant change during the collision with the thermalized surface, promoting inelastic scattering and leading to an increase in the number of molecules scattered at larger angles.

For reactive trajectories, where NO_2 molecules are desorbed after the oxidation of NO, the panel B of Figure 3.21 presents the $\Delta\theta$ distribution of NO_2 molecules compared

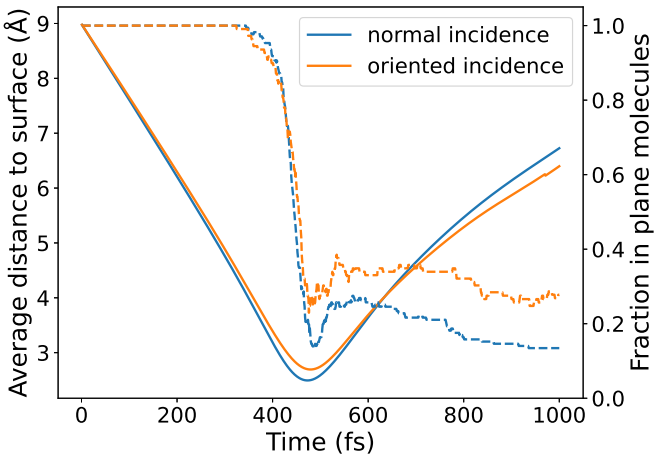

Figure 3.22: Time evolution of the average NO center-of-mass distance to the surface (solid lines, left Y axis) and fraction of molecules that remain within the scattering plane (dashed lines, right Y axis) during the first 1000 fs for oriented and normal incidence at a collision energy of 0.3 eV.

with the previously obtained $\Delta\theta$ distribution of NO molecules, but only for the oriented incidence case. Due to the limited number of NO_2 molecules produced (low reaction probability) under normal incidence conditions, there is insufficient statistical data to represent and compare the corresponding distributions. Our results show a similar shape for both distributions, indicating that the deviation angle of the scattered NO molecules or formed NO_2 molecules, in case of oriented incidence, follows the same behavior for reactive and non-reactive events. This observation could be ascribed to the fast NO oxidation process, indicated by the previously estimated reaction time of 143 fs, implying a direct reaction pathway. Consequently, the nascent NO_2 molecules pursue a trajectory similar to that of NO scattered molecules in the non-reactive scenarios.

3.5.3 Products energy distributions

Translational, vibrational and rotational energy distributions were calculated for NO scattered molecules at the highest incidence energy (0.3 eV) for both normal and oriented incidences. Additionally, the translational energy distribution was obtained for the NO_2 desorbed molecules at the same incidence energy, but only for the oriented incidence conditions.

The translational energy distributions of NO scattered molecules for both initial incidences are shown in Figure 3.23. As can be seen, both distributions have a similar bell shape, but slight differences can still be appreciated. The normal incidence distribution peaks at lower kinetic energy values (~ 0.09 eV), with a distribution mean of 0.12 eV, indicating an average translational energy loss of 60.0 %. On the other hand, for those molecules whose incidence was oriented, the distribution peaks at higher kinetic energy values (~ 0.15 eV), with a distribution mean of 0.14 eV, indicating an average translational energy loss of 53.3 %. These results are connected to the behavior previously observed in the angular distributions, where a slight preference for inelastic scattering is observed in normal incidence cases, suggesting greater energy loss during the molecule-surface collision. In normal incidence, all the translational energy is concentrated in the z-component of the molecule's velocity vector, which tends to undergo the most significant change during the collision, favoring energy loss. On the other hand, the smaller loss of translational energy in oriented incidence can be attributed to the direct interaction between the NO molecule and surface oxygen. This repulsive interaction may slightly reduce the energy lost during the collision.

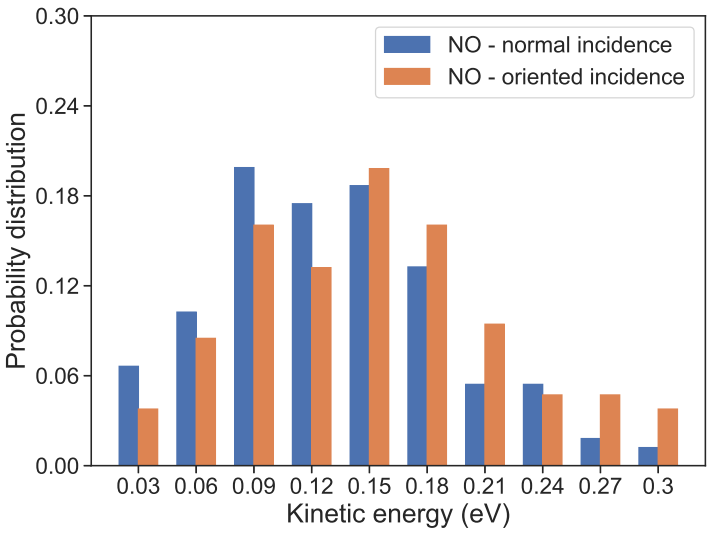

Figure 3.23: Translational energy distribution of NO molecules scattered with an initial kinetic energy of 0.3 eV at normal and oriented incidences. The distributions are normalized to the total number of scattered NO molecules and the bin size is 0.03 eV.

Furthermore, our results are also in good agreement with previous experimental observations on similar systems. Häger et al. [221] found that at low pyrographite surface temperatures (< 300 K), the NO scattered has a significantly smaller mean velocity, approximately half that of the incoming molecule. Additionally, in a recent study, Greenwood and Koehler [211] investigated the scattering of NO molecules in gold-supported room temperature graphene and observed a translational energy loss of the scattered NO molecule of approximately 80 %. In general, our AIMD simulations indicate that for a slightly different system, with the presence of an epoxy group on the graphite surface, there is also quite efficient kinetic energy transfer from the molecule to the surface at room temperature, resulting in a translational energy loss of the incoming molecule greater than 50 % after the collision.

Concerning the ro-vibrational state of the scattered NO molecules, our simulations show, in agreement with previous observations on similar systems [211, 221], that NO vibrational excitation occurs after the collision, with all detected molecules remaining in their ground vibrational state ($v = 0$), and that a small amount of energy is channeled into the rotational mode. Additionally, a small amount of energy is channeled into the rotational mode. Figure 3.24 shows the rotational energy distribution of the scattered NO with an initial kinetic energy of 0.3 eV at both incidence conditions. Once again, both distributions show a similar bell shape with slight differences in the maximum peaks. For normal incidence, the most probable rotational quantum number for the NO desorbed molecules is $j = 12$, while for oriented incidence, it is $j = 15$. These rotational quantum numbers represent rotational molecular energies of 0.03 eV and 0.05 eV respectively, indicating that only a small amount of the total energy initially available for the NO molecule (0.3 eV in translational energy + 0.1 eV in vibrational energy) was channeled to the molecule rotation after the collision with the surface. Indeed, as observed recently by Greenwood et al. [211] who studied NO scattering off graphene using surface-velocity map imaging, our simulations indicate that in NO scattering on an O@HOPG surface, most of the collision energy is transferred into the collective motion of the carbon atoms in the topmost layer of the graphite surface and is not channeled from translational to rotational modes of the molecules.

Finally, Figure 3.25 shows the translational energy distribution of the NO_2 desorbed molecules at the highest incidence energy of 0.3 eV for the oriented incidence conditions. As can be seen, the distribution presents a bimodal shape, with one peak located at 0.3

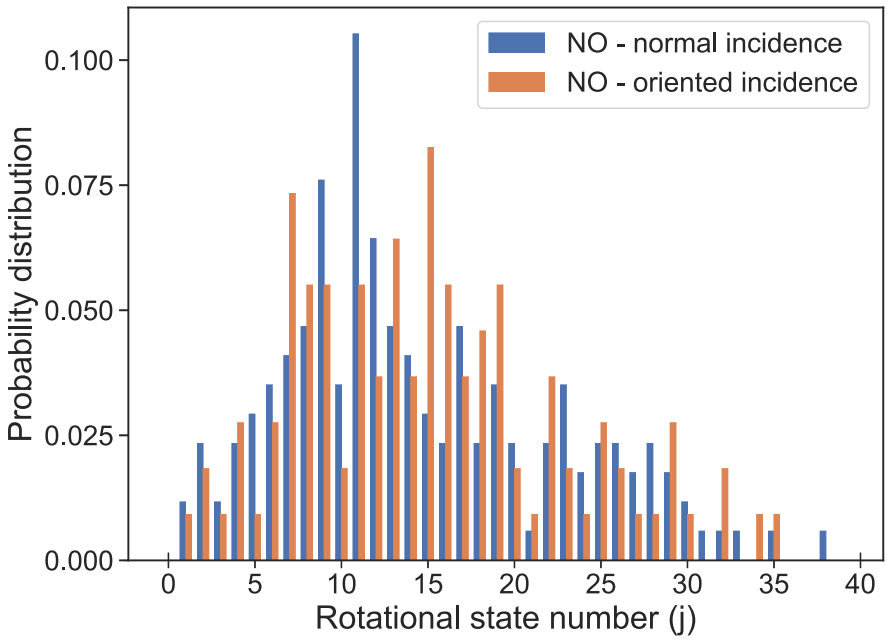

Figure 3.24: Rotational distribution of scattered NO molecules with 0.3 eV of initial kinetic energy at normal and oriented incidence. The distributions are normalized to the total number of scattered NO molecules and the bin size is 1.

eV, at the initial incidence energy, and another peak located at 0.5 eV, above the initial incidence energy. The mean value of the NO_2 kinetic energy is 0.36 eV, indicating, in contrast to previous results for scattered NO, a small increase in the translational energy of the NO_2 molecule of 20 % relative to the initial kinetic energy of the incoming NO (0.3 eV). This increase in the NO_2 molecule's translational energy is directly linked to the exothermicity of the reaction, which was estimated at approximately 1.62 eV under our molecular dynamics conditions.

The available energy resulting from the reactive process is first used to break the C-O bond of the epoxy group. Once the NO_2 molecule is formed, the desorbing molecules use part of this energy to escape from the HOPG surface. The remaining energy is then channeled into the translational and ro-vibrational modes of the nascent NO_2 molecule. Interestingly, the NO_2 molecules are mainly detected in their ground vibrational modes and are weakly rotationally excited, following the same behavior as the scattered NO molecules. In this regard, a complete analysis of the energetic characteristics of the

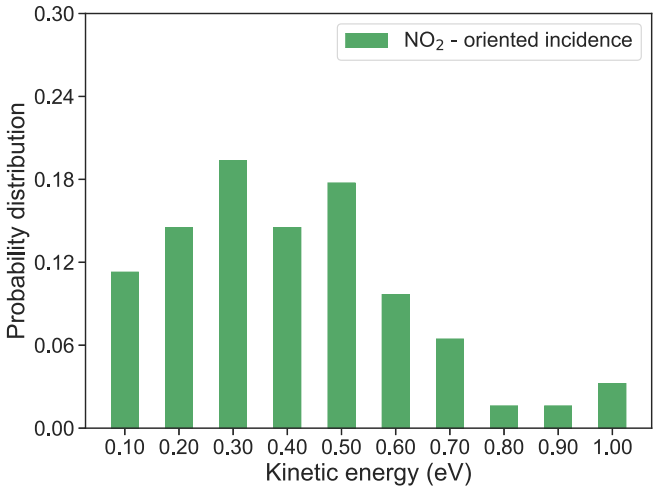

Figure 3.25: Translational energy distribution of NO_2 molecules scattered for an NO initial kinetic energy of 0.3 eV at oriented incidence. The distribution is normalized to the total number of scattered NO_2 molecules and the bin size is 0.1 eV.

NO_2 desorbed molecules, including the rotational and vibrational state populations, is reserved for future works.

3.6 Conclusion of the chapter

In summary, we performed AIMD simulations to study the oxidation of NO molecules on O@HOPG, thermalized at 300 K. A total of 2400 AIMD trajectories were analyzed, with 300 trajectories for each incidence energy (0.025 eV, 0.05 eV, 0.1 eV and 0.3 eV) at two different orientations of the incoming NO molecule (normal to the graphite and oriented to guarantee a collision with the oxygen atom of the epoxy group on the graphite surface). The reaction, adsorption, and scattering probabilities were computed. The angular and energy distributions of the products were also obtained under selected simulation conditions.

Our results reveal that NO_2 formation can occur even at the lowest collision energy investigated (0.025 eV), approximately equivalent to room temperature (300 K). This therefore highlights the impact of the dynamics on the NO oxidation process, since previous studies by Hou et al. [70] and Cen et al. [68] on graphene oxide had estimated this oxidation barrier to be around 0.1 eV, i.e. four times greater than our lowest collision

energy of 0.025 eV. Moreover, this result aligns qualitatively with previous experimental observations [60–62, 222, 223], which indicate a high likelihood of NO oxidation on oxidized graphite surfaces at room temperature. These findings underscore the efficiency of activated carbonaceous surfaces for facilitating NO conversion and NO_2 formation under mild conditions, highlighting their significance for environmental and catalytic applications.

Concerning the reaction mechanism, our exhaustive analysis shows that NO oxidation is most probable when the direct interaction between the NO molecule and the surface epoxy group is favored, suggesting that the reaction predominantly follows a direct Eley-Rideal mechanism. The key to the system's reactivity is the substantial amount of energy released once the reaction occurs; the exothermicity of the reaction under the simulation conditions was estimated to $\sim$ 1.62 eV. This energy is used to break the surface epoxy group bonds, and in some cases, is enough to overcome a physisorption barrier, allowing the nascent NO_2 molecule to escape from the surface. Furthermore, the statistical analysis performed on all the reactive trajectories, selecting configurations at specific reaction moments, allows us to describe the structural requirements for the reaction to occur under dynamical conditions. Remarkably, our results align well with the previous predictions by Hou et al. [70] and Cen et al. [68] regarding the transition state predicted for a similar system. Moreover, our analysis shows that once the energetic and configurational requirements are reached, the reaction occurs quickly ($\sim$ 143 fs) following a direct reaction mechanism.

We have also performed an analysis of the angular distributions of the NO scattered and NO_2 desorbed molecules at selected initial conditions. The results indicate slight differences between the scattered patterns of NO depending on the incidence conditions of the incoming molecule. Specular scattering is favored when the direct interaction with the O atom is forced (oriented incidence). This finding points out a possible feature that may provide clues in future experimental works to identify the interactions that occur between the incident molecule and the surface using scattering angular distributions, and potentially identify the presence of the O atom on the graphite surface. Furthermore, the similarity between the angular distributions of the NO scattered molecules and the NO_2 desorbed molecules confirms the direct nature of the NO oxidation reaction on O@HOPG surfaces.

Finally, the energetic characteristics of the NO scattered molecules obtained in our simulations confirm previous findings. There exists an efficient transfer of translational energy from the impinging NO molecule to the graphite surface, with more than 50 % of the initial NO molecule's energy being lost in the collision with the graphitic surface. The energy transfer is influenced by the presence of oxygen on the surface, with a slight decrease in energy loss observed when the interaction of the incoming NO molecule occurs directly with the epoxy group present on the surface. Furthermore, a weak rotational excitation is predicted in our simulations for the NO scattered molecules, in agreement with previous theoretical [224, 225] and experimental [210, 211] results on similar systems. Regarding the translational energy characteristics of the nascent NO_2 molecules, our findings indicate a slight increase in the available kinetic energy compared to the incoming NO molecule. In this regard, a complete analysis of the energetic characteristics of the NO_2 desorbed molecules, including the rotational and vibrational state populations, is reserved for incoming work.

Overall, our AIMD simulation results shed light on the NO oxidation mechanism on epoxy-functionalized carbon surfaces, identifying the conditions required for the reaction to occur and providing valuable insight into possible product signatures that can be targeted in future experimental studies to further characterize the reaction mechanism.

Despite the extensive number of AIMD trajectories performed, reactive events — particularly at low incidence energies — remained infrequent, even under optimized incidence conditions. Consequently, the number of successful reactive trajectories in the low-energy regime was too limited to construct statistically converged distributions for key observables such as energy partitioning and angular scattering. Capturing these distributions with sufficient statistical confidence would require a significantly larger number of trajectories, which becomes prohibitive due to the high computational cost of AIMD simulations.

In this work, all trajectories were computed on AMD EPYC nodes of the national high-performance computing facility *Très Grand Centre de Calcul du CEA* (TGCC), using a full node of 128 cores in parallel configuration. Under the employed simulation conditions, a 4000 fs trajectory required approximately 48 hours of wall time. For the 2400 AIMD trajectories used to characterize the NO + O@HOPG system, the total CPU time exceeded 4 million core-hours.

To overcome these limitations, a new methodology is required, one that preserves the accuracy of AIMD while enabling the generation of extensive trajectory datasets. Machine-learning potential energy surfaces (ML-PES), trained on *ab initio* energies and forces, offer a compelling solution to this challenge. Their development and application will be presented in the following chapters.

CHAPTER 4

Machine Learning-Based Potential Energy Surfaces for Studying Gas–Surface Interactions

Outline of the current chapter

4.1 From AIMD to ML-PES: Bridging Accuracy and Efficiency in Gas–Surface Molecular Dynamics **110**

4.2 ML-PESs for Gas– Surface Interactions **111**

This chapter introduces the theoretical foundations of machine learning (ML) for atomistic simulations, with a focus on building machine learning-based potential energy surfaces (ML-PES). These models are essential to this thesis, as they enable accurate and efficient simulations of reactive gas–surface systems. The chapter begins by explaining how data from AIMD simulations can be used to train ML-PES models, combining the accuracy of *ab initio* methods with the efficiency needed for large-scale molecular dynamics. It then describes the key components of atomistic ML models, including the choice of descriptors, neural network architectures, and loss functions. Special attention is given to approaches that learn both energies and forces from *ab initio* data. The chapter

concludes with the basic concepts needed to train and validate neural network potentials, setting the stages for their application in the simulations presented in the following chapter.

4.1 From AIMD to ML-PES: Bridging Accuracy and Efficiency in Gas–Surface Molecular Dynamics

As mentioned in the previous chapter, AIMD is computationally very expensive, as the electronic structure must be evaluated at every time step (on the order of femtoseconds or less), even when using DFT. This computational cost on the size of the system and on the time sampled by the simulations imposes strict limitations on the number of trajectories that can be simulated, thereby reducing the statistical reliability of key dynamical observables in gas-surface systems, such as adsorption and diffusion probabilities, angular and energy distributions, and energy transfer pathways.

To overcome these limitations, recent advances in ML methodologies[81, 82, 84, 85, 226], have enabled the development of high-dimensional ML-PES, allowing for efficient large-scale molecular dynamics simulations while maintaining high accuracy. These ML-PES models leverage ML to learn complex multidimensional energy and force landscapes from accurate electronic structure calculations. By significantly reducing computational costs, ML-PES enables the propagation of a much larger number of classical simulations of atom motion (trajectories), enhancing statistical accuracy while preserving the ab initio level precision. This approach effectively bridges the accuracy of AIMD with the efficiency of empirical force fields, offering a powerful alternative for gas-surface simulations[86, 87, 90, 91, 227, 228].

The development of a ML-PES is a complex process that requires training ML models on large datasets of electronic structure calculations, often performed at the DFT level for complex systems. In that context, leveraging AIMD-generated DFT data for ML model training provides a significant advantage. Unlike traditional potential energy surface construction methods, which rely on a predefined grid of points to explore the configuration space, AIMD naturally samples the most relevant configurations for the studied process. This integration of AIMD with ML-PES for molecular dynamics simulations represents

a cutting-edge methodology that optimizes both accuracy and computational efficiency, making it an indispensable tool for studying gas-surface interactions[80, 229].

In the specific case of this work, over 4.5 million core-hours of AIMD simulations yielded a dataset of more than 8 million DFT-evaluated configurations. This volume of high-quality data provides a robust foundation for training a ML-PES tailored to the NO + O@HOPG system, enabling accurate and efficient exploration of its reactive dynamics.

4.2 ML-PESs for Gas– Surface Interactions

In the last two decades, ML models have emerged as a powerful alternative for modeling PES, offering near *ab initio* accuracy at significantly reduced computational cost with a full inclusion of all nuclear degrees of freedom. Among the most widely adopted frameworks are kernel-based methods, such as Gaussian Approximation Potentials (GAP) [230] and symmetric Gradient Domain Machine Learning [231, 232], as well as neural network (NN)-based models [81, 82, 84, 233, 234]. Behler et al. [235] and Unke et al. [236] in their reviews provide a detailed comparisons of their theoretical foundations, performance characteristics and areas of applicability. Briefly, while kernel methods generally perform well in low-data regimes and offer better interpretability due to their simpler architecture, NN-based approaches tend to scale more favorably with system size and chemical complexity, making them particularly well suited for large and chemically diverse datasets.

From the perspective of gas–surface interactions, the earliest implementations of ML-PES for surface dynamics date back to the 1990s, when Blank et al. [237] employed simple feed-forward neural networks to model CO adsorption on Ni(111) and H_2 on Si(100). In this context, a major advance was the introduction of high-dimensional neural network potentials (HDNNP) by Behler and Parrinello [81], which proposed an atom-centered energy decomposition scheme. In this approach, the total energy of a system E_{tot} is expressed as a sum of atomic contributions, each determined by a neural network trained on the local atomic environment within a finite cutoff radius:

$$E_{\text{tot}} = \sum_{I=1}^{M} E_I(G_I) \tag{4.1}$$

where E_I denotes the contribution associated with atom I, M is the total number of atoms and G_I is a vector constructed to encode the local chemical environment of atom I within the predefined cutoff radius. This decomposition into local *atomic energies* offers two major advantages. First, it guarantees extensibility and scalability: the total energy of a larger system can be evaluated as the sum of atomic contributions without requiring retraining. Second, it reduces the complexity of the learning task by confining the energy prediction to the local environment of each atom, thereby rendering the high-dimensional PES more tractable.

This formulation has become the foundation of many modern ML-PES. However, it also introduces a new challenge: the construction of reliable structural descriptors that faithfully map local atomic environments into a form suitable for machine learning regression. These descriptors must encode sufficient structural and chemical information to enable smooth and accurate predictions of energies and forces, while strictly preserving the fundamental symmetries of the PES, namely invariance to global translations, rigid-body rotations, and permutations of atoms of the same species. This requirement is particularly demanding in gas-surface systems, where symmetry breaking at interfaces and the presence of chemically diverse species further complicate the design of generalizable and transferable representations.

Figure 4.1 illustrates the typical workflow followed in the construction of a ML-PES. The process begins with (1) data selection from an electronic structure calculations (e.g., configurations along AIMD trajectories), followed by (2) model construction through neural network hyperparameter selection, training, and testing. The trained model is then subject to (3) validation. If the validation criteria are met, the potential can be used for applications such as MD simulations. Otherwise, active learning is employed to expand the training dataset by sampling new, informative configurations, and the cycle is repeated until convergence. Each of these stages is described in detail in the following sections.

4.2.1 Representation of Atomic Configurations

A wide variety of descriptor strategies have been developed to meet the demands of modern ML-PES. These methods differ in their mathematical formulation, resolution, scalability, and flexibility, but all aim to efficiently transform local atomic environments

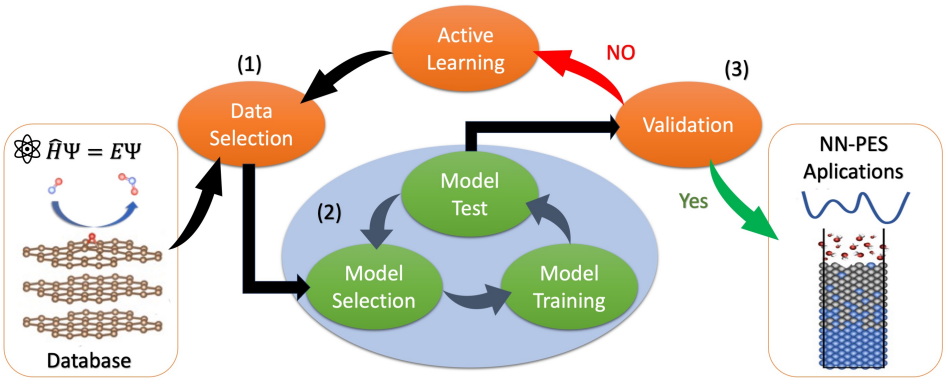

Figure 4.1: Schematic representation of the iterative workflow for constructing a ML-PES. (Adapted from Miksch et al. [238])

into feature vectors suitable for machine learning regression. In the following, we provide a brief overview of several principal approaches, emphasizing their construction principles and key advantages.

4.2.1.1 Atom-Centered Symmetry Functions (ACSFs)

Originally proposed by Behler and Parrinello in 2007 [81], atom-centered symmetry functions (ACSF) represent one of the earliest and most widely adopted descriptor families for the construction of ML-PES. These functions provide a compact and continuous representation of the local atomic environment while preserving essential physical invariances, including translation, rotation, and permutation of atoms of the same species.

In practice, the symmetry functions associated with a central atom I are typically decomposed into radial and angular contributions, denoted respectively as $G_I^{\text{rad}} = \{G_I^{\text{rad},t}\}$ and $G_I^{\text{ang}} = \{G_I^{\text{ang},s}\}$, where the index t and s labels the different symmetry functions introduced to capture the relevant features of the local atomic environment.

These functions capture two-body (radial) and three-body (angular) correlations between atom I and its neighbors J for the radial contribution and J and K for the angular one, within a finite cutoff radius R_c. The number of neighbors satisfying this criterion is denoted M_c. The general form of a radial symmetry function is given by:

$$G_I^{\text{rad}} = \sum_{\substack{I \neq J \\ R_{IJ} \leq R_c}}^{M_c} g^r(R_{IJ}) f_c(R_{IJ}) \tag{4.2}$$

where R_{IJ} is the Euclidean distance between atoms I and J, $g^r(R_{IJ})$ is a radial basis function (e.g., Gaussian or Chebyshev polynomial), and $f_c(R_{IJ})$ is a smooth cutoff function ensuring that interactions vanish continuously as $R_{IJ} \to R_c$. A commonly used cosine-based cutoff function is:

$$f_c(R_{IJ}) = \begin{cases} 0.5\left[\cos\left(\dfrac{\pi R_{IJ}}{R_c}\right) + 1\right], & R_{IJ} \leq R_c \\ 0, & R_{IJ} > R_c \end{cases} \tag{4.3}$$

The angular symmetry functions incorporate three-body correlations and are defined as:

$$G_I^{\text{ang}} = \sum_{\substack{I \neq J \neq K \\ R_{IJ}, R_{IK}, R_{JK} \leq R_c}}^{M_c} g^\theta(\theta_{IJK}) f_c(R_{IJ}) f_c(R_{IK}) f_c(R_{JK}) \tag{4.4}$$

Here, θ_{IJK} denotes the angle at atom I formed by the vectors $\vec{R}_{IJ} = \vec{R}_J - \vec{R}_I$ and $\vec{R}_{KI} = \vec{R}_K - \vec{R}_I$, and $g^\theta(\theta_{IJK})$ is an angular basis function. In the original formulation of Behler and Parrinello [81], the radial contributions were modeled using Gaussian functions centered at predefined distances, while the angular part involved products of Gaussian decays and powers of cosine angles to capture three-body correlations:

$$g^r(R_{IJ}) = e^{\left[-\eta(R_{IJ} - R_s)^2\right]} \tag{4.5}$$

$$g^\theta(\theta_{IJK}, R_{IJ}, R_{IK}, R_{JK}) = 2^{1-\zeta}\,(1 + \lambda \cos\theta_{IJK})^\zeta\, e^{\left[-\eta\left(R_{IJ}^2 + R_{IK}^2 + R_{JK}^2\right)\right]} \tag{4.6}$$

where η controls the radial decay around the central atom I (width of the Gaussian), R_s is the radial shift parameter, $\lambda \in \{-1, +1\}$ determines the angular orientation and ζ adjusts the angular resolution.

In a posterior study from Artrith et al. [82, 233], it was demonstrated that these symmetry functions can be interpreted as a basis set expansion of the local radial distribution function (RDF) and angular distribution function (ADF), respectively. Based

on this observation, they proposed replacing the original Behler–Parrinello functions with an orthonormal basis of Chebyshev polynomials, enabling a more compact representation. The Chebyshev polynomials of the first kind $T_\alpha(x)$, which form the basis of this expansion, are defined recursively as:

$$T_0(x) = 1 \tag{4.7}$$

$$T_1(x) = x \tag{4.8}$$

$$T_{\alpha+1}(x) = 2xT_\alpha(x) - T_{\alpha-1}(x) \tag{4.9}$$

and these polynomials are orthogonal over the interval $x \in [-1, 1]$ with respect to the weight function $w(x) = (1 - x^2)^{-1/2}$. To apply this basis in the context of local environments, interatomic distances and angles are first mapped to this domain via the linear transformation $x = 2R_{IJ}/R_c - 1$ for distances and $x = \theta_{IJK}/\pi$ for angles. The radial and angular symmetry functions are then expressed in terms of Chebyshev basis functions as:

$$G_I^{\text{rad}} = \sum_\alpha c_\alpha^{(2)}\, \phi_\alpha(R_{IJ}) \tag{4.10}$$

$$G_I^{\text{ang}} = \sum_\alpha c_\alpha^{(3)}\, \phi_\alpha(\theta_{IJK}) \tag{4.11}$$

where the radial basis functions $\phi_\alpha(R_{IJ})$ take the form:

$$\phi_\alpha(R_{IJ}) = \frac{k}{2\pi\sqrt{\frac{R_{IJ}}{R_c} - \left[\frac{R_{IJ}}{R_c}\right]^2}}\, T_\alpha\left(\frac{2R_{IJ}}{R_c} - 1\right) \tag{4.12}$$

and the corresponding orthogonal (dual) functions are defined as:

$$\bar{\phi}_\alpha(R_{IJ}) = T_\alpha\left(\frac{2R_{IJ}}{R_c} - 1\right) \tag{4.13}$$

with $k = 1$ for $\alpha > 0$ and $k = 2$ for $\alpha = 0$. The coefficients $c_\alpha^{(2)}$ and $c_\alpha^{(3)}$ are computed as weighted sums over the dual basis functions and cutoff-modulated distances or angles:

$$c_\alpha^{(2)} = \sum_{\substack{I \neq J \\ R_{IJ} \leq R_c}}^{M_c} \bar{\phi}_\alpha(R_{IJ}) f_c(R_{IJ}) \omega_{t_J} \tag{4.14}$$

$$c_\alpha^{(3)} = \sum_{\substack{J,K \neq I \\ R_{IJ}, R_{IK} \leq R_c}}^{M_c} \bar{\phi}_\alpha(\theta_{IJK}) f_c(R_{IJ}) f_c(R_{IK}) \omega_{t_J} \omega_{t_K} \tag{4.15}$$

where ω_{t_J} and ω_{t_K} are atom-type-dependent weights. This formulation allows a fixed descriptor dimensionality independent of the number of atomic species and is particularly advantageous for modeling chemically complex systems. A graphical representation of this methodology is shown in Figure 4.2, adapted from Miksch et al. [238]. The figure schematically illustrates the Chebyshev expansion of radial and angular distribution functions and the incorporation of species-dependent weights to encode chemical identity within the local atomic environment.

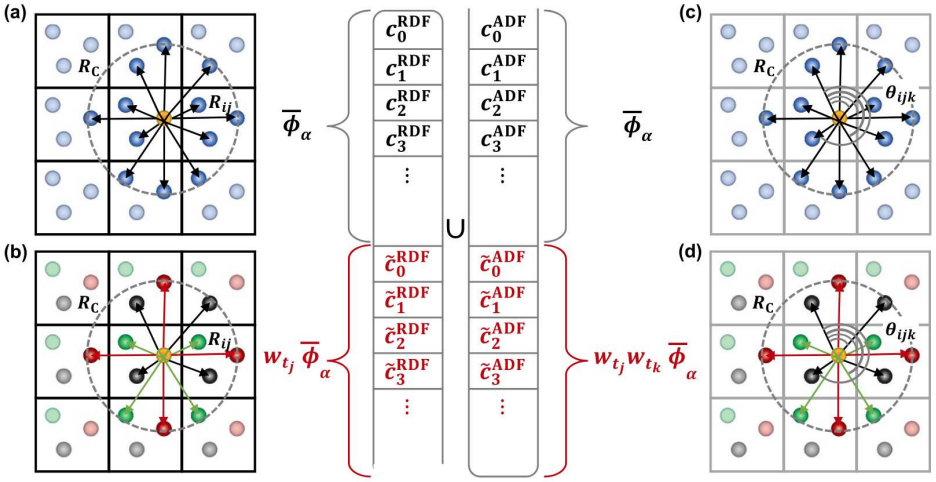

Figure 4.2: Schematic illustration of the Artrith–Urban–Ceder (AUC) descriptor as implemented in ænet [82]. The local atomic environment is encoded through a basis expansion of the radial (a) and angular (c) distribution functions using Chebyshev polynomials $\bar{\phi}_\alpha$. The corresponding expansion coefficients (b) and (d) are weighted by atom-type-specific factors ω_{t_j} and ω_{t_k}, which encode the local chemical composition. (Source Miksch et al. [238])

Another advantage from symmetry functions being defined analytically, is that forces acting on each atom can be computed directly via differentiation of the total energy with

respect to atomic coordinates. This results in an explicit expression for the atomic force on atom I, obtained by applying the chain rule:

$$\vec{F}_I = -\sum_{I=1}^{M}\sum_{u=1}^{\mu}\frac{\partial E_I}{\partial G_I^u}\frac{\partial G_I^u}{\partial \vec{R}_I} \tag{4.16}$$

where $\vec{F}_I$ is the force acting on atom I, evaluated as the negative gradient of the total energy with respect to its Cartesian coordinates $\vec{R}_I$. The summation is carried out over all M atoms in the system and over all μ components of the symmetry function vector $\vec{G}_I$ associated with atom I. Each component G_I^u is a differentiable function of atomic positions and encodes a specific radial or angular feature of the local environment, ensuring a smooth and continuous force predictions, allowing its incorporation into the training process in a rigorous and efficient manner [239].

4.2.1.2 Local Atomic Density-Based Descriptors

An alternative approach to describe the atomic environments based on the equations 4.1, is the construction of a local neighbor density function, which is then expanded in a suitable basis. Some examples include the bispectrum descriptor [230] and the Smooth Overlap of Atomic Positions (SOAP) [240]. In the bispectrum approach the local atomic environment is represented through a continuous neighbor density function $\rho_I(\vec{R})$, constructed by placing Dirac delta functions at the positions of all neighboring atoms J within a cutoff radius R_c, as well as at the position of the central atom I itself:

$$\rho_I(\vec{R}) = \delta(\vec{R}) + \sum_{R_{IJ}\leq R_c}^{M_c} f_c(R_{IJ}) w_J\, \delta(\vec{R} - \vec{R}_{IJ}) \tag{4.17}$$

Here, $f_c(R_{IJ})$ is a smooth cutoff function ensuring continuity as atoms move across the cutoff boundary, and w_J are dimensionless weights that encode the chemical identity of neighboring atoms, enabling the treatment of multicomponent systems. This formulation avoids the need for explicit parameter selection to define the spatial resolution of the descriptor, as required in approaches like ACSF. However, a major limitation is that δ-functions are sensitive to small perturbations in atomic positions, consequently even slight deviations of the positions between two environments can lead to substantial numerical differences in the resulting density.

In this context, the SOAP descriptor addresses this numerical instability by replacing the Dirac δ-functions with smooth Gaussians. Then, the neighbor density associated with a central atom I is expressed as:

$$\rho_I^{\mathrm{SOAP}}(\vec{R}) = \sum_{J=1}^{M_c} e^{\left[-\alpha\left|\vec{R}-\vec{R}_{IJ}\right|^2\right]} \tag{4.18}$$

where the summation extends over all N_{env} neighbors within a cutoff radius, and α determines the spatial width of the Gaussians. This formulation ensures a smooth and continuous representation of the local environment, improving numerical robustness. However, ρ^{SOAP} cannot be rotationally invariant, and when comparing two different neighbor densities ρ_i and ρ_i', one of the densities needs to be integrated over all possible rotations to incorporate this invariance:

$$k(\rho,\rho') = \int d\hat{R}\left|\int \rho(\vec{r})\,\rho'(\hat{R}\vec{r})\,d\vec{r}\right|^n \tag{4.19}$$

where $n = 2$ is a typical choice to preserve angular information while maintaining positive definiteness; also; a normalization factor ensures that the self-similarity $k(\rho,\rho') = 1$. This descriptor is widely adopted for its capability to capture subtle structural variations; however, this increased descriptive richness comes at the expense of higher computational cost compared to simpler alternatives such as ACSF.

4.2.1.3 Implicit atomic environment representations

A key limitation of explicit atomic environment functions (AEF), such as ACSF, SOAP, or bispectrum descriptors, lies in their limited scalability with respect to chemical complexity. As the number of atomic species increases, the combinatorial growth of basis functions makes it increasingly difficult to construct and optimize descriptors that remain both accurate and computationally efficient. To address this challenge, a new class of descriptors has emerged based on *implicit representations*, in which the structural encoding of atomic environments is learned dynamically during model training via deep neural networks.

Among the most successful frameworks are message-passing graph neural networks (GNN) [241], where atoms are represented as nodes and their pairwise interactions as edges within a cutoff radius R_c. In this formalism, each atom I is associated with a

feature vector $\vec{x}_I^t$, which is updated iteratively using information from its neighbors J. A general form of the update rule at interaction step t is given by:

$$\vec{x}_I^{t+1} = \mathcal{G}^t\left(\vec{x}_I^t + \sum_{\substack{J\neq I \\ R_{IJ}\leq R_c}} \mathcal{F}^t(\vec{x}_I^t, \vec{x}_J^t, \vec{g}(R_{IJ}))f_c(R_{IJ})\right) \tag{4.20}$$

Here, $\mathcal{F}^t$ is an interaction function that aggregates information from neighboring atoms based on their feature vectors and a radial basis embedding $\vec{g}(R_{IJ})$, while $\mathcal{G}^t$ is an atom-wise function used to refine the updated features. Both $\mathcal{F}^t$ and $\mathcal{G}^t$ are typically implemented as neural network layers. The cutoff function $f_c(R_{IJ})$, defined in Eq. (4.3), ensures smooth decay as interatomic distances approach R_c.

This iterative message-passing scheme allows each atomic representation to incorporate increasingly rich, non-local structural information, enabling the learned descriptors $\vec{x}_I$ to be sensitive to the full atomic environment. Crucially, because only interatomic distances are used and the summation is invariant to the order of neighbors, the resulting features are inherently invariant with respect to rotation, translation, and permutation of atoms—properties that are essential for physical consistency in atomistic simulations. Modern GNN-based models such as SchNet [242], NequIP [243], M3GNet [244, 245], and MACE [85] leverage these capabilities. Another prominent example of this strategy is the DeePMD-kit [84, 234], in which atomic environments are encoded as feature vectors generated by a neural network. These vectors are constructed from interatomic distances, angles, and chemical identities within a cutoff radius. Importantly, the embedding network itself is trained alongside the potential model, making the descriptor an implicit and differentiable function of atomic positions.

These implicit models have become increasingly popular due to their adaptability, chemical generalization, and potential for high accuracy. However, implicit descriptors also present notable limitations, one major challenge is the increased computational cost associated with training and evaluating deep neural networks, especially for large systems or when many layers are required to capture long-range correlations. This is particularly relevant in graph-based models, where message passing over multiple layers can lead to memory bottlenecks and reduced parallel scalability. Moreover, the abstract nature of learned embeddings reduces interpretability compared to explicit descriptors, where the physical meaning of each feature (e.g., bond distances, angular distributions)

is well understood. Additionally, because implicit representations are closely tied to the specifics of the training data and network architecture, they may suffer from poor transferability outside the domain in which they were trained [246].

In summary, the construction of meaningful atomic environment descriptors is a foundational step in the development of machine-learned interatomic potentials. A diverse range of strategies has been developed to encode local structural information in a way that respects the symmetries of the system, balances expressiveness and efficiency, and adapts to chemical complexity. From manually designed, physically interpretable functions like ACSF and bispectrum-based expansions, to more flexible and scalable approaches such as SOAP and Coulomb matrices, each representation offers trade-offs between resolution, generality, and computational cost. More recently, implicit representations based on neural networks (including DeePMD-kit style embeddings and message-passing GNN) have emerged as powerful alternatives capable of learning complex chemical environments without predefined basis sets. These developments have significantly expanded the applicability of ML-PES across chemically diverse and structurally complex systems. The effectiveness of each representation ultimately depends on how well it interfaces with the underlying machine learning model. In the next section, we examine the structure of atomistic neural networks, with a focus on how descriptors are integrated into modern potential energy models.

4.2.2 Structure and Training of ML-PES

Once the local atomic environments are encoded into descriptors, these serve as the input layer for a feedforward neural network (FFNN), as illustrated schematically in Figure 4.3, to determine the atomic contribution E_I to the total potential energy. Each atomic neural network receives as input a vector $\vec{G}_I = \{G_I^\mu\}$, containing the symmetry function components corresponding to atom I. These inputs are then propagated through one or more hidden layers comprising nonlinear activation functions and fully connected neurons.

The neurons in each hidden layer are connected to those in adjacent layers via trainable weights w_{ij}^{kl}, defining the connection strength between node i in layer k and node j in layer $l = k + 1$. In addition, each neuron receives an additive bias term b_j^l, and

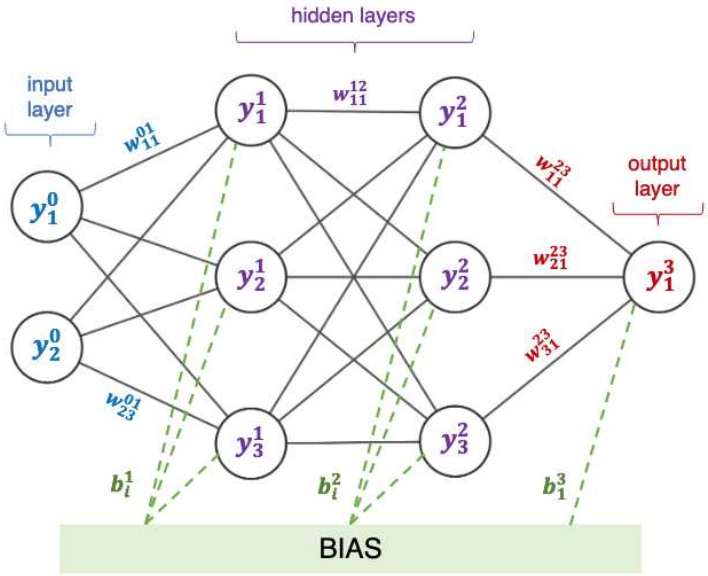

Figure 4.3: Schematic representation of a feedforward neural network with a 2–3–3–1 architecture, where the numbers indicate how many nodes are in each layer (i.e., two nodes in the input layer, two hidden layers with three nodes each, and a single output node). The input layer receives a vector of symmetry functions describing the local environment of atom I, and the output layer returns its atomic energy contribution E_I. Only a few weights $w_{i,j}^{k,l}$ are shown for clarity.

a nonlinear activation function f_j^l is applied to the weighted sum of inputs. The output of neuron y_j^l in layer l is computed as:

$$y_j^l = f_j^l \left(b_j^l + \sum_{i=1}^{N_{\text{nodes}}^k} w_{ij}^{kl} \cdot y_i^k \right) \tag{4.21}$$

where y_i^k is the value of neuron i in the previous layer. Consequently, the output E_I of the small feedforward neural network depicted in Figure 4.3 is expressed as:

$$E_I = y_1^3 = f_1^3 \left(b_1^3 + \sum_{m=1}^{3} w_{m1}^{23} \cdot f_m^2 \left(b_m^2 + \sum_{j=1}^{3} w_{jm}^{12} \cdot f_j^1 \left(b_j^1 + \sum_{i=1}^{2} w_{ij}^{01} \cdot G_i \right) \right) \right) \tag{4.22}$$

Activation functions represent a crucial component of neural networks, as they introduce the non-linearity required for the model to reproduce complex functional forms. This non-linearity enables feedforward neural networks to approximate arbitrarily complex functions, including potential energy surfaces. Several types of activation functions can be employed, such as the hyperbolic tangent, sigmoid, or Gaussian functions. All these functions share the characteristic of containing a non-linear region and being continuous and differentiable. This differentiability is essential for gradient-based optimization during network training, as well as for the accurate computation of atomic forces when the neural network potential is used in MD simulations.

For the output layer, a linear activation function is typically applied in order to ensure that the network can predict energies across the entire physically relevant range. Using a nonlinear activation (such as a sigmoid or hyperbolic tangent) in the output layer would constrain the possible output values to a limited interval (for example, between -1 and 1 for hyperbolic tangent), which would not be appropriate for representing the potentially wide variation in the atomic energies contributions. Alternatively, one could (re)scale the target energies to match the output range of the chosen activation function, but the standard practice is to use a linear output for energy prediction. The flexibility of a neural network is governed by its architecture, particularly the number of hidden layers and the number of nodes per layer. As the number of layers and nodes increases, the network's capacity to fit complex potential energy surfaces improves, albeit at the cost of a larger number of parameters.

The full model consists of a set of neural networks, one per atom, grouped by chemical element, which are trained concurrently on reference data obtained from electronic structure calculations, like for example, snapshots from AIMD, as indicated in Figure 4.4.

The training process consists of determining the optimal set of weights $\mathbf{W}_{\text{opt}}$ that minimizes the deviation between the predicted quantities—atomic energies and, optionally, forces—and the reference values obtained from *ab initio* calculations. This optimization is carried out over a training set $\{\Sigma_{\text{train}}\}$ containing Ω_{train} structures. A full pass over all training structures constitutes one *epoch*. The training objective is to minimize a total loss function $\mathcal{L}_F^{\text{tot}}$ with respect to the weight matrix $\mathbf{W}$:

$$\mathbf{W}_{\text{opt}} = \arg\min_{\mathbf{W}} \mathcal{L}_F^{\text{tot}}(\mathbf{W}; \Sigma_{\text{train}}) \tag{4.23}$$

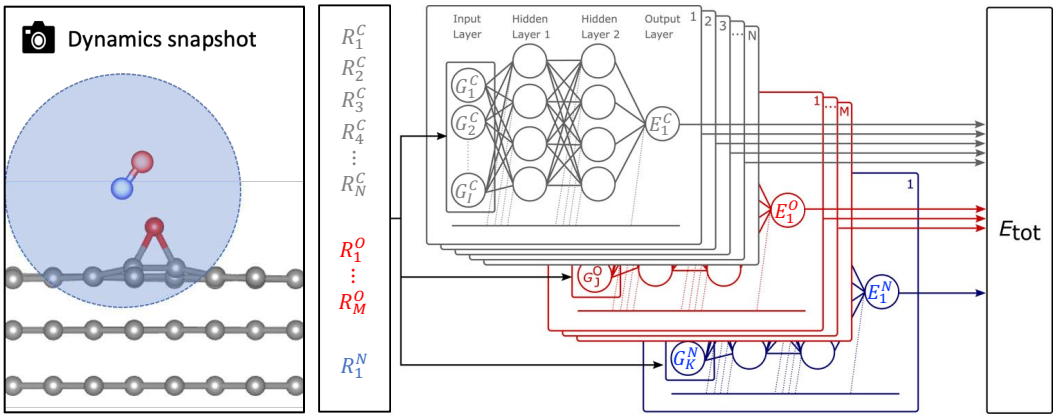

Figure 4.4: Schematic architecture of a ML-PES, each atom is represented by a vector of ACSF, which encode its local environment within a cutoff radius. These descriptors are used as input to atomic neural networks that predict atomic energies E_I^α for each atom $I = 1...N$ of type α. The total potential energy E_{tot} is obtained by summing the contributions from all atoms. Identical neural network architectures and symmetry function sets are employed for all atoms of a given element. (Adapted from Tokita et al. [247])

This total loss is commonly defined as a weighted sum of an energy loss $\mathcal{L}_E$ and a force loss $\mathcal{L}_F$:

$$\mathcal{L}_F^{\text{tot}}(\mathbf{W};\Sigma_{\text{train}}) = \alpha_E \mathcal{L}_E(\mathbf{W};\Sigma_{\text{train}}) + \alpha_F \mathcal{L}_F(\mathbf{W};\Sigma_{\text{train}}) \tag{4.24}$$

The weighting coefficients α_E and α_F are used to control the relative importance of energy and force accuracy during training. Their choice depends on the intended application of the potential, for example, to training exclusively on energies $\alpha_F = 0$, being this appropriate for cases where high force accuracy is not critical, or where forces do not vary strongly across the configurational space. Interestingly, accurate force predictions can sometimes still be obtained without explicitly including forces in the loss function [89], but this is highly dependent on the system and the training set must sufficiently samples the relevant configurational space [248].

Indeed, for applications involving molecular dynamics of gas-surface interactions involving reactions or any context requiring accurate energy gradients, it is often necessary to include forces in the loss function, either partially or fully across the training

set. The primary drawback of this approach is the increased computational cost, as force calculations require the evaluation of energy gradients with respect to atomic coordinates. Nonetheless, numerous studies have demonstrated that even incorporating force data for a subset of the training structures can significantly enhance the quality of the learned PES [242, 249].

In practice, the most common form of the loss terms is the root mean square error (RMSE), which has the advantage of being expressed in the same physical units as the target quantities. The energy loss is defined as:

$$\mathcal{L}_E(\mathbf{W};\Sigma_{\text{train}}) = \sqrt{\frac{1}{\Omega_{\text{train}}}\sum_{\omega=1}^{\Omega_{\text{train}}}\left[E_\omega^{\text{ref}}(\Sigma_{\text{train}}) - E_\omega^{\text{NN}}(\mathbf{W},\Sigma_{\text{train}})\right]^2} \tag{4.25}$$

and the force loss is computed over all Cartesian force components of all atoms in all training structures:

$$\mathcal{L}_F(\mathbf{W};\Sigma_{\text{train}}) = \sqrt{\frac{1}{3\sum_{\omega=1}^{\Omega_{\text{train}}} M_\omega}\sum_{\omega=1}^{\Omega_{\text{train}}}\sum_{A=1}^{M_\omega}\left\|\vec{F}_{A,\omega}^{\text{ref}}(\Sigma_{\text{train}}) - \vec{F}_{A,\omega}^{\text{NN}}(\mathbf{W},\Sigma_{\text{train}})\right\|^2} \tag{4.26}$$

where M_ω denotes the number of atoms in structure ω, and the norm $\|\cdot\|$ is the Euclidean norm over the force vector components.

The quality and reliability of the ML-PES must be assessed. For this purpose, errors in energies and forces with respect to the reference ab initio data are evaluated, as these metrics serve as key indicators of the model's accuracy and its suitability for dynamical simulations. The optimal errors are typically on the order of approximately 1 meV/atom for the energy and around 0.1 eV/Å for the forces, though these values depend strongly on the system and on the range of energies and forces present in the training set [247].

In practice, error convergence is monitored using a test set Σ_{test}, constructed from a subset of structures in the training set that are not employed by the neural network for training. This means that weights are optimized exclusively on the training set by minimizing the loss function with respect to that set only. At each iteration, using the ML-PES obtained during learning on the training set, the energies and/or forces of the structures in the test set are predicted. The errors obtained on this test set should be of the same order of magnitude as those of the training set, although slightly higher. More

precisely, monitoring the evolution of the RMSE on both the training and test sets over the epochs provides a means to evaluate the quality of the training process.

Three cases are typically identified when assessing the quality of training, as illustrated in Figure 4.5. The first case (a) corresponds to *underfitting*, which occurs when the model's performance is poor on both the training and test sets. This indicates that the model has not sufficiently learned the characteristics of the potential, often due to insufficient training (for example, too few epochs) and/or a neural network architecture with limited flexibility, such as too few hidden layers or neurons to represent the underlying complexity of the system.

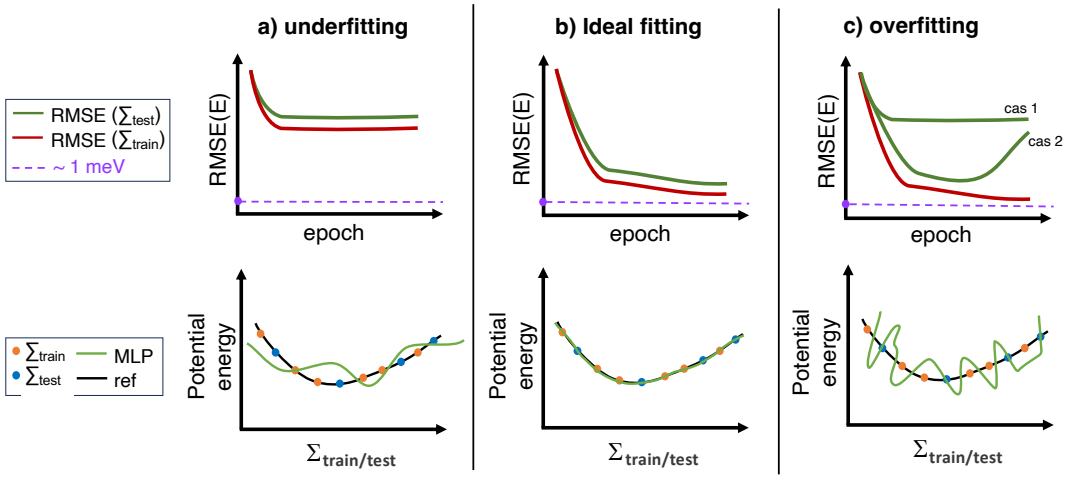

Figure 4.5: Schematic evolution of errors on the training and test sets in cases of underfitting (a), ideal fitting (b), and overfitting (c) during neural network training.

The second case (b) represents the scenario of *ideal fitting*. In this situation, the model's performance on both the training and test sets is high and comparable. This indicates that the model has successfully generalized from the training data and is capable of making accurate predictions on unseen data.

Finally, the third case (c) corresponds to *overfitting*, where the model achieves excellent performance on the training set but performs significantly worse on the test set. In this scenario, the model has essentially memorized the details of the training data, losing its ability to generalize to new configurations. Overfitting typically results from using too many epochs during training and/or an excessively complex model relative to the size of the training dataset.

4.2.3 ML-PES validation

Once the training of a ML-PES is completed and satisfactory convergence is achieved on both the training and testing sets, an essential next step is to evaluate the model's robustness through an independent validation set. The validation set comprises configurations that are entirely excluded from the training process but also from the test set, providing an unbiased assessment of the model's generalization capabilities. This evaluation is particularly critical in reactive systems, where subtle variations in atomic environments can challenge the model's predictive accuracy. By validating against a structurally diverse and physically relevant subset of the configurational space, it becomes possible to diagnose potential shortcomings in extrapolation behavior and to quantify how reliably the ML-PES reproduces key physical quantities such as energies and forces.

In gas-surface dynamics, even when the ML-PES demonstrated excellent accuracy on the validation set, it is important to recognize that MD simulations impose more stringent requirements on the quality of the potential. In contrast to isolated evaluations of energies and forces, MD relies on the continuous propagation of atomic trajectories across extended regions of the PES. Which means that even small local errors or discontinuities, imperceptible in static tests, can accumulate over time, potentially leading to nonphysical behaviors, trajectory instabilities, or incorrect predictions of reaction mechanisms and product distributions. This become particularly critical in chemically reactive systems such our system of study, NO oxidation on the O@HOPG surface, where dynamical processes are highly sensitive to subtle energetic features. For this reason, a more exhaustive validation step must be undertaken by directly comparing the ML-PES-driven dynamics against reference AIMD simulations, ensuring that the potential faithfully captures the underlying physics across the full range of accessible configurations.

4.2.4 ML-PES driven MD

To carry out MD simulations with the trained ML-PES, the interatomic forces are no longer computed from first-principles electronic structure calculations, but instead are predicted on-the-fly by the ML-PES. At each time step, the ML-PES provides the total energy and corresponding forces based on the current atomic configuration, enabling the propagation of the system according to classical mechanics. The nuclear equations of motion are integrated using algorithms like the velocity-Verlet, which ensures stable

and accurate trajectory evolution under the dynamical conditions established. Initial conditions, including the position and velocity of the atoms in the system are prepared analogously to the AIMD setup, ensuring consistency in the dynamical comparison.

Several MD packages support the integration of ML-PES, either natively or through dedicated interfaces. Among the most commonly employed tools, LAMMPS [250] offers broad compatibility with ML-PES through external libraries such as ÆNET [251], N2P2 [252], and NEQUIP [253]. Other options that enable flexible implementation of ML-PES within custom MD workflows are the TINKER software package [254, 255], SCHNETPACK [256] or frameworks like ASE [257] combined with TORCHANI [258]. The choice of software typically depends on the architecture of the trained model and the specific coupling interface available.

4.2.5 Active Learning

As previously discussed, reproducing the AIMD trajectories provides a particularly rigorous form of ML-PES validation, as this exhaustive test requires robust coverage of the configurational space explored during extended simulations. Consequently, the ML-PES may exhibit instabilities when applied beyond its training domain or in regions of configurational space that were insufficiently sampled in the original dataset. To overcome this limitation, the training set must be iteratively refined, as illustrated in the outer loop of Figure 4.1. Therefore, the simulations aimed at reproducing the AIMD results with the current ML-PES serve to identify configurations for which the model predictions are unreliable. These configurations are subsequently added to the training set, and the model is retrained to improve its generalization. This iterative cycle constitutes the foundation of active learning, a strategy that enables systematic and efficient enhancement of training datasets by targeting regions of highest uncertainty.

The primary challenge in this approach lies in identifying those configurations where the ML-PES is likely to fail, without resorting to exhaustive and computationally expensive recomputation using reference methods such as DFT. Active learning circumvents this bottleneck by selecting the most informative data points from a pool of unlabeled configurations, thereby minimizing the number of costly reference calculations. Active learning protocols [259–261] typically proceed in three iterative steps: (a) exploration of the configurational space via MD or enhanced sampling techniques, (b) selection

and labeling of critical configurations by computing their reference energies and forces, and (c) retraining the model using the enriched dataset. Two main paradigms are commonly employed: on-the-fly active learning [261–263], where retraining is performed concurrently with the simulation, and offline active learning, in which potentially informative configurations are first accumulated, followed by model retraining and subsequent re-evaluation.

A crucial component of active learning is the estimation of model uncertainty, which guides the selection of configurations to be added to the training set. Since reference energies and forces are unknown at this stage, various query strategies have been developed to estimate uncertainty without direct access to reference values. These strategies fall into three broad categories:

1. **Dataset reduction methods:** Candidate configurations are sampled and then filtered to eliminate redundancy. Only the most dissimilar configurations, relative to those already in the training set, are selected for labeling. For example, Bernstein et al. [264, 265] selected a subset of configurations from relaxation trajectories by (1) Leverage-score CUR algorithm, and (2) Flat histogram sampling (selection from low-density regions) with Boltzmann-probability bias.

2. **Query-by-committee approaches:** An ensemble of independently trained models evaluates a set of candidate configurations. The standard deviation in predicted energies (and possibly forces) across the ensemble serves as an indicator of model uncertainty [259, 266, 267]. These approaches can also be combined with dataset reduction techniques for enhanced efficiency [266].

3. **Statistical uncertainty quantification:** Statistical inference methods, such as those based on Bayesian frameworks [268], offer a principled way to estimate predictive uncertainty. For example, Gaussian Process Regression (GPR) models inherently provide uncertainty estimates that can be directly exploited during model training and configuration selection [261, 263]. In the case of ML-PES, approximate Bayesian inference can be achieved through dropout techniques applied at inference time, thereby embedding a query-by-committee strategy within a single network architecture [269]. Additionally, for ML-PES that depend linearly on their model parameters, such as Moment Tensor Potentials (MTP) [262], extrapolation

can be detected dynamically during simulations. This property can be leveraged to flag configurations lying outside the training domain and to guide the enrichment of the training dataset in an automated and data-driven manner.

In summary, while a well-tested ML-PES may appear accurate under standard validation procedures, its reliability in MD simulations must be confirmed through a combination of application testing and iterative refinement. In this context, active learning provides a systematic and scalable framework to identify and correct model deficiencies, thereby enhancing the accuracy and transferability of machine-learned potentials in complex reactive environments.

CHAPTER 5

NO Oxidation on O@HOPG: MD Simulations with a ML–PES

Outline of the current chapter

5.1 Building a ML-PES for NO oxydation over O@HOPG **132**

5.2 Validating the ML-PES with MD simulations mirroring AIMD initial conditions **144**

5.3 Active learning methodology **154**

5.4 Outlook on Large-Scale MD Applications **156**

5.5 Conclusions of the chapter **156**

This chapter provides an in-depth theoretical and computational study of the oxidation mechanism of NO on O@HOPG surfaces, investigated using a ML-PES. Building upon the methodological framework established in the previous chapter, we first describe the construction and validation of the ML-PES employed in this work. Particular attention is devoted to the generation of the training dataset from AIMD simulations, the selection of symmetry functions encoding the local atomic environments, and the

choice of neural network architecture. The resulting ML-PES is rigorously validated on an independent configuration set, with particular emphasis on its ability to accurately reproduce reference energies and forces relevant to the NO oxidation process. Subsequent sections are devoted to a detailed comparison of the dynamical behavior predicted by ML-PES MD simulations and the AIMD results, encompassing reaction probabilities, mechanistic insights, energy and angular distributions of the products, and the efficiency gains achieved through the use of the ML-PES. Finally, the chapter discusses remaining challenges, such as sources of overestimation or underestimation in certain observables from the ML-PES and highlights future directions for improving the methodology and extending its applicability.

5.1 Building a ML-PES for NO oxydation over O@HOPG

All machine learning potentials used in this work were obtained using the *ænet-PyTorch* software package [270], a graphics processing units (GPU) accelerated implementation of the atomic energy network (ænet) framework [82, 233, 239] designed for training ML-PES. This implementation, built upon the PyTorch machine learning library, offers significant computational efficiency, particularly for force-inclusive training, by leveraging parallel processing on modern GPU. The *ænet* package have been successfully implemented in the description of graphene [271] and the study of collisions of molecules with graphite, like the study from Rivero et al. [228], where they investigated specifically the scattering of oriented O_2 molecules on HOPG surface.

The reference structure data set used in this study to train the ML-PES was extracted from the AIMD simulations discussed in section 3 and the local atomic environments were encoded using Chebyshev polynomial-based symmetry functions, ensuring rotational, translational, and permutational invariance of the input descriptors. Separate neural networks were constructed for each chemical species present and the trained potentials were subsequently validated against independent DFT data, confirming their reliability in reproducing adsorption energetics and reaction pathways relevant to the NO oxidation process on O@HOPG surfaces.

5.1.1 Conformation of the training set

The accuracy and transferability of the ML-PES are fundamentally determined by the representativeness and diversity of the structures included in the training dataset. For the construction of the ML-PES used in this work, a comprehensive training set comprising 20 360 configurations was assembled. Of these, 20 000 configurations were selected from a pool of over 4.5 million snapshots generated by AIMD simulations starting from NO interacting with the O@HOPG surface described in section 3 and an additional 360 configurations were extracted from NO dissociation simulations at various heights (3.0 Å to 10 Å) above the adsorbed oxygen atom on the O@HOPG surface.

5.1.1.1 Selection criteria for the training set

In constructing the training dataset from AIMD simulations, particular attention was paid to accurately capturing the transition from the initial state, corresponding to NO + O@HOPG, to the final state of the oxidation, NO_2 + HOPG. Ensuring a balanced representation of configurations along this reaction coordinate is essential for the development of a ML-PES capable of reliably describing the dynamics and energetics of the NO oxidation process.

To capture the progression of the NO oxidation process at the structural level, representative regions along the reaction path are shown in Figure 5.1. Based on the typical evolution observed in the AIMD trajectories, three distinct stages can be identified: (i) a weak interaction region, where the NO molecule is still far from the surface; (ii) a strong interaction region, where the molecule approaches and interacts either with the surface or with the adsorbed oxygen atom; and (iii) a final region corresponding to the products, where the NO molecule is either adsorbed, scattered away, or forms NO_2 through reaction with the adsorbed oxygen. These structural distinctions serve as the basis for the segmentation scheme used in the construction of the training dataset and the analysis of dynamical behavior.

To implement this segmentation in practice, each AIMD trajectory was segmented into three regions, as depicted in Figure 5.2: weak interaction, reaction (or scattering), and products. For reactive trajectories, these regions were defined by the initial state time (t_{IS}) and the final state time (t_{FS}) of the reaction. The weak interaction region extends from the start of the dynamics to t_{IS}; the reaction region spans from t_{IS} to t_{FS}; and the

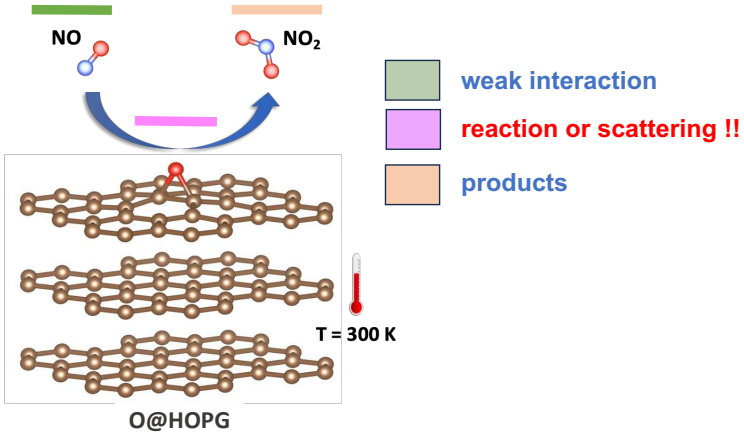

Figure 5.1: Illustration of the representative regions along the reaction pathway for NO oxidation on O@HOPG surface. Based on typical AIMD trajectories, three distinct stages are identified, (green) weak interaction, (purple) reaction or scattering, (orange) products.

products region covers the interval from t_{FS} to the end of the simulation, as shown in the left panel of Figure 5.2.

For non-reactive trajectories, a similar methodology was applied. In this case, the regions were delimited using the turning point time (t_{TP}) of the collision as a reference, with the regions defined as $t_{TP} \pm \Delta t$. The weak interaction region extends from the start of the dynamics to $t_{TP} - \Delta t$; the scattering region spans from $t_{TP} - \Delta t$ to $t_{TP} + \Delta t$; and the products region covers the interval from $t_{TP} + \Delta t$ to the end of the simulation, as shown in the right panel of Figure 5.2. For these trajectories, Δt was set to 35 fs, resulting in a total of 70 configurations within the scattering region.

To achieve a balanced and representative sampling of the relevant configurational space, configurations were randomly selected from the weak interaction, reaction (or scattering), and products regions in proportions of 20%, 50%, and 30%, respectively. This approach ensured that half of the configurations were drawn from the strong interaction region, where the energy and consequently forces show strong variations and a denser sampling is essential to accurately describe the reactive configuration space. The products region contributed 10% more configurations than the weak interaction region, as it encompasses both desorbed NO molecules, yielding configurations similar to those in

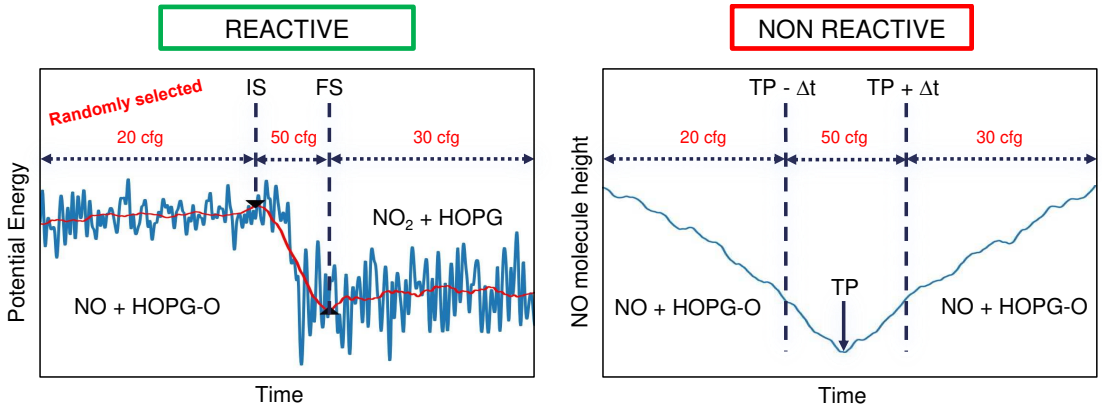

Figure 5.2: Schematic representation of the selection criteria for configurations along the AIMD trajectories. The left panel illustrates reactive trajectories, where configurations were sampled from three regions: weak interaction (20 configurations), reaction (50 configurations), and products (30 configurations), defined by the initial state time (t_{IS}) and final state time (t_{FS}). The right panel shows non-reactive trajectories, where configurations were selected based on the turning point time (t_{TP}) and a temporal window Δt, with 20 configurations from the weak interaction region (up to $t_{TP} - \Delta t$), 50 configurations from the scattering region ($t_{TP} - \Delta t$ to $t_{TP} + \Delta t$), and 30 configurations from the products region (after $t_{TP} + \Delta t$).

the weak interaction region, and adsorbed NO or NO_2 species, where the NO_2 molecule may be either desorbed or retained on the surface.

To ensure that the ML-PES correctly learns that the dissociation channel is inaccessible due to a prohibitively high energy barrier, additional structures were included in the training set. These configurations were extracted from potential energy profiles characterizing the dissociation of the NO molecule above the O@HOPG surface. These dissociation pathways were investigated within planes parallel to the O@HOPG surface, positioned at discrete heights ranging from 3 Å to 10 Å above the topmost atomic layer, with a vertical increment of 1 Å, including an additional intermediate plane at 3.5 Å. For each horizontal plane, the NO internuclear distance was varied systematically from 0.8 Å to 8.0 Å, employing a uniform increment of 0.06 Å. This rigorous sampling ensures comprehensive coverage of the relevant configurational space, essential for accurately describing the dissociation process of NO at varying proximities to the reactive surface, even though these specific configurations were not explored during the AIMD simulations.

In practice, 100 configurations per trajectory were sampled from a representative pool of AIMD trajectories at different incidence conditions and collision energies. From trajectories with oriented incidence, 70 reactive and non-reactive cases were selected, yielding 14 000 training configurations after sampling. From normal incidence conditions, 30 reactive and non-reactive trajectories were extracted, contributing an additional 6000 configurations. Finally, 120 configurations sampled along the NO dissociation coordinate—at heights of 3.0 Å, 3.5 Å, and 7.0 Å—were included to account for the NO dissociation process on the O@HOPG surface.

5.1.1.2 Assessment of training set quality

As previously mentioned, the quality of the training set is a critical factor for ensuring the accuracy, transferability, and reliability of ML-PES in simulating chemical processes. An essential requirement is the continuity and completeness of the configurational space spanned by the reference structures. Gaps or isolated groups of configurations may indicate disconnected regions along the reaction path, which could compromise the model's ability to describe transitions between states accurately.

The distribution of potential energies for the selected configurations provides essential insights into the representativeness and completeness of the training dataset. Figure 5.3 illustrates a typical potential energy distribution from the training set, demonstrating a continuous and well-sampled energy profile. This distribution smoothly encompasses configurations associated with reactants, intermediates, and products, as explicitly indicated by representative snapshots. The smooth transition across the energy spectrum confirms that the dataset sufficiently captures the configurational space necessary to accurately describe the entire reaction coordinate. This comprehensive coverage notably includes critical regions associated with bond-breaking and bond-forming events, essential for reliably modeling the reactive process.

A small number of configurations associated with NO dissociation on the O@HOPG surface lie at higher potential energies relative to the main distribution. These configurations correspond to dissociation pathways that are not accessible under the simulated conditions due to the prohibitive energetic barrier. Their inclusion serves to document these high-energy regions of the potential energy surface without introducing bias into the overall distribution, as they constitute a statistically negligible fraction of the dataset.

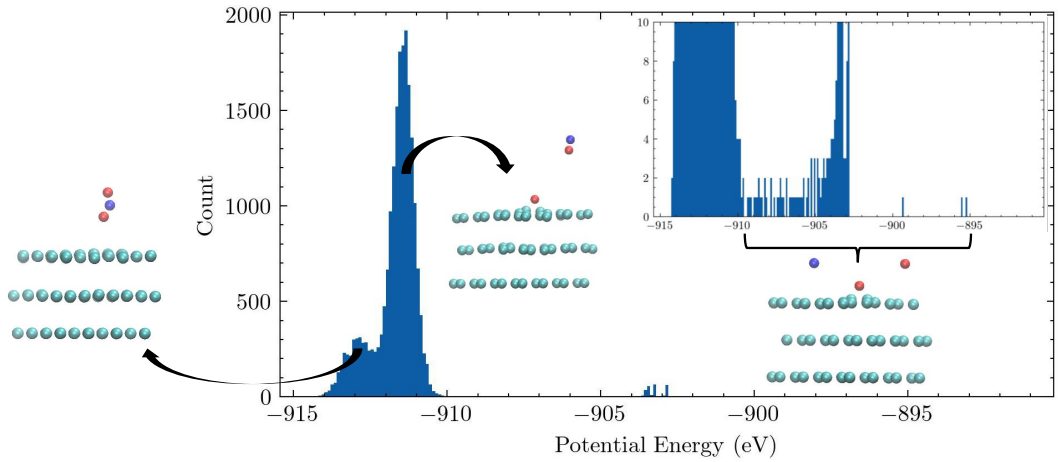

Figure 5.3: Potential energy distribution of the training set configurations, illustrating the continuous sampling of the reaction coordinate from reactants to products. Representative system geometries corresponding to different energy regions are shown alongside the histogram.

5.1.1.3 Fingerprint of the training set

Once the training data set is confirmed, the next step in the development of a ML-PES is to describe the local atomic environments in a way that is both precise and consistent across all configurations. As detailed in section 4.2.1, this description is provided by atomic fingerprints: numerical representations that capture the essential structural features of each atom's surroundings. These fingerprints are constructed to be invariant under rotations, translations, and permutations of identical atoms, ensuring that the machine learning model interprets each atomic environment uniquely and reliably.

In this work, atomic fingerprints were generated using a Chebyshev polynomial basis set, as implemented in *ænet* package, applied consistently to all species present: nitrogen (N), oxygen (O), and carbon (C). A cutoff radius of 6.0 Å was selected to define the spatial domain around each atom, chosen specifically because interactions between the NO molecule and the O@HOPG surface beyond this distance are negligible. Indeed, the NO desorption distance from the O@HOPG surface identified in our simulations is approximately 7.0 Å, which means that by adopting a cutoff radius of 6.0 Å, the fingerprints capture all significant local interactions while avoiding unnecessary inclusion of distant atoms that do not contribute meaningfully to the potential energy surface. The radial features were captured with an expansion order of 16 while the angular features

were described with an expansion order of 12. This descriptor setup offers a robust and consistent representation of atomic environments encountered along the NO oxidation reaction pathway, following criteria established in previous literature [247, 270]. A force inclusion rate of 10 %, following the best practices recommended for *ænet-PyTorch*, was used during fingerprint generation to incorporate force information into the training process

5.1.2 Training Procedure

As mention before, the ML-PES developed in this work were trained using the open-source code *ænet-PyTorch* [270], which is based on the Behler–Parrinello framework, as described in Chapter 4. This version of the original *ænet* code enables training on both, potential energies and atomic forces.

Including atomic forces in the training substantially increases the computational cost, both in terms of training time and memory requirements. For a training set consisting of X structures with M atoms each, training only on energies corresponds to fitting X target values, while including forces means fitting X energies and an additional $X \times 3M$ force components. This makes the training process considerably more demanding. However, as demonstrated in the study by López-Zorrilla et al. [270], it is not necessary to include force information for all structures to achieve an accurate potential. Instead, including forces for only a fraction of the configurations is sufficient.

The training loss function in *ænet-PyTorch* includes a weighting parameter α, which determines the relative importance of force errors in the total loss as showed in the Equation 4.24 of the chapter 4. In this work, α values between 0.1 and 0.3 were tested, in line with recommended practices, providing a good balance between force accuracy and energy precision, finally $\alpha = 0.1$ was chosen. It is important to note that when both energies and forces are included in the training, the total loss is dimensionless.

The neural network architecture followed standard designs for ML-PES, consisting of two (or, more rarely, three) hidden layers with 15 to 50 neurons per layer [82]. Most initial tests were performed using two hidden layers with 20 to 50 neurons each, where the activation function was the hyperbolic tangent. Training was carried out for up to 10 000 epochs, as in the tests reported in [270], to ensure full convergence of the network weights while avoiding overfitting. The training employed the AdamW optimizer from

the PyTorch library [272], with a learning rate of 10^{-4}, a batch size of 256 structures, and L2 regularization applied with $\lambda = 0.001$ to prevent overfitting [273]. After a benchmark the neural network architecture employed in this work consisted of two hidden layers, each containing 30 neurons.

The training process were preformed using the 90% of the configurations in the training set for training, and the remaining 10% were reserved for testing to monitor error convergence. The evolution of the total error during the training of the best ML-PES developed in this work, using the previously described setup, is shown in Figure 5.4. As shown, the training and testing errors both exhibit a consistent decreasing trend over the course of 10 000 epochs. The training error decreases smoothly, indicating steady optimization of the model parameters without abrupt fluctuations or signs of numerical instability. The testing error also follows a decreasing trend, albeit at a slightly higher magnitude than the training error, as expected. The absence of significant divergence between the two curves confirms that the model generalizes well and is not overfitting to the training data. The observed error plateau at later epochs suggests that the model has reached a convergence regime, where further training yields diminishing returns in predictive accuracy. The final errors with respect to the total energy are on the order of 0.6 meV/atom, which is well below the recommended accuracy threshold for ML-PES models, typically on the order of 1 eV/atom [247].

5.1.3 Validation of the ML-PES

Following the training procedure, the quality of the resulting ML-PES was assessed using an independent validation set. This set was constructed to ensure that it contained configurations not seen during training or testing, thereby providing an unbiased evaluation of the potential's predictive capability. A total of 2400 configurations were selected from 24 trajectories, with 100 configurations taken from each trajectory according to the same sampling criteria applied to the training set. For the oriented incidence case, the validation set included configurations from 9 reactive and 9 non-reactive trajectories, while for the normal incidence case, configurations were selected from 5 random trajectories. All these configurations were specifically chosen to be entirely distinct from those used in the training sets.

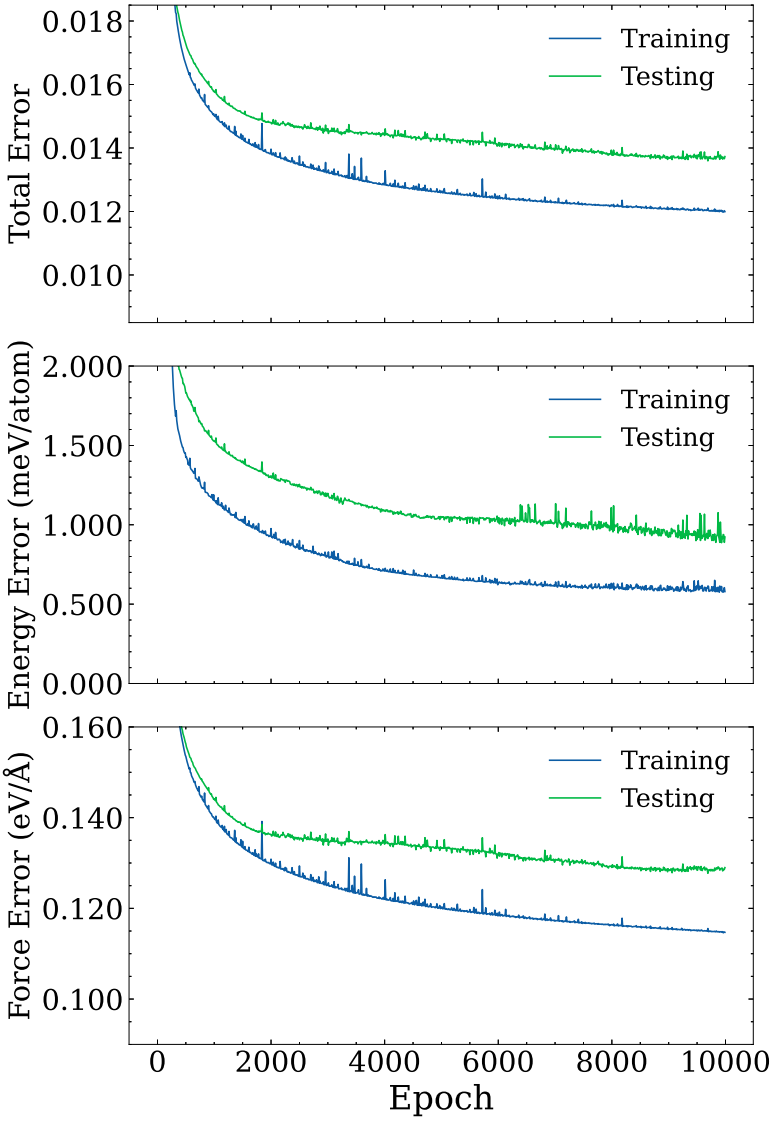

Figure 5.4: Evolution of the training and testing errors during the optimization of the ML-PES over 10 000 epochs. Panel A shows the total loss function, while Panels B and C present the separate contributions from the energy and force errors, respectively. In each case, the model achieves smooth convergence with testing errors remaining consistently low, indicating robust generalization to unseen data. Energy errors are expressed in meV per atom and force errors in eV/Å.

An essential aspect in the ML-PES validation step is ensuring that the validation set also adequately represents the configuration space explored during the reaction. In this sense, Figure 5.5 shows the distribution of DFT reference energies for all configurations in the validation set, compared against the distribution obtained from the full set of configurations generated in the AIMD simulations. The histogram shows that the validation set samples a broad range of energies, with a distribution shape that closely matches that of the full AIMD dataset. It covers both reactant-like and product-like configurations, as well as intermediate states, ensuring representativeness across the entire reaction coordinate. This confirms that the ML-PES is tested against a representative diversity of configurations spanning the full reaction coordinate.

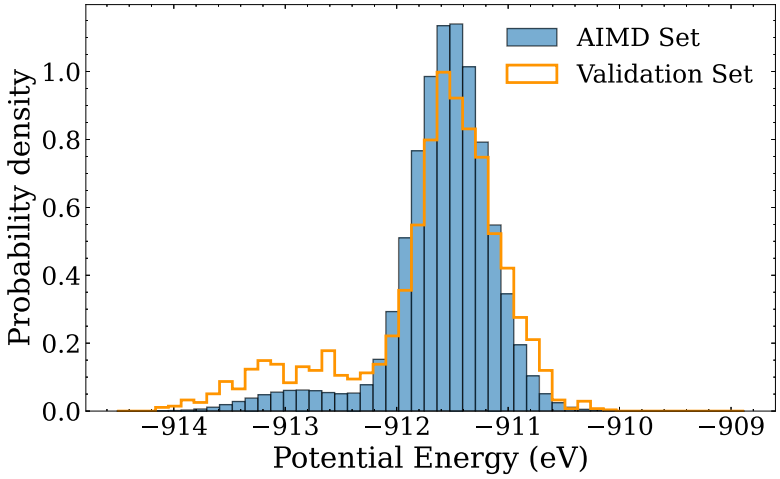

Figure 5.5: Distribution of DFT reference energies for the configurations in the validation set (orange line), compared to the energy distribution of the full AIMD dataset (blue bars).

Once the quality of the validation set is ensured, the performance of the ML-PES can be assessed by comparing its predictions with the DFT reference values. The comparison between the reference energies and the ML-PES predictions is carried out over all configurations in the validation set, evaluating both total energies and atomic forces. The comparison between the reference energies (E_{DFT}) and the energies predicted by the ML-PES ($E_{\mathrm{ML-PES}}$) is presented in Figure 5.6. The left panel displays a parity plot illustrating the correlation between predicted and reference energies across the full range of sampled configurations. The data points align closely along the diagonal, indicating excellent agreement between $E_{\mathrm{ML-PES}}$ and E_{DFT}. The right panel shows the distribution

of absolute errors ($E_{\mathrm{DFT}} - E_{\mathrm{ML-PES}}$), where these errors are mainly concentrated around zero. The statistical indicators shown on the parity plot confirm the high accuracy of the potential: the RMSE is 0.06 eV, the Mean Absolute Error (MAE) is 0.05 eV, and the coefficient of determination R^2 reaches 0.99. These error values correspond to less than 0.01% of the average absolute energy of the validation configurations, demonstrating that the potential accurately reproduces the DFT energies across the validation set.

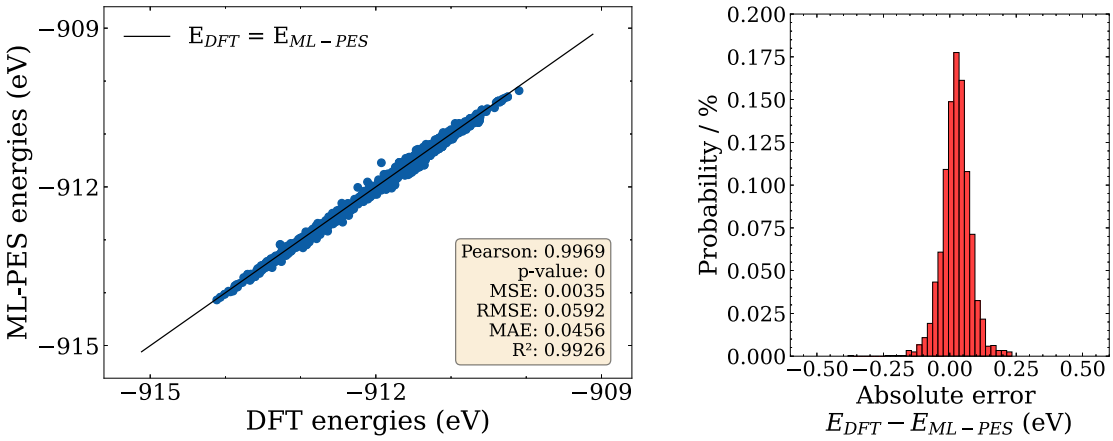

Figure 5.6: Comparison between DFT reference energies and neural network predicted energies for the validation set. The parity plot illustrates the strong correlation across the energy range, while the error histogram shows that the majority of configurations exhibit small deviations, with errors well below 0.1 eV.

The quality of the force predictions was assessed by analyzing both the magnitude and directional agreement between the forces predicted by the neural network potential ($\vec{F}_{\mathrm{ML-PES}}$) and those obtained from DFT calculations ($\vec{F}_{\mathrm{DFT}}$). Figure 5.7 presents two complementary views of this comparison. The left panel shows the joint distribution of the force magnitude ratio $\|\vec{F}_{\mathrm{DFT}}\|/\|\vec{F}_{\mathrm{ML-PES}}\|$ and the angle ϕ between the reference and predicted force vectors. This analysis highlights that approximately 85 % of configurations exhibit ratios higher than 0.85 and angular deviations smaller than 10.0°, indicating excellent correlation in both force magnitude and direction. The right panel shows the distribution of the absolute error in force magnitudes ($\|\vec{F}_{\mathrm{DFT}}\| - \|\vec{F}_{\mathrm{ML-PES}}\|$). The errors are narrowly centered around zero, confirming that the neural network potential accurately reproduces the forces across the validation set. For both energies and atomic forces, these

results are consistent with the accuracy reported in other ML-PES studies [90, 91, 227, 228] based on the same methodology.

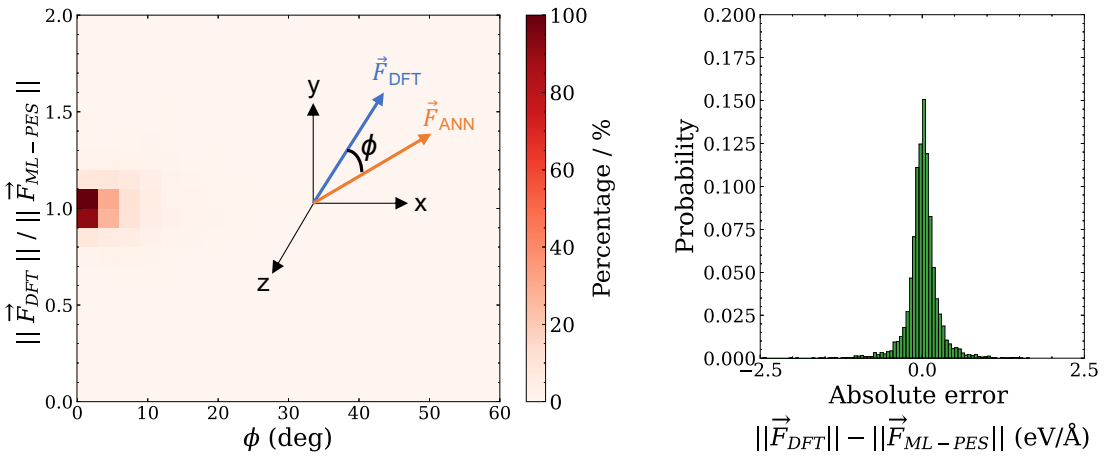

Figure 5.7: Force predictions on the independent validation set. The left panel shows the joint distribution of the force magnitude ratio $\|\vec{F}_{\mathrm{DFT}}\|/\|\vec{F}_{\mathrm{ML-PES}}\|$ and the angle ϕ between the predicted and reference force vectors, highlighting the strong correlation in both magnitude and direction. The right panel displays the distribution of absolute errors in force magnitudes ($\|\vec{F}_{\mathrm{DFT}}\| - \|\vec{F}_{\mathrm{ML-PES}}\|$), with the majority of configurations showing deviations close to zero.

To further assess the reliability of the ML-PES in describing reactive processes, we compared the dissociation profile of the NO molecule predicted by the neural network with that obtained from DFT calculations. Figure 5.8 presents this comparison for a dissociation event at a height of 3.5 Å above the O@HOPG surface. The predicted dissociation curve from the neural network potential reproduces the DFT reference profile with excellent agreement across the full N–O bond distance range, accurately capturing both the equilibrium bond region and the asymptotic dissociation limit.

The standard validation procedure based on the validation set demonstrates that the ML-PES is capable of predicting both energies and atomic forces with excellent accuracy. Moreover, the potential accurately describes key features of the NO dissociation process on the O@HOPG surface, capturing the relevant physics of bond breaking. While these results suggest that the ML-PES is sufficiently accurate for use in MD simulations, the complexity and sensitivity of reactive gas–surface dynamics warrant a more stringent

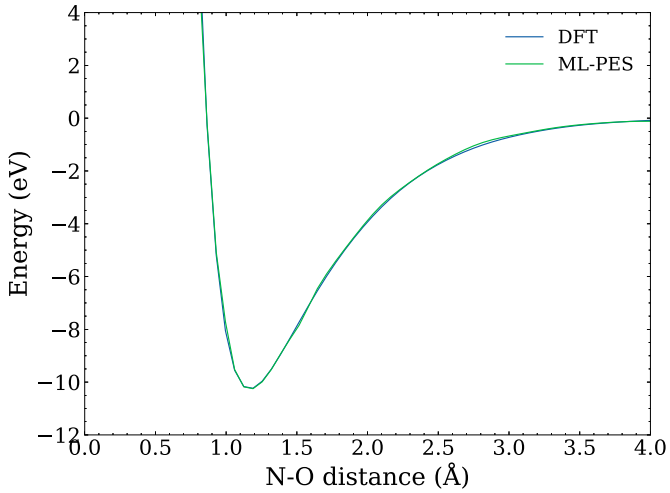

Figure 5.8: Comparison between the NO dissociation profile predicted by the neural network potential and the DFT reference at a height of 3.5 Å above the O@HOPG surface. The neural network accurately reproduces the DFT dissociation curve across the full range of N–O bond distances.

level of validation. For this reason, a higher-level assessment was performed, as described in the following section.

5.2 Validating the ML-PES with MD simulations mirroring AIMD initial conditions

As presented above, the ML-PES showed excellent results regarding the quality of the predictions on the validation set. However, MD simulations are far more sensitive to the quality of the potential, as they require not only accurate energies and forces at isolated configurations but also a consistent and smooth description of the potential energy surface across a wide range of atomic arrangements. Even small inaccuracies or discontinuities in the potential can accumulate over time, leading to significant deviations in trajectories, unphysical behaviors, or incorrect predictions of reaction mechanisms and product distributions. This sensitivity is particularly critical in complex gas-surface interactions, as is the case of NO oxidation on O@HOPG, which is why a more exhaustive validation was carried out in this work using exactly the same 2400 initial conditions employed in the AIMD simulations to perform MD simulations with the ML-PES. This approach

allows for a direct, one-to-one comparison between AIMD and ML-based trajectories under identical conditions, providing a robust benchmark of the potential's accuracy.

5.2.1 Computational Setup

In this context, MD simulations using the ML-PES were performed with the Large-scale Atomic/ Molecular Massively Parallel Simulator (LAMMPS) software package [250] interfaced with the *ænet* library, as described in detail by Chen *et al.* [251]. This setup enables efficient and accurate MD simulations with ML-PES by integrating the trained PES directly into LAMMPS. The *ænet*–LAMMPS interface takes advantage of LAMMPS's highly optimized neighbor list construction and parallelization capabilities, allowing the simulation of large systems with excellent computational performance. The ML-PES model provides energies and forces at each time step, which are used by LAMMPS to propagate the atomic trajectories using the velocity-Verlet integrator [274].

In order to set up the dynamics in LAMMPS, the initial conditions generated for VASP were adapted through an internally developed python code. This implementation enables the direct reuse of the same initial configuration of the NO + O@HOPG system employed in the AIMD simulations presented in Chapter 3. After, each MD trajectory was propagated for a total simulation time of 4 ps under micro canonical (NVE) conditions, using a time step of 0.5 fs. This choice ensures a level of energy conservation comparable to that of AIMD, with a standard deviation in total energy of the order of $\sigma \approx 0.03$ eV. In addition, in LAMMPS simulations, a termination condition was implemented in which dynamics was stopped if the center of mass of the NO_x molecule reached a height of 7.0 Å above the graphite surface. This criterion mirrors that employed in the AIMD simulations, ensuring consistency in the definition of desorption events across both methodologies.

In terms of computational performance, the execution time is reduced by several orders of magnitude relative to AIMD. For example, a 2 ps AIMD simulation with a time step of 1 fs for our 99-atom system typically requires approximately 3 days on 24 CPU cores. In contrast, the same trajectory using the ML-PES, propagated with a smaller time step of 0.5 fs and keeping all the degree of freedom of the system, can be completed in 15 minutes on the same number of cores. This means that, under equivalent computational conditions, the ML-PES simulations are approximately 800 times faster than AIMD

in the study of NO oxidation over O@HOPG. This remarkable acceleration not only underscores the efficiency of the ML-PES approach but also enables the extensive sampling of molecular trajectories, which is particularly advantageous for exploring reactive events at low collision energies. This computational efficiency makes it feasible to accumulate statistically meaningful observables across a large number of trajectories. However, beyond generating high-throughput data, it is also essential to critically assess the quality of the results obtained. To this end, a detailed comparison between the outcomes of ML-PES-driven simulations and those reported in Section 3.5 [153], with particular attention to reaction probabilities, reaction mechanism, angular and energy distribution of the products is presented.

5.2.2 Validation Results

5.2.2.1 Reaction probabilities and reaction mechanism

Starting from the methodology and results discussed in section 3.5.1, a comparison of reaction probabilities for NO collisions with the O@HOPG surface between ML-PES driven MD (dashed lines) and the AIMD simulations (solid lines) is presented in Figure 5.9. Normal incidence conditions are represented in blue while oriented incidence conditions are orange. Our results show that the ML-PES successfully reproduce the qualitative trends observed in AIMD: the reaction probability increases with collision energy for oriented incidence, while it remains nearly constant and low for normal incidence. These trends are consistent with the mechanistic interpretation provided by the AIMD data, where the orientation of the NO velocity vector enhancing the probability of direct encounters with the surface oxygen atoms plays a crucial role in the NO_2 formation.

In more details, Figure 5.9 show that for the higher collision energy studied (0.3 eV), the reaction probability is very well predicted by the ML-PES at bot incidence conditions, with a slightly higher overestimation at oriented incidence. However, at lower collision energies, the ML-PES systematically underestimates the overall reactivity compared to AIMD, being the underestimation at oriented incidence stronger. This behavior may be attributed to an overestimation of the C–O bond strength in the ML-PES, which would inhibit the departure of the oxygen atom from the surface and thus reduce the likelihood of NO_2 formation. Such overestimation may stem from sampling that is

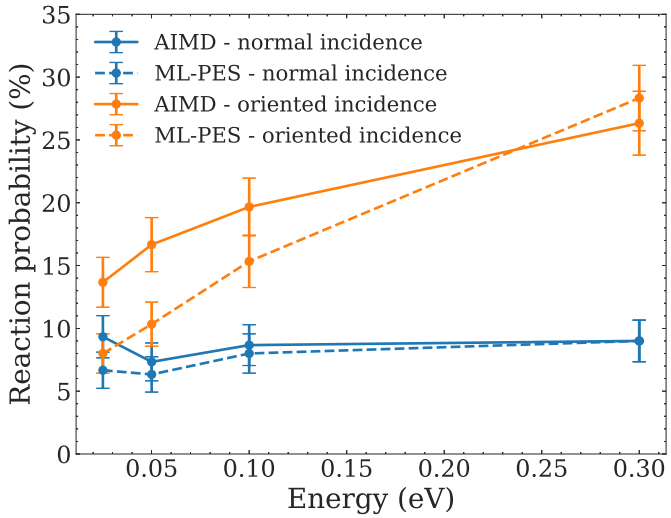

Figure 5.9: Comparison of reaction probabilities as a function of collision energy for NO impinging on the O@HOPG surface, obtained from AIMD simulations (solid lines) and ML-PES driven molecular dynamics (dashed lines). Results are shown for both normal (blue) and oriented (orange) incidence conditions. The probability errors were calculated by the usual formula for binomial distributions.

not fully sufficient to capture the complexity of this region of the configuration space, particularly near the point where epoxy bond breaking is initiated. Although this region is well represented in our training data, the results from MD simulations using the multidimensional ML-PES indicate that, for systems as complex as this, achieving complete and representative coverage of the configurational space explored by AIMD remains a major challenge. Since the training set is built from the sampled configurations, the quality and diversity of the sampling directly determine the ability of the ML-PES to generalize across the reaction landscape. This suggests that while the ML-PES accurately captures the global features of the reaction dynamics, further improvement may require targeted data enrichment to improve quantitative agreement and fully reproduce the statistical behavior observed in AIMD. These findings highlight the importance of combining complementary methodologies to enhance the robustness and completeness of the training set.

These results could also raise questions about the transferability of initial conditions from AIMD to ML-PES-based MD simulations. Although the same initial positions and velocities were used in the ML-PES simulations to ensure consistency in the comparative

validation, the slight mismatch in predicted reactivity suggests that, for applications beyond validation, it may be necessary to generate initial conditions—such as surface relaxation, thermalization, and NO dissociation—directly using the ML-PES itself. This approach would ensure full internal consistency between the PES used for MD propagation and the ensemble of initial configurations. In this work, however, the use of AIMD-derived initial conditions was intentional and appropriate for quality control purposes, enabling a strict one-to-one comparison with AIMD results in the validation phase.

To further evaluate the reliability of the ML-PES in capturing the dynamical features of the NO oxidation mechanism, we compare its predictions with AIMD reference data based on the methodology established in Section 3.5.1. In that section, the reaction pathway was characterized using two key observables: the change in potential energy during the reaction (ΔE_{pot}) and the temporal delay between the onset and completion of the reaction event. The initial state (IS) was defined as the point at which the potential energy begins to decrease sharply, indicating the onset of bond breaking, while the final state (FS) corresponds to the moment when the potential energy stabilizes after the reaction has occurred and the NO_2 product is formed. Figure 5.10 presents a direct comparison of the ML-PES MD simulations (black triangles) and the AIMD results (colored histograms) for these two observables. The left panel shows the distribution of ΔE_{pot}, representing the energy released between the IS and FS, while the right panel displays the distribution of the corresponding delay times, $\Delta t = t_{\text{FS}} - t_{\text{IS}}$. These comparisons allow for a detailed assessment of how accurately the ML-PES captures the mechanism of the reaction.

As shown in both panels of Figure 5.10, the ML-PES MD simulations reproduce the peak positions and general shapes of the distributions with good agreement compared to the AIMD reference data. While some quantitative deviations are observed, particularly a slight shift and broadening in the potential energy distribution, these may be explained by the lower number of reactive trajectories sampled in the ML-PES MD simulations, which could affect statistical convergence. Despite these minor discrepancies, the exothermicity predicted by the ML-PES is −1.78 eV, which is in good agreement with the AIMD value of −1.63 eV. Even closer agreement is obtained for the reaction time, where both approaches yield the same average duration of 143 fs. These results indicate that the ML-PES not only captures the correct energetic landscape of the reaction but also reproduce

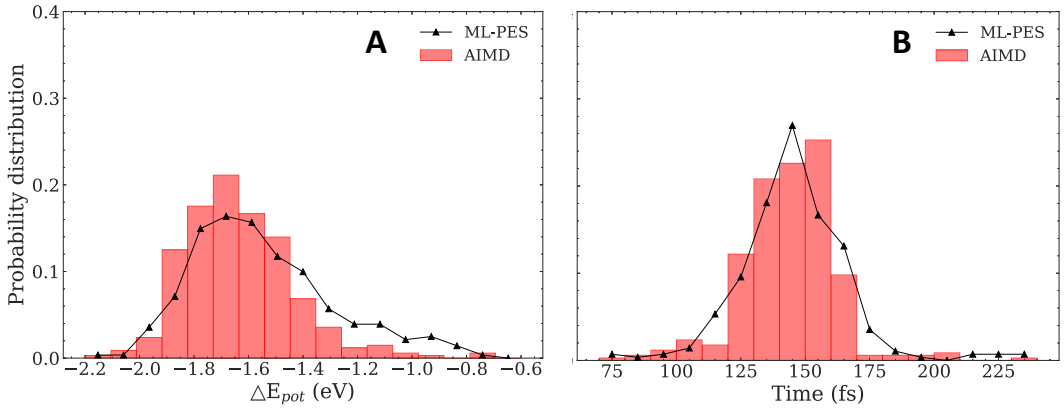

Figure 5.10: Comparison of reaction mechanism observables between AIMD (colored bars) and ML-PES (black triangles): (A) potential energy change, (B) delay times between IS and FS.

the fast and direct character of the NO oxidation mechanism observed in the AIMD simulations.

5.2.2.2 Angular and energetic distribution of scattered NO molecules

Another important aspect to assess in the evaluation of the ML-PES quality is its ability to accurately reproduce non-reactive trajectories. This can be addressed through the analysis of product observables such as angular deviation and translational energy distributions, which offer insight into how well the ML-PES captures the underlying mechanisms of momentum transfer and energy dissipation during scattering events. In this context, and following the methodology described in section 3.5.2, Figure 5.11 presents a comparison between the angular deviation distributions obtained from ML-PES MD simulations (black triangles for normal incidence and black circles for oriented incidence) and the corresponding AIMD reference results (blue and orange bars, respectively).

Our results show that the ML-PES MD simulations successfully reproduce the main features of the NO angular deviation distributions obtained from AIMD for both incidence conditions. In particular, the model captures the prominent peaks at low scattering angles, which corresponds to the dominant specular scattering regime of scattered NO molecules observed in the AIMD simulations. Furthermore, at the tail of the distribution, the ML-PES correctly reflects the overall decreasing trend observed in

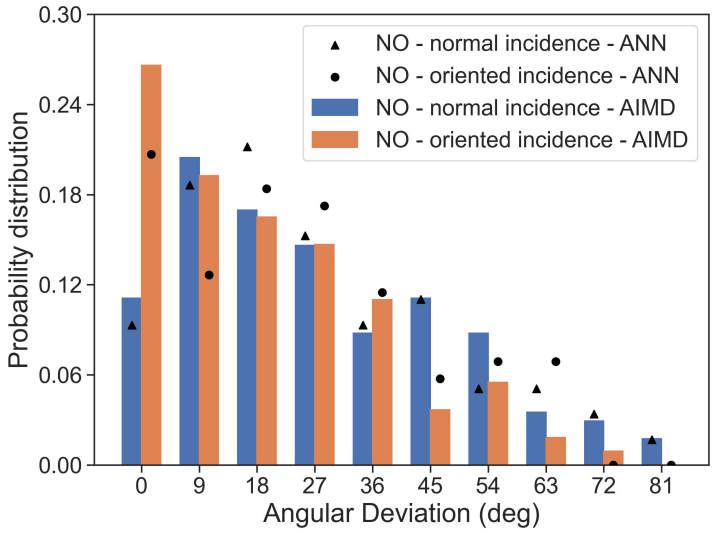

Figure 5.11: Scattering angle distribution of NO at an initial translational energy of 0.3 eV. The bars correspond to the reference AIMD results (blue for normal incidence, orange for oriented incidence), while black triangles and circles indicate the ML-PES predictions for normal and oriented incidence, respectively.

the AIMD results. Although, some quantitative discrepancies appear in the distribution, for the oriented incidence case, the ML-PES slightly underestimates the probability for very small deviation angles (0° and 9°), while overestimating the contributions in the range between 18° and 63°. For normal incidence, the same behavior is observed but with a less pronounced effect, with the more representative overestimation and underestimations at 18° and 54° respectively.

This results suggest that the ML-PES overestimate the angular deviation experienced by the NO molecules during non-reactive collisions with the O@HOPG surface. This effect is more pronounced under oriented incidence, where the distribution is broader and the probability density is redistributed toward intermediate scattering angles. From a physical standpoint, such deviations could be associated with an overestimation of the lateral momentum transfer during the collision. In particular, an exaggerated surface corrugation or an inaccurate description of short-range repulsive interactions may result in stronger deflection forces acting on the molecule as it approaches and interacts with surface atoms. In the context of the ML-PES, these discrepancies may arise from insufficient sampling of the configuration space associated with non-reactive interactions,

especially those involving close-range encounters with either the adsorbed oxygen atom or the carbon lattice of the functionalized surface, where the ML-PES may extrapolate these interactions and incorrectly predict the resulting angular momentum transfer. Moreover, it is worth noting that large angular deviations are statistically less frequent in the AIMD trajectories, which leads to a reduced sampling of these events during training. This scarcity of data may hinder the ML-PES's ability to generalize accurately in these regions and contribute to the minor discrepancies observed at the tails of the angular distributions.

Following the analysis of angular deviations, the translational energy profile of scattered NO molecules were examined to further evaluate the performance of the ML-PES in capturing energy redistribution during non-reactive scattering events. Figure 5.12 displays the probability distributions of the final translational energy for the NO molecules with an initial collision energy of 0.3 eV, comparing the ML-PES MD predictions (black triangles for normal incidence and black circles for oriented incidence) with the AIMD reference results (blue and orange bars, respectively) discussed in the section 3.5.3.

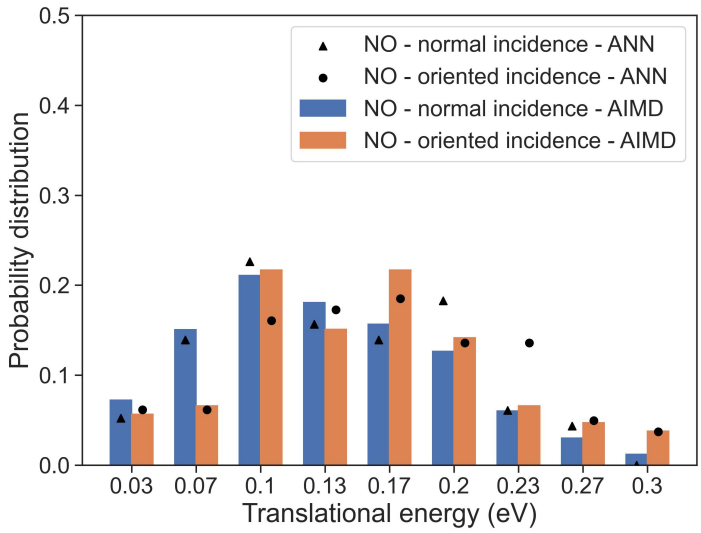

Figure 5.12: Translational energy distribution of scattered NO at an initial translational energy of 0.3 eV. The bars correspond to the reference AIMD results (blue for normal incidence, orange for oriented incidence), while black triangles and circles indicate the ML-PES predictions for normal and oriented incidence, respectively.

The results presented in Figure 5.12 demonstrate that the ML-PES accurately captures the main features of the translational energy distributions obtained from AIMD simulations. In particular, the overall shapes and peak positions of the distributions are well reproduced for both incidence conditions. For normal incidence, both methods exhibit a peak at 0.10 eV, while for oriented incidence the peak appears at 0.17 eV. From a quantitative standpoint, the ML-PES predicts slightly higher final translational energies compared to the AIMD results, as observed in normal incidence at 0.2 eV and oriented incidence at 0.23 eV, leading to a modest underestimation of the translational energy loss associated with the scattering process. For instance, at normal incidence, the ML-PES yields a mean translational energy of 0.13 eV, corresponding to a 55.4 % reduction relative to the initial energy, whereas the AIMD simulations report a slightly lower mean value of 0.12 eV, indicating a 58.9 % energy loss. Similarly, under oriented incidence conditions, the ML-PES predicts a mean energy of 0.16 eV (48.0 % loss), compared to 0.14 eV (52.5 % loss) obtained from AIMD.

Overall, these results support the ability of the ML-PES to reflect the distinct scattering dynamics associated with each incidence condition observed in the AIMD simulations. Under normal incidence, the alignment of the NO velocity vector with the surface normal enhances momentum transfer into the substrate, leading to more pronounced energy loss, while for oriented incidence, the presence of more oblique collision geometries and repulsive interactions with surface oxygen atoms results in a less pronounced translational energy loss. The minor discrepancies observed between the ML-PES and AIMD energy distributions, such as the slight overestimation of high-energy scattering, may be associated with the same factors discussed previously in the context of angular deviations. In particular, they could originate from insufficient sampling of the configurational space corresponding to close-range, non-reactive encounters between the NO molecule and surface atoms, including the adsorbed oxygen. In such regions, where the interaction forces are highly sensitive to subtle changes in geometry, the ML-PES may exhibit localized extrapolation behavior, leading to small deviations in the predicted momentum and energy transfer during the scattering process.

5.2.2.3 Assessment and Implications

It is important to recall that, in this validation stage, the number of ML-PES MD trajectories was intentionally kept equal to that of the AIMD simulations. As a result, no statistical enhancement was introduced at this point, and the comparison reflects a one-to-one reproduction framework rather than an exploitation of the ML-PES's computational efficiency. This reinforces that the results discussed here still belong to the validation phase, where the focus is on consistency and accuracy, rather than on statistical sampling improvements.

The validation of a ML-PES through the reproduction of dynamical observables represents one of the most stringent and comprehensive tests of its reliability. Unlike static benchmarks, such as force and energy errors on isolated configurations, MD simulations require the ML-PES to provide not only accurate predictions but also a smooth and consistent representation of the potential energy surface across extended regions of configuration space. In this context, even small inaccuracies in the prediction of a single configuration can accumulate over time, propagating through successive integration steps and ultimately leading to deviations in trajectory outcomes. This sensitivity makes dynamical reproduction a particularly demanding criterion and highlights the importance of identifying and correcting regions where the ML-PES may extrapolate or fail to generalize adequately.

In this context, the results obtained in the validation of the ML-PES constructed with the *ænet-pytorch* suit using the LAMMPS software as propagator, are in good agreement with the AIMD simulations. However, minor deviations have been identified in the ML-PES predictions across several observables, notably in reaction probabilities, scattering behavior, and energy redistribution. These findings indicates that, although the current ML-PES provides a solid foundation and captures key dynamical trends with good qualitative agreement, it may not yet be the optimal choice for large-scale MD applications. The ML-PES will be used to autonomously generate initial conditions—such as surface optimization, thermalization, and NO placement on the O@HOPG surface—and to propagate thousands of trajectories using tools like LAMMPS. However, for such a workflow to be fully reliable and efficient, the ML-PES should first pass the rigorous validation step imposed in this work with consistently accurate results.

Addressing this need for refinement leads naturally to the exploration of iterative strategies aimed at improving the training set and model performance. While the training set constructed from AIMD simulations proved effective in capturing key features of the NO + O@HOPG system, the discrepancies observed during validation highlight the difficulty of fully representing the configurational diversity in complex reactive environments. In this context, integrating active learning techniques becomes a promising and potentially essential strategy to iteratively enhance the training dataset.

5.3 Active learning methodology

In this thesis, the implementation of an active learning protocol is currently underway with the goal of iteratively enhancing the ML-PES by systematically addressing its limitations in regions of high uncertainty. The approach follows a data-driven refinement scheme in which short MD trajectories are propagated using the current ML-PES to sample both reactive and non-reactive configurations. A committee of neural network models, typically a set of models (e.g., 3) trained on the same data but with different initial weights, is used to estimate local uncertainty based on the variance in energy and force predictions. Configurations exhibiting intermediate levels of disagreement (e.g., force deviations between 0.1 and 0.8 eV/Å) are selected as candidates for enrichment, as they are most likely to improve the quality of the ML-PES upon retraining. In contrast, configurations associated with excessively large prediction errors are discarded if such high-error correspond to unphysical or highly distorted geometries. An example could be instances where the NO molecule penetrates unnaturally deep into the graphite surface, which is outside the relevant chemical space and could introduce numerical instabilities if included in the training set.

Although the AIMD trajectories naturally capture a broad range of reactive events, including adsorption, desorption, bond formation, and even complete reaction pathways, there remains the possibility that certain critical regions of the configuration space, such as rare scattering geometries or alternative mechanistic channels, are underrepresented due to the stochastic nature of trajectory sampling. To take this into account and ensure comprehensive coverage of all chemically relevant configurations, enhanced sampling techniques such as metadynamics [275] or umbrella sampling [276] will be implemented in future iterations of the active learning workflow. These methods employ carefully

chosen collective variables (CV), such as the distance between the nitrogen atom of NO and an epoxy oxygen (N–O_{epoxy}) or the O–N–O angle, to bias the exploration toward relevant reaction pathways and rare events. The configurations sampled in this way are subsequently evaluated by the neural network committee in the same manner as those from standard MD; those exhibiting substantial model uncertainty are selected as candidates for high-level reference calculations and subsequent inclusion in the training set.

Among all candidates configurations obtained from MD and enhanced sampling, a representative subset is selected to ensure decorrelation and structural diversity. To achieve this, dimensionality reduction techniques such as principal Component Analysis (PCA) [277] or Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP) [278] will be employed to project the descriptor space, followed by clustering algorithms to identify and select structurally distinct candidates. These subsampled configurations are then labeled with reference energies and forces at the same DFT level as the original training set and incorporated into the dataset. Using this enriched and decorrelated training set, a new generation of ML-PES models is subsequently trained, thus completing one iteration of the active learning workflow.

Convergence of the active learning protocol will be assessed using a combination of complementary criteria. First, the process is considered to have converged when the number of configurations identified by the neural network committee for labeling and addition to the training set falls below a user-defined threshold after an exploration phase, indicating that the model is sufficiently accurate in the sampled regions. Additionally, convergence will be monitored through the systematic reduction of prediction uncertainty across new MD trajectories, improvements in validation metrics such as the force RMSE, and the narrowing of the gap between training and test errors. This multifaceted approach ensures that the resulting ML-PES is robust and transferable, providing an accurate description of reactive dynamics and energy dissipation mechanisms.

This active learning protocol is currently under implementation, and its application to the NO + O@HOPG system will form the basis of future work. While the specific performance improvements of the refined ML-PES remain to be quantified, the methodology outlined here provides a rigorous and automated route to systematically enhance the model's transferability and accuracy in dynamically relevant regimes.

5.4 Outlook on Large-Scale MD Applications

Once the refinement of the ML-PES through active learning and improved sampling protocols is complete, and the model reaches the level of accuracy and robustness required, it will be ready for reliable and efficient large-scale MD simulations using the LAMMPS package. In this final phase, initial conditions—such as surface optimization, thermalization, and NO placement—will be generated autonomously within the ML-PES framework, allowing for the propagation of tens of thousands of trajectories under varied incidence conditions. This workflow, designed to fully exploit the computational efficiency of ML-PES, will enable the extraction of statistically converged observables that remain inaccessible to AIMD simulations due to their high computational cost. The infrastructure developed in this thesis provides a solid foundation for this transition, which will support the large-scale application of ML-driven dynamics to the specific case of NO oxidation on oxygen-functionalized HOPG. This will allow for a detailed statistical characterization of the reaction mechanism, energy transfer pathways, and scattering dynamics, offering new insights into this prototypical gas–surface reactive system.

5.5 Conclusions of the chapter

In this chapter, the construction of a ML-PES for the NO oxidation reaction on an O@HOPG surface was presented, using the *ænet-PyTorch* framework. The training dataset was carefully selected from AIMD trajectories, including representative configurations spanning the full dynamical pathway, along with explicit sampling of the high-energy NO dissociation channel. Atomic environments were encoded using Chebyshev polynomial-based descriptors, and a force-inclusive training strategy was employed to ensure accuracy in both energy and force predictions. The final ML architecture, consisting of two hidden layers with 30 neurons each, achieved strong performance on the validation sets, with energy errors below 0.1 eV and force predictions exhibiting tight angular and magnitude agreement with DFT.

When deployed in MD simulations, the ML-PES demonstrated excellent computational efficiency, enabling trajectory propagation up to 800 times faster than AIMD under comparable conditions. This acceleration made it possible to perform extensive simula-

tions and obtain statistically meaningful sampling of scattering and reactive outcomes, including low-probability events. A detailed comparison with AIMD results confirmed that the ML-PES reproduces well the main features of the NO + O@HOPG reaction system. The model captures the correct qualitative trends in reactivity and product branching, including the increase in NO_2 formation with collision energy at oriented incidence and its near-constant behavior under normal incidence. It also provides a consistent description of the reaction mechanism, closely matching the exothermicity and reaction times observed in AIMD, and for non-reactive events, the ML-PES reproduces the angular and translational energy distributions of scattered NO molecules with reasonable agreement. These results demonstrate the good predictive capability of the model and highlight its potential as a reliable and efficient tool for simulating complex gas–surface interactions beyond the reach of conventional AIMD approaches.

While the results presented in this chapter are highly encouraging, some aspects of the ML-PES can still be improved to achieve closer quantitative agreement with AIMD. In particular, the current model slightly underestimates the reaction probability and tends to overpredict NO adsorption. To address these discrepancies, future work will explore the incorporation of additional configurations into the training set via active learning strategies, with the goal of explicitly sampling regions of the configurational space where extrapolation may be occurring. Nevertheless, the present results are highly promising, and we remain confident in the potential to further refine the ML-PES to achieve excellent agreement with AIMD simulations.

Conclusions and perspectives

This doctoral thesis aimed to advance the atomic-scale theoretical understanding of NO oxidation on oxygen-functionalized carbon-based materials by addressing several key scientific challenges. In particular, it sought to elucidate the influence of surface thermalization on reaction probabilities, to characterize the atomistic mechanism underlying the oxidation process, and to investigate energy transfer pathways during gas–surface collisions. To meet these objectives, a combined theoretical framework based on ab initio molecular dynamics (AIMD) and machine learning potential energy surfaces (ML-PES) was employed.

Using the AIMD approach, simulation of the NO oxidation on the O@HOPG surface thermalized at room temperature, for a wide range of NO incidence energies and orientations was performed. In this study was possible to characterize the fundamental dynamical mechanisms that govern reaction, scattering, and energy redistribution processes. One of the key findings was the observation of NO_2 formation at incidence energies as low as 0.025 eV, indicating that the oxidation reaction can proceed efficiently even under mild thermal conditions. This result suggests that the reaction may be more accessible than previously predicted by static DFT studies [68, 70] in a similar system, while qualitatively supporting earlier experimental observations of efficient NO conversion on oxidized graphite surfaces at room temperature [60, 62, 223]. These insights directly contribute to the first objective of the thesis, namely clarifying the role of thermal effects and molecular dynamics in enabling NO oxidation. In more details, our results confirm that the reaction follows a direct Eley–Rideal mechanism when the incoming NO molecule interacts with the adsorbed oxygen in the surface, proceeding rapidly and with an exothermicity of approximately 1.6 eV. These observations fulfill the second thesis objective by elucidating the atomistic mechanism and identifying structural requirements for the reaction to take place.

Beyond reactivity, the AIMD simulations also provided valuable insight into energy transfer processes during non-reactive collisions. Scattered NO molecules typically exhibited substantial translational energy loss, with a dependency on the incidence

orientation: collisions at normal incidence led to stronger momentum transfer and greater energy loss than oriented incidence. Angular distributions of scattered NO and desorbed NO_2 products were found to be closely aligned, reinforcing the direct nature of the reaction and suggesting that scattering observables could serve as experimental indicators of surface functionalization and reactivity. These findings contribute to the third thesis objective, providing a detailed picture of momentum and energy redistribution during gas–surface interactions. Overall, the AIMD results provide valuable guidance for interpreting observables related to product formation and scattering in future experimental studies. More broadly, the findings highlight the efficiency of oxidized carbonaceous materials in promoting NO-to-NO_2 conversion under mild conditions, reinforcing their relevance for environmental catalysis and pollution mitigation strategies.

This study also provides a high-quality reference dataset for the development and validation of ML-PES, allowing for more efficient simulations maintaining the accuracy of the DFT-level of theory. Therefore, the second part of this thesis focused on the development of an ML-PES to enable larger-scale MD simulations of NO oxidation on the O@HOPG surface. The ML-PES was constructed using representative configurations spanning the full dynamical landscape of the reaction, from weakly physisorbed states to fully formed products. Once the model was trained, an intensive validation of the ML-PES was performed by reproducing the AIMD simulations. In this context, the use of ML-PES for trajectory propagation resulted in an efficiency gain of nearly 800 times compared to AIMD. A detailed comparison between results from both approaches, revealed good agreement across a range of dynamical observables. The ML-PES reasonably reproduced the reaction probabilities and correctly captured their dependence on incidence energy and molecular orientation. Reaction times and exothermicity were also well reproduced, further confirming the model's robustness in capturing both the thermodynamic and kinetic features of the process. Moreover, the ML-PES also demonstrated good predictive performance in describing non-reactive scattering events, including angular and energy distributions.

Overall, these results indicate that the obtained ML-PES captures the main features of the reaction dynamics with good reliability. However, they also reveal important lessons about the limitations and challenges of applying ML-PES models to complex reactive gas–surface systems, lessons that shape the methodological perspective adopted throughout this thesis. One of the central messages that emerges is that statistical validation alone is not sufficient to guarantee the reliability of a ML-PES to perform MD simulations in gas-surface systems. In systems involving reactive gas–surface interactions, the configuration space is not only high-dimensional by nature but also structurally complex and highly specific. In such cases, rigorous dynamical validation is essential to ensure that the ML-PES can reliably capture the relevant physical and chemical processes. This thesis demonstrates that relying solely on statistical metrics or validation against static configurations is not sufficient; proper assessment of a ML-PES

must include MD simulations that test its behavior under realistic conditions. Even with a carefully designed and well-balanced training set, the foundation upon which the ML-PES is trained, subtle deviations in reactivity and energy transfer can still emerge when using the ML-PES to perform MD simulations. This underscores the intrinsic difficulty of achieving full configurational coverage when constructing a perfectly representative training set for systems as complex as gas–surface interfaces, particularly when chemical reactivity is involved.

Addressing this challenge calls for more adaptive and systematic sampling strategies capable of enriching the training set where it matters most. Looking ahead, the continued refinement of the ML-PES for NO oxidation on O@HOPG will rely on the integration of advanced sampling strategies, notably those based on active learning. Such approaches promise a more rigorous and automated exploration of the configurational space, which should significantly improve the transferability and predictive accuracy of the ML-PES, enabling more reliable and efficient MD simulations. Beyond the framework developed in this thesis, several alternative machine learning potentials merit consideration for future work. Approaches based on deep neural networks, graph-based architectures, and physically informed models have shown considerable promise in the broader field of reactive dynamics, offering complementary perspectives to the Behler-Parrinello scheme.

Finally, the versatility of the methodology developed in this thesis opens up numerous avenues for future research. In particular, the integrated AIMD and ML-PES framework established here can be readily extended to investigate a variety of gas–surface systems relevant to atmospheric and astrochemical processes. For instance, future studies may address the oxidation and heterogeneous reactivity of other atmospherically relevant molecules, such as SO_2, O_3, or NO_2, on graphite surfaces functionalized not only with oxygen, but also with typical atmospheric groups such as hydroxyl, carbonyl, or carboxyl functionalities. Such research directions are particularly relevant given the prevalence of these functional groups on atmospheric soot particles following oxidation processes. Furthermore, interactions with small molecules — such as water vapor (H_2O), ammonia (NH_3), and VOC — are highly relevant, their integration into model systems poses significant challenges and necessitates a progressive, iterative approach. Nonetheless, recent advances in numerical techniques, particularly foundational ML-PES, hold promise for overcoming these barriers. These developments are expected to facilitate more accurate modeling of molecular complexity and expand simulations across broader spatial and temporal scales, ultimately helping to bridge the gap between computational models, laboratory experiments, and field observations.

Bibliography

[1] C. Díaz, "Theoretical Simulations of Reactive and Nonreactive Scattering of Light Diatomic Molecules from Metal Surfaces: Past, Present, and Future", Advances in Chemistry **2014**, 1–21 (2014).

[2] V. Bortolani, N. H. March, and M. P. Tosi, *Interaction of Atoms and Molecules with Solid Surfaces* (Springer US, Boston, MA, 1990).

[3] B. I. Bentley, "Scattering, Adsorption, and Langmuir-Hinshelwood Desorption Models for Physisorptive and Chemisorptive Gas-Surface Systems", PhD thesis (Theses and Dissertations, 2013).

[4] D. Borodin, I. Rahinov, P. R. Shirhatti, M. Huang, A. Kandratsenka, D. J. Auerbach, T. Zhong, H. Guo, D. Schwarzer, T. N. Kitsopoulos, and A. M. Wodtke, "Following the microscopic pathway to adsorption through chemisorption and physisorption wells", Science **369**, 1461–1465 (2020).

[5] I. Langmuir, "The adsorption of gases on plane surfaces of glass, mica and platinum", Journal of the American Chemical Society **40**, 1361–1403 (1918).

[6] S. Hinshelwood Cyril Norman, *The kinetics of chemical change* (Clarendon Press, 1940).

[7] E. K. Rideal, "A note on a simple molecular mechanism for heterogeneous catalytic reactions", Mathematical Proceedings of the Cambridge Philosophical Society **35**, 130–132 (1939).

[8] D. D. Eley and E. K. Rideal, "Parahydrogen conversion on tungsten", Nature **146**, 401–402 (1940).

[9] T. Kammler, D. Kolovos-Vellianitis, and J. Küppers, "A hot-atom reaction kinetic model for H abstraction from solid surfaces", Surface Science **460**, 91–100 (2000).

[10]C. E. Kolb, R. A. Cox, J. P. D. Abbatt, M. Ammann, E. J. Davis, D. J. Donaldson, B. C. Garrett, C. George, P. T. Griffiths, D. R. Hanson, M. Kulmala, G. McFiggans, U. Pöschl, I. Riipinen, M. J. Rossi, Y. Rudich, P. E. Wagner, P. M. Winkler, D. R. Worsnop, and C. D. O' Dowd, "An overview of current issues in the uptake of atmospheric trace gases by aerosols and clouds", Atmospheric Chemistry and Physics **10**, 10561–10605 (2010).

[11]J. Zhang, Y. Wei, and Z. Fang, "Ozone Pollution: A Major Health Hazard Worldwide", Frontiers in Immunology **10**, 2518 (2019).

[12]İ. A. Reşitoğlu, "No$_x$ pollutants from diesel vehicles and trends in the control technologies", in *Diesel and gasoline engines* (IntechOpen, Rijeka, 2018) Chap. 8.

[13]W. de Vries, "Impacts of nitrogen emissions on ecosystems and human health: a mini review", Current Opinion in Environmental Science & Health **21**, 100249 (2021).

[14]A. L. D. Villarejo, "Nitrogen oxides and their role in air pollution", Anales de la Real Academia Nacional de Farmacia **88**, 345–350 (2022).

[15]O. Hertel, C. A. Skjøth, S. Reis, A. Bleeker, R. M. Harrison, J. N. Cape, D. Fowler, U. Skiba, D. Simpson, T. Jickells, M. Kulmala, S. Gyldenkærne, L. L. Sørensen, J. W. Erisman, and M. A. Sutton, "Governing processes for reactive nitrogen compounds in the european atmosphere", Biogeosciences **9**, 4921–4954 (2012).

[16]M. Ammann, M. Kalberer, D. T. Jost, L. Tobler, E. Rössler, D. Piguet, H. W. Gäggeler, and U. Baltensperger, "Heterogeneous production of nitrous acid on soot in polluted air masses", Nature **395**, 157–160 (1998).

[17]K. Tabor, L. Gutzwiller, and M. J. Rossi, "Heterogeneous chemical kinetics of NO2 on amorphous carbon at ambient temperature", The Journal of Physical Chemistry **98**, 6172–6186 (1994).

[18]J. Xi and B.-J. Zhong, "Soot in Diesel Combustion Systems", Chemical Engineering & Technology **29**, 665–673 (2006).

[19]M. N. Laritchev and J. C. Petit, "Soot particles from different combustion sources: composition, surface groups, oxidation under atmospheric conditions", in *Global atmospheric change and its impact on regional air quality* (Springer Netherlands, Dordrecht, 2002), pp. 129–135.

[20]X. Cheng, L. Chen, F. Yan, and S. Dong, "Study on soot formation characteristics in the diesel combustion process based on an improved detailed soot model", Energy Conversion and Management **75**, 1–10 (2013).

[21]H. Michelsen, "Probing soot formation, chemical and physical evolution, and oxidation: A review of in situ diagnostic techniques and needs", Proceedings of the Combustion Institute **36**, 717–735 (2017).

[22]J. Wu, A. Faccinetto, S. Batut, M. Cazaunau, E. Pangui, N. Nuns, B. Hanoune, J.-F. Doussin, P. Desgroux, and D. Petitprez, "On the correlation between hygroscopic properties and chemical composition of cloud condensation nuclei obtained from the chemical aging of soot particles with O3 and SO2", Science of The Total Environment **906**, 167745 (2024).

[23]A. N. Nguyen, *Oxidation and atmospheric transformation of vehicle exhaust particles* (University of California, Riverside, 2014).

[24]A. Zygogianni, M. Syrigou, A. G. Konstandopoulos, and M. Kostoglou, "Oxidative reactivity of particulate samples from different diesel combustion systems and its relation to structural and spectral characteristics of soot", Emission Control Science and Technology **5**, 99–123 (2019).

[25]J. C. Kaiser, N. Riemer, and D. A. Knopf, "Detailed heterogeneous oxidation of soot surfaces in a particle-resolved aerosol model", Atmospheric Chemistry and Physics **11**, 4505–4520 (2011).

[26]T. Bierkandt, P. Oßwald, T. Schripp, and M. Köhler, "Experimental Investigation of Soot Oxidation under Well-Controlled Conditions in a High-Temperature Flow Reactor", Combustion Science and Technology **191**, 1499–1519 (2019).

[27]A. Messerer, R. Niessner, and U. Pöschl, "Comprehensive kinetic characterization of the oxidation and gasification of model and real diesel soot by nitrogen oxides and oxygen under engine exhaust conditions: Measurement, Langmuir–Hinshelwood, and Arrhenius parameters", Carbon **44**, 307–324 (2006).

[28]C. Garcia-Fernandez, S. Picaud, M. T. Rayez, J. C. Rayez, and J. Rubayo-Soneira, "First-Principles Study of the Interaction between NO and Large Carbonaceous Clusters Modeling the Soot Surface", The Journal of Physical Chemistry A **118**, 1443–1450 (2014).

[29]Q. Mao, A. C. Van Duin, and K. Luo, "Formation of incipient soot particles from polycyclic aromatic hydrocarbons: A ReaxFF molecular dynamics study", Carbon **121**, 380–388 (2017).

[30]Z. Li, W. Zhang, Z. Chen, Q. Zhang, X. Yang, S. Mao, and W. Jian, "Reaction mechanism for NO oxidation on the soot surface using a quantum chemistry", Fuel **313**, 123032 (2022).

[31]F. Hagen, F. Hardock, S. Koch, N. Sebbar, H. Bockhorn, A. Loukou, H. Kubach, R. Suntz, D. Trimis, and T. Koch, "Why soot is not alike soot: a molecular/nanostructural approach to low temperature soot oxidation", Flow, Turbulence and Combustion **106**, 295–329 (2021).

[32]J. Xi, G. Yang, J. Cai, and Z. Gu, "A Review of Recent Research Results on Soot: The Formation of a Kind of Carbon-Based Material in Flames", Frontiers in Materials **8**, 695485 (2021).

[33]G. Ghigo, A. Maranzana, G. Tonachini, C. M. Zicovich-Wilson, and M. Causà, "Modeling soot and its functionalization under atmospheric or combustion conditions by density functional theory within molecular (polycyclic-aromatic-hydrocarbon-like) and periodic methodologies", The Journal of Physical Chemistry B **108**, 3215–3223 (2004).

[34]M. Seipenbusch, J. V. Erven, T. Schalow, A. P. Weber, A. D. V. Langeveld, J. C. Marijnissen, and S. K. Friedlander, "Catalytic soot oxidation in microscale experiments", Applied Catalysis B: Environmental **55**, 31–37 (2005).

[35]S. Morisset, N. Rougeau, and D. Teillet-Billy, "Hydrogenation reactions and adsorption : from CO to methanol on a graphene surface", Molecular Astrophysics **14**, 1–9 (2019).

[36]F. Dulieu, S. Morisset, A.-S. Ibrahim Mohamed, L. Boshman, S. Cazaux, D. Teillet-Billy, S. Baouche, and N. Rougeau, "Reactivity of coronene with O-atoms, a possible route to ketene in the interstellar medium", Molecular Astrophysics **17**, 100054 (2019).

[37]D. Kang, "Lightning nox emissions and the implications for surface air quality over the contiguous united states", EM: Air and Waste Management Association's Magazine for Environmental Managers **11** (2018).

[38]J. Li and Y. Wang, "Inferring the anthropogenic no_x emission trend over the united states during 2003-2017 from satellite observations: was there a flattening of the emission trend after the great recession?", Atmospheric Chemistry and Physics **19**, 15339–15352 (2019).

[39]A. J. Kunov-Kruse, P. L. Thomassen, A. Riisager, S. Mossin, and R. Fehrmann, "Absorption and oxidation of nitrogen oxide in ionic liquids", Chemistry – A European Journal **22**, 11745–11755 (2016).

[40]G. Li, B. Wang, W. Q. Xu, Y. Li, Y. Han, and Q. Sun, "Simultaneous removal of SO_2 and NO_x from flue gas by wet scrubbing using a urea solution", Environmental Technology **40**, 2620–2632 (2019).

[41]Z. Zhu and B. Xu, "Purification Technologies for NOx Removal from Flue Gas: A Review", Separations **9**, 307 (2022).

[42]Y. Wang, T. Wang, Q. Gu, and J. Shang, "Adsorption removal of NO_2 under low-temperature and low-concentration conditions: a review of adsorbents and adsorption mechanisms", Advanced Materials **37**, 2401623 (2025).

[43]K. Silas, W. A. W. A. K. Ghani, T. S. Choong, and U. Rashid, "Carbonaceous materials modified catalysts for simultaneous SO_2/NO_xremoval from flue gas: a review", Catalysis Reviews **61**, 134–161 (2019).

[44]Y. Liu, X. Wu, X. Yang, H. Tao, J. Li, C. Zhang, R. T. Yang, and Z. Li, "Enhancement of NOx adsorption performance on zeolite via a facile modification strategy", Journal of Hazardous Materials **443**, 130225 (2023).

[45]M.-W. Liu, H. Zhang, J. Li, X.-C. Qi, Y.-F. Wang, and J. Pang, "Recent progress and future prospects of metal-organic frameworks for adsorption, separation and catalytic removal of NO_x and N_2O", ChemCatChem, e202400922 (2024).

[46]P.-M. Park, Y.-K. Park, and J.-I. Dong, "Reaction characteristics of NOx and N2O in selective non-catalytic reduction using various reducing agents and additives", Atmosphere **12**, 1175 (2021).

[47]Z. Gholami, G. Luo, F. Gholami, and F. Yang, "Recent advances in selective catalytic reduction of NO_x by carbon monoxide for flue gas cleaning process: a review", Catalysis Reviews **63**, 68–119 (2021).

[48]B. Guan, R. Zhan, H. Lin, and Z. Huang, "Review of state of the art technologies of selective catalytic reduction of NOx from diesel engine exhaust", Applied Thermal Engineering **66**, 395–414 (2014).

[49]A. V. Karre, R. K. Garlapalli, A. Jena, and N. Tripathi, "State of the art developments in oxidation performance and deactivation of diesel oxidation catalyst (DOC)", Catalysis Communications **179**, 106682 (2023).

[50]A. Russell and W. S. Epling, "Diesel oxidation catalysts", Catalysis Reviews **53**, 337–423 (2011).

[51]M. Singh, M. Srilomsak, Y. Wang, K. Hanamura, and R. Vander Wal, "Nanostructure changes in diesel soot during NO_2–O_2 oxidation under diesel particulate filter-like conditions toward filter regeneration", International Journal of Engine Research **20**, 953–966 (2019).

[52]G. Qi and W. Li, "NOx adsorption and reduction over LaMnO3 based lean NOx trap catalysts", Catalysis Letters **144**, 639–647 (2014).

[53]H. Kim and K. B. Lee, "Enhanced low-temperature NOx storage and reduction performance using cu-loaded lean NOx trap catalysts", Chemical Engineering Journal **470**, 143895 (2023).

[54]Z. Zhang, D. Li, G. Lan, Z. Yin, M. Pan, F. Jiang, J. Li, and D. Tan, "Development and evaluation of mechanistic model for standard SCR, fast SCR, and NO2 SCR of NH3-SCR over cu-ZSM-5", Energy **306**, 132544 (2024).

[55]Y. Shen, X. Ge, and M. Chen, "Catalytic oxidation of nitric oxide (NO) with carbonaceous materials", RSC Advances **6**, 8469–8482 (2016).

[56]Z. Hong, Z. Wang, and X. Li, "Catalytic oxidation of nitric oxide (NO) over different catalysts: an overview", Catalysis Science & Technology **7**, 3440–3452 (2017).

[57]E. Bouleghlimat, P. R. Davies, R. J. Davies, R. Howarth, J. Kulhavy, and D. J. Morgan, "The effect of acid treatment on the surface chemistry and topography of graphite", Carbon **61**, 124–133 (2013).

[58]Y.-R. Shin, S.-M. Jung, I.-Y. Jeon, and J.-B. Baek, "The oxidation mechanism of highly ordered pyrolytic graphite in a nitric acid/sulfuric acid mixture", Carbon **52**, 493–498 (2013).

[59]J. P. Sousa, M. F. Pereira, and J. L. Figueiredo, "Catalytic oxidation of NO to NO_2 on N-doped activated carbons", Catalysis Today **176**, 383–387 (2011).

[60]I. Mochida, S. Kisamori, M. Hironaka, S. Kawano, Y. Matsumura, and M. Yoshikawa, "Oxidation of NO into NO2 over Active Carbon Fibers", Energy & Fuels 1994, 8, 1341-1344, 4 (1994).

[61]I. Mochida, N. Shirahama, S. Kawano, Y. Korai, A. Yasutake, M. Tanoura, S. Fujii, and M. Yoshikawa, "NO oxidation over activated carbon fiber (ACF). Part 1. Extended kinetics over a pitch based ACF of very large surface area", Fuel **79**, 1713–1723 (2000).

[62]S. Adapa, V. Gaur, and N. Verma, "Catalytic oxidation of NO by activated carbon fiber (ACF)", Chemical Engineering Journal **116**, 25–37 (2006).

[63]Z. Zhang, J. D. Atkinson, B. Jiang, M. J. Rood, and Z. Yan, "Nitric oxide oxidation catalyzed by microporous activated carbon fiber cloth: an updated reaction mechanism", Applied Catalysis B: Environmental **148–149**, 573–581 (2014).

[64]J. D. Atkinson, Z. Zhang, Z. Yan, and M. J. Rood, "Evolution and impact of acidic oxygen functional groups on activated carbon fiber cloth during NO oxidation", Carbon **54**, 444–453 (2013).

[65]I. D. Rosca, F. Watari, M. Uo, and T. Akasaka, "Oxidation of multiwalled carbon nanotubes by nitric acid", Carbon **43**, 3124–3131 (2005).

[66]J. P. Sousa, M. F. Pereira, and J. L. Figueiredo, "NO oxidation over nitrogen doped carbon xerogels", Applied Catalysis B: Environmental **125**, 398–408 (2012).

[67]Z. Zeng, P. Lu, C. Li, L. Mai, Z. Li, and Y. Zhang, "Removal of NO by carbonaceous materials at room temperature: A review", Catalysis Science & Technology **2**, 2188 (2012).

[68]W. Cen, M. Hou, J. Liu, S. Yuan, Y. Liu, and Y. Chu, "Oxidation of SO_2 and NO by epoxy groups on graphene oxides: the role of the hydroxyl group", RSC Adv. **5**, 22802–22810 (2015).

[69]S. Tang and Z. Cao, "Adsorption of nitrogen oxides on graphene and graphene oxides: Insights from density functional calculations", The Journal of Chemical Physics **134**, 044710 (2011).

[70]M. Hou, W. Cen, H. Zhang, J. Liu, H. Yin, and F. Wei, "Adsorption and oxidation of NO on graphene oxides: A dispersion corrected density functional theory investigation", Applied Surface Science **339**, 55–61 (2015).

[71]P. Nowicki and R. Pietrzak, "Carbonaceous adsorbents prepared by physical activation of pine sawdust and their application for removal of NO2 in dry and wet conditions", Bioresource Technology **101**, 5802–5807 (2010).

[72]T. Suzuki, T. Kyotani, and A. Tomita, "Study on the Carbon-Nitric Oxide Reaction in the Presence of Oxygen", Industrial & Engineering Chemistry Research **33**, 2840–2845 (1994).

[73]R. S. Rathore, D. K. Srivastava, A. K. Agarwal, and N. Verma, "Development of surface functionalized activated carbon fiber for control of NO and particulate matter", Journal of Hazardous Materials **173**, 211–222 (2010).

[74]S. Ahmed, R. Baldwin, F. Derbyshire, B. McEnaney, and J. Stencel, "Catalytic reduction of nitric oxide over activated carbons", Fuel **72**, 287–292 (1993).

[75]Z. Zhang, J. D. Atkinson, B. Jiang, M. J. Rood, and Z. Yan, "Nitric oxide oxidation catalyzed by microporous activated carbon fiber cloth: an updated reaction mechanism", Applied Catalysis B-environmental **148**, 573–581 (2014).

[76]Z. Hong, Z. L. Wang, and X. Li, "Catalytic oxidation of nitric oxide (no) over different catalysts: an overview", Catalysis Science & Technology **7**, 3440–3452 (2017).

[77]M. Elstner and G. Seifert, "Density functional tight binding", Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **372**, 20120483 (2014).

[78]M. Zheng and M. P. Waller, "Adaptive quantum mechanics/molecular mechanics methods", WIREs Computational Molecular Science **6**, 369–385 (2016).

[79]I. Leven, H. Hao, S. Tan, X. Guan, K. A. Penrod, D. Akbarian, B. Evangelisti, M. J. Hossain, M. M. Islam, J. P. Koski, S. Moore, H. M. Aktulga, A. C. T. Van Duin, and T. Head-Gordon, "Recent advances for improving the accuracy, transferability, and efficiency of reactive force fields", Journal of Chemical Theory and Computation **17**, 3237–3251 (2021).

[80]J. Xia, Y. Zhang, and B. Jiang, "The evolution of machine learning potentials for molecules, reactions and materials", Chemical Society Reviews **54**, 4790–4821 (2025).

[81]J. Behler and M. Parrinello, "Generalized neural-network representation of high-dimensional potential-energy surfaces", Physical Review Letters **98**, 146401 (2007).

[82]N. Artrith and A. Urban, "An implementation of artificial neural-network potentials for atomistic materials simulations: performance for tio2", Comput. Mater. Sci **114**, 135–150 (2016).

[83]Y. Zhang, C. Hu, and B. Jiang, "Embedded atom neural network potentials: efficient and accurate machine learning with a physically inspired representation", The Journal of Physical Chemistry Letters **10**, 4962–4967 (2019).

[84]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 **228**, 178–184 (2018).

[85]I. Batatia, P. Benner, Y. Chiang, A. M. Elena, D. P. Kovács, J. Riebesell, X. R. Advincula, M. Asta, M. Avaylon, W. J. Baldwin, F. Berger, N. Bernstein, A. Bhowmik, S. M. Blau, V. Cărare, J. P. Darby, S. De, F. Della Pia, V. L. Deringer, R. Elijošius, Z. El-Machachi, F. Falcioni, E. Fako, A. C. Ferrari, A. Genreith-Schriever, J. George, R. E. A. Goodall, C. P. Grey, P. Grigorev, S. Han, W. Handley, H. H. Heenen, K. Hermansson, C. Holm, J. Jaafar, S. Hofmann, K. S. Jakob, H. Jung, V. Kapil, A. D. Kaplan, N. Karimitari, J. R. Kermode, N. Kroupa, J. Kullgren, M. C. Kuner, D. Kuryla, G. Liepuoniute, J. T. Margraf, I.-B. Magdău, A. Michaelides, J. H. Moore, A. A. Naik, S. P. Niblett, S. W. Norwood, N. O'Neill, C. Ortner, K. A. Persson, K. Reuter, A. S. Rosen, L. L. Schaaf, C. Schran, B. X. Shi, E. Sivonxay, T. K. Stenczel, V. Svahn, C. Sutton, T. D. Swinburne, J. Tilly, C. van der Oord, E. Varga-Umbrich, T. Vegge, M. Vondrák, Y. Wang, W. C. Witt, F. Zills, and G. Csányi, "A foundation model for atomistic materials chemistry", arXiv **2401.00096** (2024).

[86]B. Kolb, X. Luo, X. Zhou, B. Jiang, and H. Guo, "High-dimensional atomistic neural network potentials for molecule–surface interactions: hcl scattering from au(111)", The Journal of Physical Chemistry Letters **8**, 666–672 (2017).

[87]K. Shakouri, J. Behler, J. Meyer, and G.-J. Kroes, "Accurate neural network description of surface phonons in reactive gas–surface dynamics: n_2 + ru(0001)", The Journal of Physical Chemistry Letters **8**, 2131–2136 (2017).

[88]B. Jiang, J. Li, and H. Guo, "High-fidelity potential energy surfaces for gas-phase and gas–surface scattering processes from machine learning", The Journal of Physical Chemistry Letters **11**, 5120–5131 (2020).

[89]A. Rivero Santamaría, M. Alducin, R. Díez Muiño, J. I. Juaristi, P. Larrégaray, and L. Bonnet, "High-dimensional atomistic neural network potential to study the alignment-resolved o_2 scattering from highly oriented pyrolytic graphite", The Journal of Physical Chemistry A **125**, 2588–2600 (2021).

[90]A. Serrano Jiménez, A. P. Sánchez Muzas, Y. Zhang, J. Ovčar, B. Jiang, I. Lončarić, J. I. Juaristi, and M. Alducin, "Photoinduced desorption dynamics of co from pd(111): a neural network approach", Journal of Chemical Theory and Computation **17**, 4648–4659 (2021).

[91]I. Žugec, A. Tetenoire, M. Alducin, J. I. Juaristi, M. Blanco-Rey, and D. Novko, "Understanding the photoinduced desorption and oxidation of co on ru(0001) using a neural network potential energy surface", JACS Au **4**, 1997–2004 (2024).

[92]R. Díez Muiño and H. F. Busnengo, eds., *Dynamics of Gas-Surface Interactions: Atomic-level Understanding of Scattering Processes at Surfaces*, Vol. 50, Springer Series in Surface Sciences (Springer Berlin Heidelberg, Berlin, Heidelberg, 2013).

[93]B. I. Bentley, *Scattering, adsorption, and langmuir-hinshelwood desorption models for physisorptive and chemisorptive gas-surface systems* (Air Force Institute of Technology, 2013).

[94]J. Ischtwan and M. A. Collins, "Molecular potential energy surfaces by interpolation", The Journal of Chemical Physics **100**, 8080–8088 (1994).

[95]K. C. Thompson and M. A. Collins, "Molecular potential-energy surfaces by interpolation: further refinements", Journal of the Chemical Society, Faraday Transactions **93**, 871–878 (1997).

[96]B. J. Braams and J. M. Bowman, "Permutationally invariant potential energy surfaces in high dimensionality", International Reviews in Physical Chemistry **28**, 577–606 (2009).

[97]J. M. Bowman, G. Czakó, and B. Fu, "High-dimensional ab initio potential energy surfaces for reaction dynamics calculations", Physical Chemistry Chemical Physics **13**, 8094 (2011).

[98]H.-Y. Kwon, Z. Morrow, C. T. Kelley, and E. Jakubikova, "Interpolation methods for molecular potential energy surface construction", The Journal of Physical Chemistry A **125**, 9725–9735 (2021).

[99]H. F. Busnengo, A. Salin, and W. Dong, "Representation of the 6D potential energy surface for a diatomic molecule near a solid surface", The Journal of Chemical Physics **112**, 7641–7651 (2000).

[100]P. Spiering, M. Wijzenbroek, and M. F. Somers, "An improved static corrugation model", The Journal of Chemical Physics **149**, 234702 (2018).

[101]B. Smits and M. F. Somers, "Beyond the static corrugation model: dynamic surfaces with the embedded atom method", The Journal of Chemical Physics **154**, 074710 (2021).

[102]C. Crespos, M. A. Collins, E. Pijper, and G. J. Kroes, "Application of the modified Shepard interpolation method to the determination of the potential energy surface for a molecule–surface reaction: H2+Pt(111)", The Journal of Chemical Physics **120**, 2392–2404 (2004).

[103]W. G. Stark, C. van der Oord, I. Batatia, Y. Zhang, B. Jiang, G. Csányi, and R. J. Maurer, "Benchmarking of machine learning interatomic potentials for reactive hydrogen dynamics at metal surfaces", Machine Learning: Science and Technology **5**, 030501 (2024).

[104]M. E. Tuckerman, "Ab initio molecular dynamics: basic concepts, current trends and novel applications", Journal of Physics: Condensed Matter **14**, R1297 (2002).

[105]R. Car and M. Parrinello, "Unified approach for molecular dynamics and density-functional theory", Phys. Rev. Lett. **55**, 2471–2474 (1985).

[106]M. Blanco-Rey, J. I. Juaristi, R. Díez Muiño, H. F. Busnengo, G. J. Kroes, and M. Alducin, "Electronic friction dominates hydrogen hot-atom relaxation on pd(100)", Physical Review Letters **112**, 103203 (2014).

[107]E. Schrödinger, "Quantisierung als eigenwertproblem", Annalen der Physik **384**, 361–376 (1926).

[108]H. Shull and G. G. Hall, "Atomic units", Nature **184**, 1559–1560 (1959).

[109]M. Born and R. Oppenheimer, "Zur quantentheorie der molekeln", Annalen der Physik **389**, 457–484 (1927).

[110]C. C. J. Roothaan, "New developments in molecular orbital theory", Reviews of Modern Physics **23**, 69–89 (1951).

[111]T. Kato, "On the eigenfunctions of many-particle systems in quantum mechanics", Communications on Pure and Applied Mathematics **10**, 151–177 (1957).

[112]B. Simon, "Tosio kato's work on non-relativistic quantum mechanics, part 2", Bulletin of Mathematical Sciences **9**, 1950005 (2019).

[113]P. Hohenberg and W. Kohn, "Inhomogeneous electron gas", Physical Review **136**, B864–B871 (1964).

[114]W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects", Physical Review **140**, A1133–A1138 (1965).

[115]J. P. Perdew and K. Schmidt, "Jacob's ladder of density functional approximations for the exchange-correlation energy", AIP Conference Proceedings **577**, 1–20 (2001).

[116]M. Bursch, J.-M. Mewes, A. Hansen, and S. Grimme, "Best-practice DFT protocols for basic molecular computational chemistry**", Angewandte Chemie International Edition **61**, e202205735 (2022).

[117]D. M. Ceperley and B. J. Alder, "Ground state of the electron gas by a stochastic method", Physical Review Letters **45**, 566–569 (1980).

[118]R. G. Parr and Y. Weitao, *Density-functional theory of atoms and molecules* (Oxford University Press, 1989).

[119]J. P. Perdew, "Accurate density functional for the energy: real-space cutoff of the gradient expansion for the exchange hole", Physical Review Letters **55**, 1665 (1985).

[120]J. P. Perdew and W. Yue, "Accurate and simple density functional for the electronic exchange energy: generalized gradient approximation", Physical review B **33**, 8800 (1986).

[121]M. Filatov and W. Thiel, "A new gradient-corrected exchange-correlation density functional", Molecular Physics **91**, 847–860 (1997).

[122]J. P. Perdew, "Unified theory of exchange and correlation beyond the local density approximation", Electronic structure of solids' 91 **11** (1991).

[123]K. Burke, J. P. Perdew, and Y. Wang, "Derivation of a generalized gradient approximation: the pw91 density functional", in *Electronic density functional theory: recent progress and new directions* (Springer, 1998), pp. 81–111.

[124]G. J. Laming, V. Termath, and N. C. Handy, "A general purpose exchange-correlation energy functional", The Journal of chemical physics **99**, 8765–8773 (1993).

[125]A. D. Becke, "Density functional calculations of molecular bond energies", The Journal of Chemical Physics **84**, 4524–4529 (1986).

[126]J. P. Perdew, "Density-functional approximation for the correlation energy of the inhomogeneous electron gas", Physical review B **33**, 8822 (1986).

[127]D. J. Lacks and R. G. Gordon, "Pair interactions of rare-gas atoms as a test of exchange-energy-density functionals in regions of large density gradients", Physical Review A **47**, 4681 (1993).

[128]J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized Gradient Approximation Made Simple", Phys. Rev. Lett. **77**, 3865–3868 (1996).

[129]J. P. Perdew and Y. Wang, "Accurate and simple analytic representation of the electron-gas correlation energy", Physical review B **45**, 13244 (1992).

[130]C. Lee, W. Yang, and R. G. Parr, "Development of the colle-salvetti correlation-energy formula into a functional of the electron density", Physical review B **37**, 785 (1988).

[131]J. P. Perdew, S. Kurth, A. Zupan, and P. Blaha, "Accurate density functional with correct formal properties: a step beyond the generalized gradient approximation", Physical review letters **82**, 2544 (1999).

[132]E. I. Proynov, A. Vela, and D. R. Salahub, "Nonlocal correlation functional involving the laplacian of the density", Chemical physics letters **230**, 419–428 (1994).

[133]A. D. Becke, "Density-functional thermochemistry. III. The role of exact exchange", The Journal of Chemical Physics **98**, 5648–5652 (1993).

[134]C. Adamo and V. Barone, "Toward reliable density functional methods without adjustable parameters: the pbe0 model", The Journal of chemical physics **110**, 6158–6170 (1999).

[135]J. Heyd, G. E. Scuseria, and M. Ernzerhof, "Hybrid functionals based on a screened coulomb potential", The Journal of Chemical Physics **118**, 8207–8215 (2003).

[136]J. C. Sancho-García and C. Adamo, "Double-hybrid density functionals: merging wavefunction and density approaches to get the best of both worlds", Physical Chemistry Chemical Physics **15**, 14581 (2013).

[137]C. Møller and M. S. Plesset, "Note on an approximation treatment for many-electron systems", Physical Review **46**, 618–622 (1934).

[138]S. Grimme, "Semiempirical hybrid density functional with perturbative second-order correlation", The Journal of Chemical Physics **124**, 3 (2006).

[139]A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior", Physical Review A **38**, 3098–3100 (1988).

[140]S. Grimme, "Density functional theory with London dispersion corrections", WIREs Computational Molecular Science **1**, 211–228 (2011).

[141]S. Grimme, "Accurate description of van der waals complexes by density functional theory including empirical corrections", Journal of Computational Chemistry **25**, 1463–1473 (2004).

[142]S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu", The Journal of Chemical Physics **132**, 154104 (2010).

[143]E. R. Johnson, I. D. Mackie, and G. A. DiLabio, "Dispersion interactions in density-functional theory", Journal of Physical Organic Chemistry **22**, 1127–1135 (2009).

[144]E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth, and S. Grimme, "A generally applicable atomic-charge dependent london dispersion correction", The Journal of Chemical Physics **150**, 15 (2019).

[145]A. D. Becke and E. R. Johnson, "Exchange-hole dipole moment and the dispersion interaction: high-order dispersion coefficients", The Journal of Chemical Physics **124**, 1 (2006).

[146]R. Iftimie, P. Minary, and M. E. Tuckerman, "Ab initio molecular dynamics: concepts, recent developments, and future trends", Proceedings of the National Academy of Sciences **102**, 6654–6659 (2005).

[147]E. Paquet and H. L. Viktor, "Computational methods for ab initio molecular dynamics", Advances in Chemistry **2018**, 9839641 (2018).

[148]D. Marx and J. Hutter, “Ab initio molecular dynamics: theory and implementation”, in *Modern methods and algorithms of quantum chemistry*, NIC series 1, Num Pages: 562 (NIC, 2000).

[149]R. P. Feynman, “Forces in molecules”, Phys. Rev. **56**, 340–343 (1939).

[150]R. M. Martin, “Electronic structure: basic theory and practical methods”, in Electronic structure: basic theory and practical methods (2004).

[151]D. Frenkel and B. Smit, *Understanding molecular simulation: from algorithms to applications* (Elsevier, 2023).

[152]L. Verlet, “Computer "experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules”, Phys. Rev. **159**, 98–103 (1967).

[153]G. A. Alou Angulo, A. Rivero Santamaría, C. Toubin, and M. Monnerville, “Ab initio molecular dynamics calculations on NO oxidation over oxygen-functionalized highly oriented pyrolytic graphite”, The Journal of Physical Chemistry C **128**, 17894–17904 (2024).

[154]G. Kresse and J. Furthmüller, “Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set”, Comput. Mater. Sci. **6**, 15–50 (1996).

[155]G. Kresse and J. Furthmüller, “Efficient iterative schemes for ab-initio total-energy calculations using a plane-wave basis set”, Physical Review B **54**, 11169–11186 (1996).

[156]B. Hammer, L. B. Hansen, and J. K. Nørskov, “Improved adsorption energetics within density-functional theory using revised perdew-burke-ernzerhof functionals”, Physical Review B **59**, 7413–7421 (1999).

[157]B. Liu, L. Cheng, L. Curtiss, and J. Greeley, “Effects of van der waals density functional corrections on trends in furfural adsorption and hydrogenation on close-packed transition metal surfaces”, Surface Science **622**, 51–59 (2014).

[158]D. Loffreda, C. Michel, F. Delbecq, and P. Sautet, “Tuning catalytic reactivity on metal surfaces: insights from DFT”, Journal of Catalysis **308**, 374–385 (2013).

[159]O. Mohan, Q. T. Trinh, A. Banerjee, and S. H. Mushrif, “Predicting CO_2 adsorption and reactivity on transition metal surfaces using popular density functional theory methods”, Molecular Simulation **45**, 1163–1172 (2019).

[160]J. Wellendorff, T. L. Silbaugh, D. Garcia-Pintos, J. K. Nørskov, T. Bligaard, F. Studt, and C. T. Campbell, “A benchmark database for adsorption bond energies to transition metal surfaces and comparison to selected DFT functionals”, Surface Science **640**, 36–44 (2015).

[161] Y. Yu and C. Dai, "DFT study of gas adsorption and sensing based on noble metal (Ag, Au and Pt) functionalized boron selenide nanosheets", Physica E: Low-dimensional Systems and Nanostructures **125**, 114409 (2021).

[162] S. Rangarajan and M. Mavrikakis, "A comparative analysis of different van der waals treatments for molecular adsorption on the basal plane of 2H-MoS2", Surface Science **729**, 122226 (2023).

[163] C. Lechner, B. Pannier, P. Baranek, N. C. Forero-Martinez, and H. Vach, "First-Principles Study of the Structural, Electronic, Dynamic, and Mechanical Properties of HOPG and Diamond: Influence of Exchange–Correlation Functionals and Dispersion Interactions", The Journal of Physical Chemistry C **120**, 5083–5100 (2016).

[164] S. Grimme, S. Ehrlich, and L. Goerigk, "Effect of the damping function in dispersion corrected density functional theory", J. Comput. Chem. **32**, 1456–1465 (2011).

[165] A. Rivero Santamaría, M. Alducin, R. Díez Muiño, and J. I. Juaristi, "Ab Initio Molecular Dynamics Study of Alignment-Resolved O $_2$ Scattering from Highly Oriented Pyrolytic Graphite", The Journal of Physical Chemistry C **123**, 31094–31102 (2019).

[166] T. A. Pham, K. E. Kweon, A. Samanta, M. T. Ong, V. Lordi, and J. E. Pask, "Intercalation of lithium into graphite: insights from first-principles simulations", The Journal of Physical Chemistry C **124**, 21985–21992 (2020).

[167] T. Baba, K. Sodeyama, Y. Kawamura, and Y. Tateyama, "Li-ion transport at the interface between a graphite anode and Li_2 CO_3 solid electrolyte interphase: *ab initio* molecular dynamics study", Physical Chemistry Chemical Physics **22**, 10764–10774 (2020).

[168] R. Mittal, M. K. Gupta, and S. L. Chaplot, "Phase transition mechanism of hexagonal graphite to hexagonal and cubic diamond: ab initio simulation", Journal of Physics: Condensed Matter **33**, 425403 (2021).

[169] M. T. Ajide, P. Naeiji, J. Klug, and N. J. English, "Machine learning force field predictions of structural and dynamical properties in HOPG defects and the HOPG-water interface with electronic structure analysis", ACS Omega **10**, 25962–25980 (2025).

[170] P. E. Blöchl, "Projector augmented-wave method", Physical Review B **50**, 17953–17979 (1994).

[171] G. Kresse and D. Joubert, "From ultrasoft pseudopotentials to the projector augmented-wave method", Physical Review B **59**, 1758–1775 (1999).

[172] H. J. Monkhorst and J. D. Pack, "Special points for Brillouin-zone integrations", Physical Review B **13**, 5188–5192 (1976).

[173]G. Teobaldi, H. Ohnishi, K. Tanimura, and A. L. Shluger, "The effect of van der waals interactions on the properties of intrinsic defects in graphite", Carbon **48**, 4145–4161 (2010).

[174]T. Bučko, J. Hafner, S. Lebègue, and J. G. Ángyán, "Improved description of the structure of molecular and layered crystals: ab initio DFT calculations with van der waals corrections", The Journal of Physical Chemistry A **114**, 11814–11824 (2010).

[175]Y. Baskin and L. Meyer, "Lattice constants of graphite at low temperatures", Phys. Rev. **100**, 544–544 (1955).

[176]P. Trucano and R. Chen, "Structure of graphite by neutron diffraction", Nature **258**, 136–137 (1975).

[177]C. L. Fu and K. M. Ho, "First-principles calculation of the equilibrium ground-state properties of transition metals: Applications to Nb and Mo", Physical Review B **28**, 5480–5486 (1983).

[178]M. J. Mehl, J. E. Osburn, D. A. Papaconstantopoulos, and B. M. Klein, "Structural properties of ordered high-melting-temperature intermetallic alloys from first-principles total-energy calculations", Physical Review B **41**, 10311–10323 (1990).

[179]F. Birch, "Finite Elastic Strain of Cubic Crystals", Physical Review **71**, 809–824 (1947).

[180]M. Hebbache and M. Zemzemi, "*Ab initio* study of high-pressure behavior of a low compressibility metal and a hard material: Osmium and diamond", Physical Review B **70**, 224107 (2004).

[181]O. L. Blakslee, D. G. Proctor, E. J. Seldin, G. B. Spence, and T. Weng, "Elastic constants of compression-annealed pyrolytic graphite", Journal of Applied Physics **41**, 3373–3382 (1970).

[182]A. Incze, A. Pasturel, and C. Chatillon, "First-principles study of the atomic oxygen adsorption on the (0 0 0 1) graphite surface and dissolution", Applied Surface Science **177**, 226–229 (2001).

[183]A. Incze, A. Pasturel, and C. Chatillon, "Oxidation of graphite by atomic oxygen: a first-principles approach", Surface Science **537**, 55–63 (2003).

[184]A. Barinov, O. B. Malcioğlu, S. Fabris, T. Sun, L. Gregoratti, M. Dalmiglio, and M. Kiskinova, "Initial Stages of Oxidation on Graphitic Surfaces: Photoemission Study and Density Functional Theory Calculations", The Journal of Physical Chemistry C **113**, 9009–9013 (2009).

[185]Y. G. Polynskaya, N. A. Matsokin, A. S. Sinitsa, A. A. Knizhnik, and B. V. Potapkin, "First-principles investigation of interaction between the atomic oxygen species and carbon nanostructures", Carbon Trends **9**, 100201 (2022).

[186]V. Morón, P. Gamallo, and R. Sayós, "DFT and kinetics study of O/O2 mixtures reacting over a graphite (0001) basal surface", Theoretical Chemistry Accounts **128**, 683–694 (2011).

[187]Ž. Šljivančanin, A. S. Milošević, Z. S. Popović, and F. R. Vukajlović, "Binding of atomic oxygen on graphene from small epoxy clusters to a fully oxidized surface", Carbon **54**, 482–488 (2013).

[188]V. Morón, L. Martin-Gondre, C. Crespos, P. Larregaray, P. Gamallo, and R. Sayós, "Classical dynamics study of atomic oxygen over graphite (0001) with new interpolated and analytical potential energy surfaces", Computational and Theoretical Chemistry **990**, 132–143 (2012).

[189]J. Tian-Tian, S. Bao-Zhen, L. Hua-Xiang, L. Yi, and C. Wen-Kai, "Bonding of Hydroxyl and Epoxy Groups on Graphene: Insights from Density Functional Calculations", JIEGOU HUAXUE **32** (2013).

[190]D. Vasic Anicijevic, I. Perovic, S. Maslovara, S. Brkovic, D. Zugic, Z. Lausevic, and M. Marceta Kaninski, "Ab initio study of graphene interaction with O2, O and O-", Macedonian Journal of Chemistry and Chemical Engineering **35**, 271 (2016).

[191]F. Mehmood, R. Pachter, W. Lu, and J. J. Boeckl, "Adsorption and Diffusion of Oxygen on Single-Layer Graphene with Topological Defects", The Journal of Physical Chemistry C **117**, 10366–10374 (2013).

[192]D. C. Sorescu, K. D. Jordan, and P. Avouris, "Theoretical Study of Oxygen Adsorption on Graphite and the (8,0) Single-walled Carbon Nanotube", The Journal of Physical Chemistry B **105**, 11227–11232 (2001).

[193]H. Kinoshita, M. Umeno, M. Tagawa, and N. Ohmae, "Hyperthermal atomic oxygen beam-induced etching of HOPG (0001) studied by X-ray photoelectron spectroscopy and scanning tunneling microscopy", Surface Science **440**, 49–59 (1999).

[194]R. Larciprete, S. Fabris, T. Sun, P. Lacovig, A. Baraldi, and S. Lizzit, "Dual Path Mechanism in the Thermal Reduction of Graphene Oxide", J. Am. Chem. Soc. **133**, 17315–17321 (2011).

[195]R. Larciprete, P. Lacovig, S. Gardonio, A. Baraldi, and S. Lizzit, "Atomic Oxygen on Graphite: Chemical Characterization and Thermal Reduction", The Journal of Physical Chemistry C **116**, 9900–9908 (2012).

[196]J. T. Paci, H. P. Upadhyaya, J. Zhang, G. C. Schatz, and T. K. Minton, "Theoretical and Experimental Studies of the Reactions between Hyperthermal O(^{3}P) and Graphite: Graphene-Based Direct Dynamics and Beam-Surface Scattering Approaches", The Journal of Physical Chemistry A **113**, 4677–4685 (2009).

[197]W. E. Koerner and F. Daniels, "The heat of formation of nitric oxide and phosphorus trioxide", The Journal of Chemical Physics **20**, 113–115 (1952).

[198]R. D. Bach and H. B. Schlegel, "The Bond Dissociation Energy of the N–O Bond", The Journal of Physical Chemistry A **125**, 5014–5021 (2021).

[199]J. A. Montgomery, M. J. Frisch, J. W. Ochterski, and G. A. Petersson, "A complete basis set model chemistry. VII. Use of the minimum population localization method", The Journal of Chemical Physics **112**, 6532–6542 (2000).

[200]J. W. Ochterski, G. A. Petersson, and J. A. Montgomery, "A complete basis set model chemistry. V. Extensions to six or more heavy atoms", The Journal of Chemical Physics **104**, 2598–2619 (1996).

[201]L. A. Curtiss, P. C. Redfern, and K. Raghavachari, "Gaussian-4 theory", The Journal of Chemical Physics **126**, 84108 (2007).

[202]K. Huber and G. Herzberg, *Constants of diatomic molecules. In: Molecular Spectra and Molecular Structure.* (Springer, Boston, MA, 1979), pp. 466–480.

[203]J. Demaison, J. Liévin, and A. Perrin, "Equilibrium structure of nitrogen dioxide revisited", Molecular Physics **122**, e2318016 (2024).

[204]H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. DiNola, and J. R. Haak, "Molecular dynamics with coupling to an external bath", The Journal of Chemical Physics **81**, 3684–3690 (1984).

[205]H. C. Andersen, "Molecular dynamics simulations at constant pressure and/or temperature", The Journal of Chemical Physics **72**, 2384–2393 (1980).

[206]W. G. Hoover, "Canonical dynamics: Equilibrium phase-space distributions", Phys. Rev. A **31**, 1695–1697 (1985).

[207]S. Nosé, "A unified formulation of the constant temperature molecular dynamics methods", J. Chem. Phys. **81**, 511–519 (1984).

[208]G. G. Balint-Kurti, R. N. Dixon, and C. C. Marston, "Grid methods for solving the schrödinger equation and time dependent quantum dynamics of molecular photofragmentation and reactive scattering processes", International Reviews in Physical Chemistry **11**, 317–344 (1992).

[209]S. Kallush and R. Kosloff, "Improved methods for mapped grids: applied to highly excited vibrational states of diatomic molecules", Chemical Physics Letters **433**, 221–227 (2006).

[210]J. Häger, M. Fink, and H. Walther, "Scattering of NO from a graphite surface at cryogenic temperatures", Surface Science **550**, 35–45 (2004).

[211]T. Greenwood and S. P. K. Koehler, "Nitric Oxide Scattering off Graphene Using Surface-Velocity Map Imaging", The Journal of Physical Chemistry C **125**, 17853–17860 (2021).

[212]T. Greenwood and S. P. K. Koehler, "Molecular Dynamics Simulations of Nitric Oxide Scattering Off Graphene", ChemPhysChem **23**, e202200216 (2022).

[213]G. D. Billing, *Dynamics of Molecule Surface Interaction* (John Wiley & Sons, 2000).

[214]W. H. Miller, "Classical-limit quantum mechanics and the theory of molecular collisions", in *Advances in Chemical Physics*, Vol. 25, edited by I. Prigogine and S. A. Rice, 1st ed. (Wiley, Jan. 1974), pp. 69–177.

[215]G. D. Billing, "The dynamics of molecule-surface interaction", Computer Physics Reports **12**, 383–450 (1990).

[216]M. L. González-Martínez, L. Bonnet, P. Larrégaray, and J.-C. Rayez, "Classical treatment of molecular collisions: striking improvement of the description of recoil energy distributions using gaussian weighted trajectories", The Journal of Chemical Physics **126**, 41102 (2007).

[217]R. K. Pingak, A. Z. Johannes, F. Nitti, and M. Z. Ndii, "A theoretical study on vibrational energies of molecular hydrogen and its isotopes using a semi-classical approximation", Indonesian Journal of Chemistry **21**, 725 (2021).

[218]T. Sun, S. Fabris, and S. Baroni, "Surface Precursors and Reaction Mechanisms for the Thermal Reduction of Graphene Basal Surfaces Oxidized by Atomic Oxygen", J. Phys. Chem. C **115**, 4730–4737 (2011).

[219]L. Xiaowei, R. Jean-Charles, and Y. Suyuan, "Effect of temperature on graphite oxidation behavior", Nuclear Engineering and Design **227**, 273–280 (2004).

[220]R. Schafer, "What is a savitzky-golay filter? [lecture notes]", IEEE Signal Processing Magazine **28**, 111–117 (2011).

[221]J. Hager, Y. R. Shen, and H. Walther, "State-selective velocity and angular distributions of no molecules scattered from a graphite surface", Physical Review A **31**, 1962–1964 (1985).

[222]S. A. Dastgheib, H. Salih, T. Ilangovan, and J. Mock, "NO Oxidation by Activated Carbon Catalysts: Impact of Carbon Characteristics, Pressure, and the Presence of Water", ACS Omega **5**, 21172–21180 (2020).

[223]Z.-y. Guo, X.-c. Liu, H.-d. Huang, and P.-y. Yang, "Catalytic oxidation of no at low concentrations by carbon nanofibers near room temperature", New Carbon Materials **36**, 401–408 (2021).

[224]G. Nyman, L. Holmlid, and J. B. C. Pettersson, "Surface scattering of NO from graphite: A statistical description of energy distributions", The Journal of Chemical Physics **93**, 845–853 (1990).

[225]T. Greenwood and S. P. K. Koehler, "Molecular dynamics simulations of nitric oxide scattering off graphene", ChemPhysChem **23**, e202200216 (2022).

[226]Y. Zhang, C. Hu, and B. Jiang, "Embedded atom neural network potentials: efficient and accurate machine learning with a physically inspired representation", The Journal of Physical Chemistry Letters **10**, 4962–4967 (2019).

[227]B. Jiang, J. Li, and H. Guo, "High-fidelity potential energy surfaces for gas-phase and gas–surface scattering processes from machine learning", The Journal of Physical Chemistry Letters **11**, 5120–5131 (2020).

[228]A. Rivero Santamaría, M. Ramos, M. Alducin, H. F. Busnengo, R. Díez Muiño, and J. I. Juaristi, "High-Dimensional Atomistic Neural Network Potential to Study the Alignment-Resolved O_2 Scattering from Highly Oriented Pyrolytic Graphite", The Journal of Physical Chemistry A **125**, 2588–2600 (2021).

[229]A. Omranpour, J. Elsner, K. N. Lausch, and J. Behler, "Machine learning potentials for heterogeneous catalysis", ACS Catalysis **15**, 1616–1634 (2025).

[230]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 **104**, 136403 (2010).

[231]S. Chmiela, H. E. Sauceda, K.-R. Müller, and A. Tkatchenko, "Towards exact molecular dynamics simulations with machine-learned force fields", Nature Communications **9**, 3887 (2018).

[232]H. E. Sauceda, S. Chmiela, I. Poltavsky, K.-R. Müller, and A. Tkatchenko, "Molecular force fields with gradient-domain machine learning: construction and application to dynamics of small molecules with coupled cluster forces", The Journal of Chemical Physics **150**, 114102 (2019).

[233]N. Artrith, A. Urban, and G. Ceder, "Efficient and accurate machine-learning interpolation of atomic energies in compositions with many species", Physical Review B **96**, 014112 (2017).

[234]J. Zeng, D. Zhang, D. Lu, P. Mo, Z. Li, Y. Chen, M. Rynik, L. Huang, Z. Li, S. Shi, Y. Wang, H. Ye, P. Tuo, J. Yang, Y. Ding, Y. Li, D. Tisi, Q. Zeng, H. Bao, Y. Xia, J. Huang, K. Muraoka, Y. Wang, J. Chang, F. Yuan, S. L. Bore, C. Cai, Y. Lin, B. Wang, J. Xu, J.-X. Zhu, C. Luo, Y. Zhang, R. E. A. Goodall, W. Liang, A. K. Singh, S. Yao, J. Zhang, R. Wentzcovitch, J. Han, J. Liu, W. Jia, D. M. York, W. E, R. Car, L. Zhang, and H. Wang, "DeePMD-kit v2: a software package for deep potential models", The Journal of Chemical Physics **159**, 54801 (2023).

[235]J. Behler, "Four generations of high-dimensional neural network potentials", Chemical Reviews **121**, 10037–10072 (2021).

[236]O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko, and K.-R. Müller, "Machine learning force fields", Chemical Reviews **121**, 10142–10186 (2021).

[237]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 **103**, 4129–4137 (1995).

[238]A. M. Miksch, T. Morawietz, J. Kästner, A. Urban, and N. Artrith, "Strategies for the construction of machine-learning potentials for accurate and efficient atomic-scale simulations", Machine Learning: Science and Technology **2**, 031001 (2021).

[239]A. M. Cooper, J. Kästner, A. Urban, and N. Artrith, "Efficient training of ANN potentials by including atomic forces via Taylor expansion and application to water and a transition-metal oxide", npj Computational Materials **6**, 54 (2020).

[240]A. P. Bartók, R. Kondor, and G. Csányi, "On representing chemical environments", Physical Review B **87**, 184115 (2013).

[241]A. Bochkarev, Y. Lysogorskiy, and R. Drautz, "Graph atomic cluster expansion for semilocal interactions beyond equivariant message passing", Physical Review X **14**, 21036 (2024).

[242]K. T. Schütt, F. Arbabzadah, S. Chmiela, K. R. Müller, and A. Tkatchenko, "Quantum-chemical insights from deep tensor neural networks", Nature Communications **8**, 13890 (2017).

[243]S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, and B. Kozinsky, "E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials", Nature Communications **13**, 2453 (2022).

[244]C. Chen, W. Ye, Y. Zuo, C. Zheng, and S. P. Ong, "Graph networks as a universal machine learning framework for molecules and crystals", Chemistry of Materials **31**, 3564–3572 (2019).

[245]C. Chen and S. P. Ong, "AtomSets as a hierarchical transfer learning framework for small and large materials datasets", npj Computational Materials **7**, 173 (2021).

[246]R. Jacobs, D. Morgan, S. Attarian, J. Meng, C. Shen, Z. Wu, C. Y. Xie, J. H. Yang, N. Artrith, B. Blaiszik, G. Ceder, K. Choudhary, G. Csanyi, E. D. Cubuk, B. Deng, R. Drautz, X. Fu, J. Godwin, V. Honavar, O. Isayev, A. Johansson, B. Kozinsky, S. Martiniani, S. P. Ong, I. Poltavsky, K. Schmidt, S. Takamoto, A. P. Thompson, J. Westermayr, and B. M. Wood, "A practical guide to machine learning interatomic potentials – Status and future", Current Opinion in Solid State and Materials Science **35**, 101214 (2025).

[247]A. M. Tokita and J. Behler, "How to train a neural network potential", The Journal of Chemical Physics **159**, 121501 (2023).

[248]J. Behler, "Constructing high-dimensional neural network potentials: A tutorial review", International Journal of Quantum Chemistry **115**, 1032–1050 (2015).

[249]A. Singraber, T. Morawietz, J. Behler, and C. Dellago, "Parallel multistream training of high-dimensional neural network potentials", Journal of Chemical Theory and Computation **15**, 3075–3092 (2019).

[250]A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. In 'T Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton, "LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales", Computer Physics Communications **271**, 108171 (2022).

[251]M. S. Chen, T. Morawietz, H. Mori, T. E. Markland, and N. Artrith, "AENET–LAMMPS and AENET–TINKER: Interfaces for accurate and efficient molecular dynamics simulations with machine learning potentials", The Journal of Chemical Physics **155**, 074801 (2021).

[252]A. Singraber, J. Behler, and C. Dellago, "Library-based *LAMMPS* implementation of high-dimensional neural network potentials", Journal of Chemical Theory and Computation **15**, 1827–1840 (2019).

[253]B. Kozinsky, A. Musaelian, A. Johansson, and S. Batzner, "Scaling the leading accuracy of deep equivariant models to biomolecular simulations of realistic size", in Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (Nov. 2023), pp. 1–12.

[254]L. Lagardère, L.-H. Jolly, F. Lipparini, F. Aviat, B. Stamm, Z. F. Jing, M. Harger, H. Torabifard, G. A. Cisneros, M. J. Schnieders, N. Gresh, Y. Maday, P. Y. Ren, J. W. Ponder, and J.-P. Piquemal, "Tinker-HP: a massively parallel molecular dynamics package for multiscale simulations of large complex systems with advanced point dipole polarizable force fields", Chemical Science **9**, 956–972 (2018).

[255]O. Adjoua, L. Lagardère, L.-H. Jolly, A. Durocher, T. Very, I. Dupays, Z. Wang, T. J. Inizan, F. Célerse, P. Ren, J. W. Ponder, and J.-P. Piquemal, "Tinker-HP: accelerating molecular dynamics simulations of large complex systems with advanced point dipole polarizable force fields using GPUs and multi-GPU systems", Journal of Chemical Theory and Computation **17**, 2034–2053 (2021).

[256]K. T. Schütt, P. Kessel, M. Gastegger, K. A. Nicoli, A. Tkatchenko, and K.-R. Müller, "SchNetPack: a deep learning toolbox for atomistic systems", Journal of Chemical Theory and Computation **15**, 448–455 (2019).

[257]A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T.

Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, and K. W. Jacobsen, "The atomic simulation environment—a python library for working with atoms", Journal of Physics: Condensed Matter **29**, 273002 (2017).

[258]X. Gao, F. Ramezanghorbani, O. Isayev, J. S. Smith, and A. E. Roitberg, "TorchANI: a free and open source PyTorch-based deep learning implementation of the ANI neural network potentials", Journal of Chemical Information and Modeling **60**, 3408–3415 (2020).

[259]N. Artrith and J. Behler, "High-dimensional neural network potentials for metal surfaces: a prototype study for copper", Physical Review B **85**, 45439 (2012).

[260]T. Lookman, P. V. Balachandran, D. Xue, and R. Yuan, "Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design", npj Computational Materials **5**, 21 (2019).

[261]R. Jinnouchi, K. Miwa, F. Karsai, G. Kresse, and R. Asahi, "On-the-fly active learning of interatomic potentials for large-scale atomistic simulations", The Journal of Physical Chemistry Letters **11**, 6946–6955 (2020).

[262]E. V. Podryabinkin and A. V. Shapeev, "Active learning of linearly parametrized interatomic potentials", Computational Materials Science **140**, 171–180 (2017).

[263]J. Vandermause, S. B. Torrisi, S. Batzner, Y. Xie, L. Sun, A. M. Kolpak, and B. Kozinsky, "On-the-fly active learning of interpretable bayesian force fields for atomistic rare events", npj Computational Materials **6**, 20 (2020).

[264]N. Bernstein, G. Csányi, and V. L. Deringer, "De novo exploration and self-guided learning of potential-energy surfaces", npj Computational Materials **5**, 99 (2019).

[265]V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti, and G. Csányi, "Gaussian process regression for materials and molecules", Chemical Reviews **121**, 10073–10141 (2021).

[266]J. S. Smith, B. Nebgen, N. Lubbers, O. Isayev, and A. E. Roitberg, "Less is more: sampling chemical space with active learning", The Journal of Chemical Physics **148**, 241733 (2018).

[267]C. Schran, K. Brezina, and O. Marsalek, "Committee neural network potentials control generalization errors and enable active learning", The Journal of Chemical Physics **153**, 104105 (2020).

[268]G. E. Box and G. C. Tiao, *Bayesian inference in statistical analysis*, 1st ed. (Wiley, Apr. 1992).

[269]M. Wen and E. B. Tadmor, "Uncertainty quantification in molecular simulations with dropout neural network potentials", npj Computational Materials **6**, 124 (2020).

[270]J. López-Zorrilla, X. M. Aretxabaleta, I. W. Yeu, I. Etxebarria, H. Manzano, and N. Artrith, "Ænet-PyTorch: A GPU-supported implementation for machine learning atomic potentials training", The Journal of Chemical Physics **158**, 164105 (2023).

[271]A. Singh and Y. Li, "Reliable machine learning potentials based on artificial neural network for graphene", Computational Materials Science **227**, 112272 (2023).

[272]I. Loshchilov and F. Hutter, "Decoupled weight decay regularization", arXiv, 1711.05101 (2017).

[273]A. Krogh and J. Hertz, "A simple weight decay can improve generalization", Advances in neural information processing systems **4** (1991).

[274]M. Allen and D. J. Tildesley, *Computer Simulation of Liquid* (Oxford University Press, 1987).

[275]A. Laio and M. Parrinello, "Escaping free-energy minima", Proceedings of the National Academy of Sciences **99**, 12562–12566 (2002).

[276]G. Torrie and J. Valleau, "Nonphysical sampling distributions in monte carlo free-energy estimation: umbrella sampling", Journal of Computational Physics **23**, 187–199 (1977).

[277]X. Chang, D. Zhang, Q. Chu, and D. Chen, "Minimizing redundancy and data requirements of machine learning potential: a case study in interface combustion", Journal of Chemical Theory and Computation **20**, 6813–6825 (2024).

[278]L. McInnes, J. Healy, and J. Melville, "UMAP: uniform manifold approximation and projection for dimension reduction", arXiv, 1802.03426 (2018).

From Ab Initio Molecular Dynamics to Machine Learning Potential Energy Surface-Based Simulations: NO Oxidation on Oxygen-Functionalized Graphite

Abstract

Oxygen-functionalized graphite surfaces (O@HOPG) are commonly found in the atmosphere due to exposure to oxidizing pollutants such as ozone (O_3) and hydrogen peroxide (H_2O_2). These surface-bound oxygen groups play a crucial role in atmospheric chemistry, as they can react with a broad range of pollutants, including hydrocarbons and nitrogen oxides (NO_x). Among these, NO_x are major contributors to smog and acid rain; nitric oxide (NO) is a precursor to nitrogen dioxide (NO_2), which also promotes the formation of secondary organic aerosols—known to negatively impact both human health and the environment.

This thesis presents a theoretical investigation of the oxidation of NO molecules on epoxy-functionalized HOPG, with relevance to heterogeneous atmospheric chemistry and pollution abatement. The study relies on an approach combining Density Functional Theory (DFT)-based ab initio Molecular Dynamics (AIMD) with a machine learning potential energy surface (ML-PES). Four collision energies (0.025, 0.05, 0.1, and 0.3 eV) and two different orientations were analyzed where the reaction, adsorption, and scattering probabilities were computed.

AIMD simulations reveal that NO_2 formation can occur even at the lowest collision energy investigated (0.025 eV), approximately equivalent to room temperature (300 K), which agrees qualitatively with the experimental results. Additionally, simulations show that scattered NO molecules exhibit low specular reflection, lose half of their initial translational energy, and remain vibrationally cold with minimal rotational excitation. Furthermore, a statistical analysis of all reactive trajectories, elucidated the structural requirements for the reaction to efficiently occur. Finally, this theoretical study demonstrates the potential of oxygen-doped carbon surfaces for the conversion of NO to NO_2 highlighting the importance of explicitly accounting for dynamical effects on the NO oxidation.

To circumvent the high computational cost of AIMD simulations and the limited statistical sampling, a ML-PES was constructed and trained on configurations extracted from AIMD trajectories. The accuracy of the ML-PES was rigorously validated by direct comparison with AIMD results, confirming its ability to reproduce both reactive and non-reactive events with high fidelity. The ML-PES successfully captures the essential dynamical and energetic features of the NO + O@HOPG $\longrightarrow$ NO_2 + HOPG reaction, accurately identifying reaction channels and characterizing energy transfer mechanisms, while reducing the computational cost by approximately two orders of magnitude compared to conventional AIMD simulations. The high efficiency of ML-PES simulation makes it possible to enhance the sampling of reactive events, thereby providing statistically robust insights into the underlying mechanisms of NO oxidation. Nevertheless, limitations persist in regions of the configurational space that are underrepresented in the training data. To overcome this, an active learning strategy is proposed to iteratively refine the ML-PES by incorporating dynamically important configurations, thereby improving the model's robustness and transferability.

This work establishes a robust computational framework that combines AIMD and ML-PES for the atomistic simulation of gas–surface reactions. While this thesis concentrated on the NO oxidation on O@HOPG, the developed methodology is broadly transferable and can be extended to study other gas-surface reactions of atmospheric relevance within the context of depollution or catalysis.

Keywords: ab initio molecular dynamics, machine learning potentials, gas–surface reactions, atmospheric chemistry, graphite, nitric oxide, dft

DES SIMULATIONS DE DYNAMIQUE MOLÉCULAIRE AB INITIO AUX SURFACES D'ÉNERGIE POTENTIELLE BASÉES SUR L'APPRENTISSAGE AUTOMATIQUE : OXYDATION DE NO SUR UNE SURFACE DE GRAPHITE OXYDÉE

Résumé

Suite aux émissions anthropogéniques, les particules de carbone sont abondamment présentes dans l'atmosphère et particulièrement sous forme oxydées par des oxydants tels que l'ozone (O_3) et le peroxyde d'hydrogène (H_2O_2). Les groupes oxygénés en surface jouent un rôle clé, réagissant avec divers polluants, notamment les hydrocarbures et les oxydes d'azote (NO_x). Ces derniers sont des précurseurs majeurs du smog et des pluies acides ; le monoxyde d'azote (NO), en particulier, se transforme en dioxyde d'azote (NO_2), favorisant la formation d'aérosols organiques secondaires, nocifs pour la santé humaine et l'environnement.

Cette thèse présente une étude théorique de l'oxydation des molécules de NO sur du graphite pyrolytique hautement orienté (HOPG) fonctionnalisé par des groupes époxy. L'étude repose sur une approche combinant la dynamique moléculaire ab initio (AIMD) basée sur la théorie de la fonctionnelle de la densité (DFT) et une surface d'énergie potentielle obtenue par apprentissage automatique (ML-PES). Les probabilités de réaction, d'adsorption et de diffusion ont été caractérisées pour quatre énergies de collision et deux orientations différentes. Les simulations AIMD révèlent que la collision réactive, conduisant à la formation de NO_2, peut se produire même à la plus faible énergie de collision étudiée (0.025 eV), correspondant approximativement à la température ambiante (300 K), ce qui s'accorde qualitativement avec les résultats expérimentaux. Ce travail met en évidence l'influence des effets dynamiques sur le processus d'oxydation du NO à la surface de graphite activée. Les distributions angulaires et énergétiques des produits révèlent que Les molécules de NO diffusées présentent une faible réflexion spéculaire, perdent la moitié de leur énergie translationnelle initiale et restent vibrationnellement froides, avec une excitation rotationnelle minimale. En outre, une analyse statistique de toutes les trajectoires réactives, a permis d'élucider les exigences structurales pour que la réaction ait lieu dans des conditions dynamiques. Cependant, le coût numérique élevé des simulations AIMD limite l'échantillonnage statistique des événements réactifs rares. Pour pallier à cette limitation, une ML-PES a été construite et entraînée sur des configurations extraites des trajectoires AIMD. L'exactitude de la ML-PES a été rigoureusement validée par une comparaison directe avec les résultats AIMD, confirmant sa capacité à reproduire avec une grande fidélité aussi bien les événements réactifs que non réactifs. La ML-PES capture avec succès les caractéristiques dynamiques et énergétiques essentielles de la réaction NO + O@HOPG $\longrightarrow$ NO_2 + HOPG, identifiant précisément les canaux réactionnels et caractérisant les mécanismes de transfert d'énergie, tout en réduisant le coût computationnel d'environ deux ordres de grandeur par rapport aux simulations AIMD conventionnelles. Cette réduction spectaculaire du coût numérique rend possible l'exploration d'un nombre bien plus important d'événements réactifs, offrant ainsi une vision statistiquement robuste des mécanismes sous-jacents de l'oxydation du NO. Néanmoins, certaines régions de l'espace des configurations restent sous-échantillonnées dans les données d'entraînement. Pour y remédier, une stratégie d'apprentissage actif est proposée afin d'affiner itérativement la ML-PES en intégrant des configurations dynamiquement importantes, améliorant ainsi la robustesse et la transférabilité du modèle. Ce travail établit une méthodologie solide combinant AIMD et ML-PES pour la simulation atomique des processus gaz–surface. Bien que cette thèse se concentre sur l'oxydation du NO sur les surfaces O@HOPG, les développements proposés sont largement transférables. En particulier, l'approche peut être facilement étendue à l'étude d'autres réactions avec des surfaces carbonées fonctionnalisées d'intérêt atmosphérique.

Mots clés : dynamique moléculaire ab initio, potentiels d'apprentissage automatique, réactions gaz–surface, chimie atmosphérique, graphite, monoxyde d'azote, dft