SCUOLA DI DOTTORATO
UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Dipartimento di / Department of
Materials Science

Dottorato di Ricerca in / PhD program Materials Science and Nanotechnology Ciclo / Cycle 38°

Curriculum in (se presente / if it is) _____

Machine learning methods for capturing the dynamic nature of heterogeneous catalysts under operando conditions

Cognome / Surname Tosello Gardini Nome / Name Axel

Matricola / Registration number 896514

Tutore / Tutor: Prof. Marco Bernasconi

Cotutore / Co-tutor: _____
(se presente / if there is one)

Supervisor: Prof. Michele Parrinello
(se presente / if there is one)

Coordinatore / Coordinator: Prof. Francesco Montalenti

ANNO ACCADEMICO / ACADEMIC YEAR 2024-2025

University of Milano-Bicocca
Department of Materials Science

Italian Institute of Technology
Atomistic Simulations

PhD program in Materials Science and Nanotechnology
· XXXVIII Cycle ·

Machine learning methods for capturing the dynamic nature of heterogeneous catalysts under operando conditions

by

Axel Tosello Gardini
Mat.896514

A dissertation submitted for the degree of
Dottore di Ricerca Ph.D.

Supervisor:
Prof. Michele Parrinello
Co-Supervisor:
Dr. Umberto Raucci
Tutor:
Prof. Marco Bernasconi
Coordinator:
Prof. Francesco Montalenti

ACADEMIC YEAR 2024/2025

Contents

Abstract

Computational models in heterogeneous catalysis have traditionally treated catalysts as static entities with well-defined active sites. While such an approach has led to important insights, it neglects the profound influence of *operando* conditions, such as temperature, pressure, and reactant flow typically encountered in industrial environments, on the catalyst structure and behavior. Increasing experimental evidence suggests that catalysts are inherently dynamic systems, whose activity is closely tied to their ability to undergo structural transformations under reaction conditions. This motivates the urgent need to rethink the computational approaches in catalysis.

With advances in computational power and the integration of machine learning algorithms, the field of computational catalysis is undergoing a paradigm shift. Simulations are now increasingly capable of capturing the behavior of catalysts under realistic and relevant *operando* conditions, providing unprecedented atomistic insights into their dynamic nature. Within this emerging view, catalysis must be regarded as a collective, dynamic, and non-local phenomenon, in which the entire system evolves in response to *operando* variables and actively participates in reactions.

In this thesis, we aim to deepen the understanding of catalytic processes by supporting the view that catalysis is driven by dynamics arising from *operando* conditions. To this end, we employ molecular dynamics simulations enhanced by machine learning interatomic potentials as main computational tool. In particular, we present three case studies in which the dynamics of the catalysts, namely their ability to undergo continuous structural rearrangements, plays a central role in enabling or enhancing reactivity.

The first case is the ammonia synthesis catalyzed by barium hydride (BaH_2) within a chemical looping process. Under operando conditions, BaH_2 undergoes a dramatic chemical transformation that is not confined to the catalyst surface but involves the entire bulk material, indicating that the whole solid participates actively in the catalytic process.

As a second case of study, we investigate the effect of the electrochemical potential on solvent dynamics and its implications for reaction mechanisms in CO_2 electroreduction on the $Ag(111)/H_2O$ interface. Our simulations reveal that the applied electrochemical potential induces a strong reorganization of the interfacial water layer, modifying its orientation and hydrogen-bonding network. This collective solvent response plays an active role in stabilizing key intermediates, facilitating proton transfer, and regulating reaction kinetics.

Lastly, we explore ammonia synthesis in lithium-sodium alloys, a system that we proposed for its intrinsic dynamics as a liquid metal catalyst. The highly fluxional interface between the two metallic components gives rise to a microheterogeneous environment, where continuous interfacial rearrangements and synergy between the two components drive catalytic activity.

Through these studies, we highlight how *operando* conditions shape the structure and dynamics of catalysts, ultimately driving their reactivity. These features remain inaccessible to traditional static models. Only by adopting a dynamic, microscopic view of catalysis, made possible by advances in machine learning-based molecular dynamics simulations, we can deepen our understanding of catalytic process and open new opportunities for the rational design of more efficient catalysts for industrial applications.

Riassunto

I modelli computazionali nella catalisi eterogenea hanno tradizionalmente trattato i catalizzatori come entità statiche con siti attivi ben definiti. Sebbene questo approccio abbia fornito importanti contributi alla comprensione dei meccanismi catalitici, si trascura l'influenza delle condizioni operative, come temperatura, pressione e flusso dei reagenti tipicamente usate in ambito industriale, sulla struttura e sul comportamento dei catalizzatori. Evidenze sperimentali sempre più numerose suggeriscono che i catalizzatori siano sistemi intrinsecamente dinamici, la cui attività è strettamente legata alla loro capacità di subire trasformazioni strutturali in condizioni di reazione. Ciò motiva l'urgente necessità di ripensare gli approcci computazionali alla catalisi.
Con l'aumento della potenza computazionale e l'integrazione di algoritmi di *machine learning*, il campo della catalisi computazionale sta attraversando un cambiamento di paradigma. Le simulazioni sono oggi sempre più in grado di catturare il comportamento dei catalizzatori in condizioni operative realistiche e rilevanti, offrendo una comprensione atomistica senza precedenti della loro natura dinamica. In questa nuova prospettiva, la catalisi deve essere considerata come un fenomeno collettivo, dinamico e non locale, in cui l'intero sistema evolve in risposta alle variabili operative e partecipa attivamente al processo reattivo.
In questa tesi, si intende contribuire alla comprensione dei processi catalitici sostenendo l'idea che la catalisi sia guidata dalla dinamica che emerge proprio dalle condizioni di reazione. A tal fine, viene usato come principale strumento computazionale la dinamica molecolare, potenziata da potenziali interatomici basati su *machine learning*.
Vengono presentati in particolare tre casi di studio in cui la dinamica dei catalizzatori, ovvero la loro capacità di subire riorganizzazioni strutturali continue, riveste un ruolo centrale nel permettere o potenziare la reattività.
Il primo caso riguarda la sintesi dell'ammoniaca catalizzata dall'idruro di bario (BaH_2) all'interno di un processo di *chemical looping*. In condizioni operative, BaH_2 subisce una trasformazione chimica drastica che non è confinata alla superficie, ma coinvolge l'intero materiale, indicando che tutto il solido partecipa attivamente al processo catalitico.
Il secondo caso di studio analizza l'effetto del potenziale elettrochimico sulla dinamica del solvente e le sue implicazioni sui meccanismi di reazione nella riduzione elettrochimica della CO_2 all'interfaccia $Ag(111)/H_2O$. Le simulazioni mostrano che il potenziale applicato induce una forte riorganizzazione dello strato d'acqua interfacciale, modificandone l'orientazione e la rete di legami a idrogeno. Questa risposta collettiva del solvente gioca un ruolo attivo nella stabilizzazione degli intermedi, nel facilitare il trasferimento di protoni e nel modulare la cinetica di reazione.
Infine, viene esplorata la sintesi dell'ammoniaca in leghe litio-sodio, un sistema che viene pro-

posto come catalizzatore liquido grazie alla sua dinamica intrinseca. L'interfaccia altamente flussionale tra i due componenti metallici genera un ambiente microeterogeneo in cui continue riorganizzazioni interfacciali e la sinergia tra i due componenti guidano l'attività catalitica.
Attraverso questi studi, si vuole mettere in evidenza come le condizioni operative plasmino la struttura e la dinamica dei catalizzatori, determinandone la reattività. Comportamenti di questo tipo non possono essere catturati dai modelli statici tradizionali. Solo adottando una visione microscopica e dinamica della catalisi, resa possibile dai recenti progressi nella dinamica molecolare basata su *machine learning*, è possibile approfondire la comprensione dei processi catalitici e aprire nuove opportunità per la progettazione razionale di materiali catalitici più efficienti per applicazioni industriali.

Chapter 1

Introduction

The term *catalysis* was first introduced in 1835 by the Swedish chemist Jöns Jacob Berzelius [1], yet it took several decades before Wilhelm Ostwald offered a functional definition in 1894: "Catalysis is the acceleration of a slow chemical process by the presence of a foreign material" [2]. This simple description laid the groundwork for a field that has since evolved into a central pillar of modern chemistry.

The *material* in Ostwald's definition refers to what we call a *catalyst*, and there is a general consensus that the catalyst is not consumed during the reaction. But does this mean it remains unchanged throughout the process?

Catalysts do respond dynamically to changes in the chemical potential of their environment. Many examples can be found in homogeneous catalysis, where the catalyst is dissolved in the same phase as the reactants. In reactions such as Suzuki, Stille, and Heck C–C coupling processes, a precatalyst is used. This precatalyst undergoes *in situ* transformation under reaction conditions, generating the actual active species responsible for catalysis. Furthermore, the performance of the active species depends on various factors such as the solvent, pH, temperature, and the nature of the ligands, all of which can significantly influence both the reaction yield and the products distribution.

In heterogeneous catalysis, the process takes place at the interface between a solid catalyst and a liquid or gaseous reactant phase. In many relevant industrial processes, such Haber-Bosch, Fischer-Tropsch or dry reforming, the catalyst operates under harsh working conditions. Also in this case, the catalyst often is activated only under reaction conditions. This raises a number of questions: Are catalytic events uniformly distributed across the active catalyst surface, or are they instead localized at specific regions (*active sites*)? Are these regions well-defined atomic ensembles, or do they emerge from local disorder, defects, or dynamic surface restructuring? To what extent are the catalyst transformations confined to the surface, or do they involve subsurface layers and even the bulk material? How can the catalyst sustain its activity over an extended lifetime under such working conditions? And many more questions.

Answering these questions is not merely an academic exercise; it has concrete practical implications. Since catalytic processes are essential for the production of countless materials and chemicals, understanding how catalysts actually work paves the way for improving their performance and discovering better alternatives that meet the evolving needs of society [3, 4]. To achieve this,

combined efforts from both experimentalists and theoreticians are being made.

A comprehensive understanding of catalysis requires analysis across multiple spatial and temporal scales. Catalytic phenomena arise from complex, multiscale interactions, ranging from atomic level events, such as bond formation and cleavage, to larger scale processes including surface restructuring, diffusion, and phase transitions. Bridging these scales presents two fundamental challenges: first, capturing the elementary steps that constitute the catalytic mechanism; and second, recognizing that catalytic activity does not stem from a single, static configuration, but rather from a distribution of dynamic, condition-dependent structural motifs. These complexities blur the one-to-one relationship between atomic structure and catalytic function, necessitating statistical and systems-level approaches.

This multiscale complexity also invites a fundamental question: *what is an active site*?

Broadly speaking, an active site can be defined as a structural entity that directly participates in the catalytic process and is typically present in extremely small amounts. This last point implies that, on the catalyst surface, there also exist spectator species that are present in large amounts but do not directly participate in the catalytic cycle. Early models of catalysis, such as Langmuir's, assumed that large portions of the surface participate uniformly in the reaction [5]. However, subsequent studies revealed that only a small fraction of the surface is truly active. In 1925, Taylor proposed that catalysis occurs at specific, under-coordinated atoms, often located at edges, corners, or defects [6]. These rare sites possess enhanced reactivity due to their low coordination and can adsorb multiple reactants simultaneously. Taylor's theory thus marked a turning point, introducing the idea that catalytic activity emerges from local heterogeneities [7].

Later on, Spencer emphasized how thermal and adsorbate-driven dynamics shape catalytic activity [8]. Already at approximately one-third of the melting temperature, known as the Hüttig temperature, metal surfaces can enter a dynamic equilibrium in which surface atoms become mobile, leaving their lattice positions and enabling local reconstructions and transformations. This atomic mobility allows the catalyst surface to continuously reorganize under reaction conditions. Spencer proposed that, for most supported metal catalysts, such atomic mobility is sufficient to reach a quasi-equilibrium structure. The ability of metal surfaces to reorganize dynamically ensures that catalytically relevant configurations are continuously regenerated and stabilized under working conditions. In contrast, when the system operates below the Hüttig temperature, or in the absence of strong adsorbate-induced reconstruction, surface atoms may remain kinetically trapped in out-of-equilibrium configurations. These are defined as metastable metal surfaces and can lead to significant variability in both activity and selectivity. Spencer's work underscores a key insight: catalytic performance emerges not only from the presence of local heterogeneities but also from their dynamic accessibility and structural evolution over time. This adds an important temporal dimension to the definition of active sites, suggesting that they are not static entities, but rather transient features stabilized by the interplay between temperature, adsorbates, and support effects [9, 10].

The ability of surface atoms to reorganize under thermal and chemical stimuli suggests that catalytic activity may not stem from a predefined structure, but instead emerges from a structural transition that takes place during the reaction. This leads to the broader concept of the *active phase*, a distinct structural or compositional state of the catalyst that forms under working con-

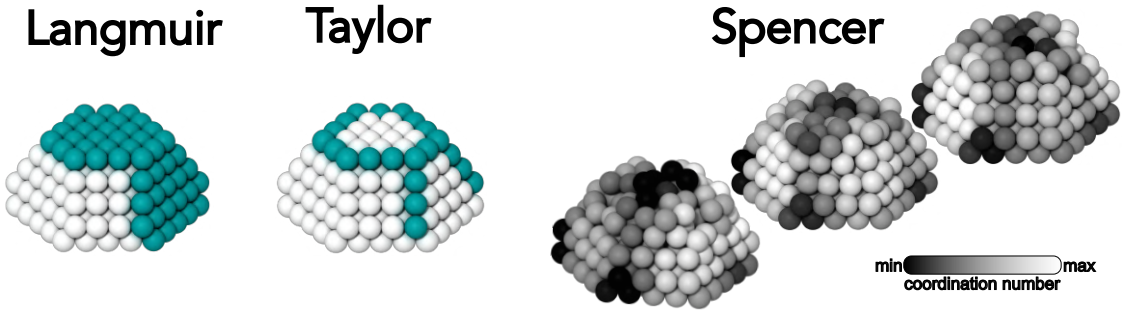

Figure 1.1: **Different models of catalysis.** Schematic representation of the Langmuir, Taylor and Spencer's models of catalytic surfaces.

ditions and is responsible for catalytic activity. As emphasized by Schlögl, the active phase is not pre-existing but is dynamically induced and sustained by the chemical potential of the environment [9].

Capturing the emergence of active phases and the evolution of catalytic sites under working conditions requires experimental techniques that go beyond traditional *ex situ* or static *in situ* characterization. In this context, *operando* spectroscopy has proven essential [10]. Techniques such as Raman, UV–Vis, X-ray absorption (XAS), and infrared spectroscopy, when applied under real reaction conditions with simultaneous catalytic performance monitoring, allow for the direct correlation between structural changes and catalytic activity in real time. These methods have revealed that the active phase is a transient state, dynamically induced by the reaction environment and often reversible once the stimulus is removed. Notable examples include the operando Raman spectroscopy study of the decomposition of chlorinated hydrocarbons over La_2O_3-based catalysts, where the formation of an active LaOCl phase was observed only under reaction conditions [11]. When the feed was switched back, the original La_2O_3 phase reappeared, demonstrating the reversible, condition-dependent nature of the active phase. Similarly, in the operando UV–Vis study of Cr/Al_2O_3 catalysts used for alkane dehydrogenation, the catalytically active species was identified as a pseudo-octahedral Cr^{3+} complex, which only formed under a specific temperature and gas composition [12, 13]. These findings highlight that active phases are not present in the fresh catalyst, but emerge dynamically during the reaction.
The field of operando spectroscopy has evolved significantly, with increasingly advanced techniques offering improved spatial and temporal resolution. Ideally, experimentalists aim to achieve a "molecular movie" that captures the structural evolution of catalysts in real time during exposure to reactants, but this remains an ambitious goal [14]. Identifying and characterizing these elusive active sites remains a major challenge in heterogeneous catalysis. A truly comprehensive understanding of catalytic function requires operando experiments that integrate multiple spectroscopic techniques, each probing different aspects of the system: from oxidation state and coordination environment to local geometry and electronic structure. When applied simultaneously under realistic reaction conditions, these tools can begin to track the evolving structure of the catalyst during operation [15, 16, 17]. However, the dynamic nature of active sites and the short

lifetimes of key intermediates often demand ultrafast, time-resolved measurements, ideally with sub-picosecond temporal resolution, to capture the fleeting configurations that drive reactivity. Compounding the difficulty, many of these reactive species are structurally ill-defined or completely unknown, making their unambiguous spectroscopic identification particularly difficult.
The inherent fluxionality of the active phase makes the quantification of active sites particularly elusive. After promoting a reaction event, an active site does not necessarily return to its catalytically active state instantaneously. Instead, it may require a finite reactivation time before it can contribute to catalysis again. This fundamentally complicates the definition of catalytic efficiency. Common kinetic descriptors such as the turnover frequency (TOF) rely on the assumption of a constant and fully accessible population of active sites, each ready to sustain continuous reaction cycles. In practice, however, the timescale associated with site regeneration can exceed the characteristic TOF [18]. As a result, TOF becomes a highly idealized metric, which may significantly overestimate the actual per-site activity [19]. The assumption of a steady-state ensemble of uniformly available sites breaks down when site accessibility fluctuates over time or when reactivation processes become rate-limiting. Therefore, comparing TOF values across different catalytic systems, or using them to assess performance, is only meaningful when the underlying structural and dynamical context is well understood. Without such information, TOF loses its mechanistic relevance and serves merely as a formal descriptor of ideal behavior. In this context, computational simulations emerge as a powerful and often indispensable complement to experimental methods.
The challenge of modeling such dynamic behavior is intrinsically linked to multiscale nature of catalysis, which aim to bridge the gap between atomistic and macroscopic phenomena.

Computational approaches have been historically instrumental in catalysis, providing atomic-scale insights into reaction mechanisms [20, 21]. However, despite their usefulness, traditional computational studies often rely on simplifying assumptions that limit their relevance and predictive accuracy under real-world industrial conditions [22]. One major approximation is the use of optimized geometries corresponding to local minima of the potential energy surface (PES). This oversimplification neglects the broad configurational variability typical of fluxional systems, as we have described. Secondly, these local minima are often treated using the harmonic approximation, where atomic vibrations as independent harmonic oscillators around equilibrium positions and models the PES as locally quadratic. However, these approximations are inadequate in industrial operando catalytic environments, where large-amplitude fluctuations, surface mobility, and dynamic restructuring are common. In such systems, especially at elevated temperatures, pressure or under confinement, atoms and adsorbates sample regions of the PES that are far from the local minima, with pronounced anharmonic behavior. As a result, thermodynamic predictions based on static representation of the PES are often unreliable. Furthermore, reaction mechanisms may involve not a single, well-defined transition state, but rather a broad ensemble of activated configurations. The free energy landscape at finite temperature can differ substantially from the potential energy surface derived at 0K, and reaction coordinates that appear dominant in static calculations may become irrelevant when entropic effects are accounted for [22].
To overcome these limitations, molecular dynamics has emerged as a powerful tool to study these systems at atomistic level [23]. Molecular dynamics allows to account for temperature-dependent

atomic motion, enabling treatment of anharmonicity and entropic effects. Furthermore, molecular dynamics can be combined with enhanced sampling techniques to access rare events, for example chemical reactions, or phase transitions. Such approaches also allow for the simultaneous exploration of multiple reaction pathways in systems where a high degree of configurational freedom is present.
A striking example of how molecular dynamics can capture the dynamic nature of active sites is provided by the work of Wang et al. who investigated ceria-supported gold nanoparticles under CO oxidation conditions [24]. Their study revealed that in the presence of CO, a single Au atom can transiently detach from the nanoparticle, migrate onto the ceria surface, and form a cationic $Au^{+}-CO$ species. This dynamic entity acts as the true catalytic center by facilitating CO oxidation through a redox mechanism involving lattice oxygen. After the reaction, the gold atom return to the nanoparticle, completing a fully reversible catalytic cycle. Importantly, this reactive species does not exist *a priori*, but is induced and stabilized by the chemical environment.
Beyond capturing atomic-scale fluctuations and dynamic restructuring of solid surfaces, molecular dynamics also enables the explicit simulation of solvents at solid–liquid interfaces. Notable examples are found in the literature, particularly in electrochemical systems and catalysis within confined materials, where solvent reorganization plays a critical role [25, 26, 27]. In these environments, entropic contributions arising from solvent dynamics can dominate the reaction free energy landscape, often favoring pathways that would not be anticipated from static 0K potential energy surfaces alone.

As discussed in chapter 3.1, the PES in molecular dynamics simulations can be modeled using different levels of approximation. In the works mentioned above, authors employed *ab initio* molecular dynamics (AIMD), where the PES is computed from first-principles electronic structure calculations. This approach accurately captures the evolution of the electronic structure during chemical reactions. However, the high computational cost imposes limitations on the system size ($\sim 10^2$ atoms) and accessible timescales ($\sim$ 100ps), making AIMD seemingly incompatible with the spatial and temporal resolutions required for realistic catalysis modeling. Recent developments in machine learning offer a promising path to overcome these limitations. Machine learning interatomic potentials (MLPs) are data-driven models trained on quantum mechanical reference calculations. By learning the mapping between atomic configurations and their corresponding energies and forces, MLPs enable efficient simulations with near *ab initio* accuracy, while reducing the computational cost by several orders of magnitude. In this framework, a representation of the PES is constructed via nonlinear regression techniques, such as neural networks or Gaussian process models, based on a representative dataset of quantum-mechanical configurations. Once trained, these MLPs can be used in large-scale molecular dynamics simulations to explore extended timescales and system sizes that are otherwise inaccessible to AIMD.
Thanks to these methodological advances, notable progress has been made in the study of heterogeneous catalysis. Several examples are presented in chapter 2 of this thesis. While not intended to be exhaustive, these case studies underscore the central idea that catalysis is not a static or localized phenomenon, but rather a dynamic, collective, and non-local process. Understanding these complex behaviors requires the adoption of advanced computational methodologies capable of going beyond static, 0K energy landscapes. The theoretical background underpinning

these methodologies are described in chapter 3 and form the basis of the works presented in the following chapters. In chapter 4, we present simulations of the reversible phase transformation of BaH_2 during ammonia synthesis via chemical looping process [28]. In chapter 5, MLPs are used to investigate the CO_2 electroreduction reaction on the Ag(111)/H_2O at two potential conditions [29]. In chapter 6, we present simulations of ammonia synthesis catalyzed by a liquid Li—Na alloy. The system features a spontaneously formed, dynamic liquid–liquid interface where nitrogen and hydrogen are activated and converted into ammonia [30].
These results further support the view that heterogeneous catalysis must be regarded as a dynamic, collective, and highly nonlocal phenomenon, where the entire catalytic system actively participates in driving the chemical transformations.

Chapter 2

Examples of dynamic behaviors from surfaces to liquid catalysts

In this chapter, we review selected case studies in which molecular dynamics (MD), both AIMD and MLPs-driven MD, has been used to elucidate the dynamic mechanisms of catalysis in ammonia synthesis and decomposition, electrocatalysis, and liquid-phase catalysis. These examples are examined because they directly relate to the results presented in chapters 4, 5, and 6.

2.1 Surface dynamics during N_2 adsorption in ammonia synthesis

Ammonia (NH_3) has long been recognized as one of the most essential industrial chemicals. Globally, approximately 150 million tonnes of ammonia are produced each year, and production is expected to grow at an annual rate of about 2.3%. However, the main challenge associated with ammonia lies in the way it is produced. Over 96% of global ammonia is synthesized through the Haber–Bosch process [31, 32]. Traditionally, this process uses a iron based metal catalyst and is highly energy-intensive as it requires extreme temperature (350-550°C) and pressure (140-320 atm) conditions to convert molecular nitrogen and hydrogen into ammonia:

$$N_2 + 3\,H_2 \longrightarrow 2\,NH_3$$

This results in a substantial carbon footprint: on average, conventional ammonia production emits approximately 1.5 to 1.6 tonnes of CO_2 per tonne of NH_3 synthesized and accounts for ∼1.2% of global anthropogenic CO_2 emissions [31, 32].
In this context, understanding the fundamental mechanisms that guide ammonia synthesis is of paramount importance for designing more sustainable catalytic routes.
The mechanistic picture of N_2 adsorption on Fe originates from pioneering studies by Gerhard Ertl and co-workers. Vibrational spectroscopy showed that N_2 binds to Fe(111) in two limiting orientations: end-on (approximately perpendicular) and side-on (approximately parallel). The end-on states (γ and δ) are weakly bound, whereas the side-on α state displays a strongly red-shifted N−N stretch, marking substantial bond activation and identifying a molecular precursor to dissociation (Fig. 2.1) [33, 34, 35]. Early electronic-structure work rationalized these observations:

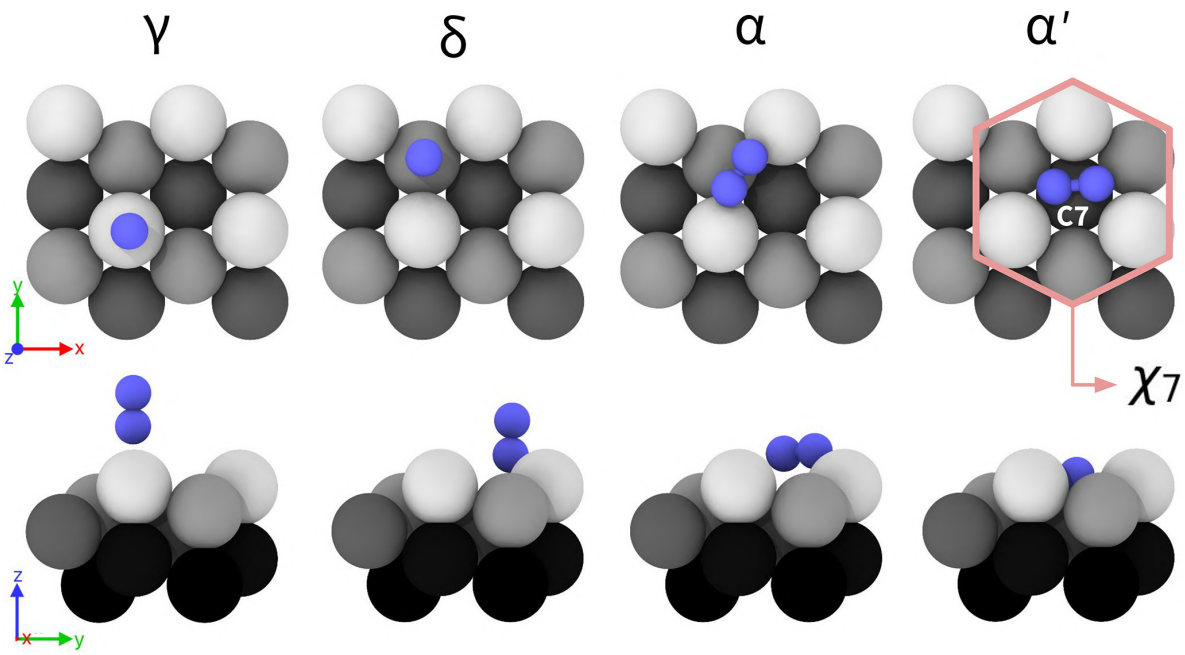

Figure 2.1: **N_2 adsorption sites on Fe(111) from 0K calculations**. Top and side views. Fe are colored according to their z-coordinate. Reprinted with permission from ref. [38]. Copyright 2025 National Academy of Sciences. Courtesy of Luigi Bonati.

N_2-Fe bonding involves σ donation from N_2 to the metal and π-backdonation from Fe d states. This results in a weakening of the N≡N triple bond and easier dissociation [36]. Subsequent DFT calculations by Mortensen et al. provided an atomistic picture of the nitrogen adsorption and dissociation process on Fe(111) [37]. The work identified five distinct adsorption states, including three previously known molecular states (γ, δ and α) and a newly proposed intermediate, the α' state. The latter is located above a third-layer of atom where the N_2 molecule is coordinated by a seven-fold Fe environment, denoted as $\chi 7$. This state was identified as the immediate precursor to dissociation in the multistep mechanism $\gamma \rightarrow \delta \rightarrow \alpha \rightarrow \alpha' \rightarrow \beta$, where β denotes the dissociated atomic state (Fig. 2.1). However, most of the evidence above comes from low temperature and high vacuum measurements or 0K static calculations. Under industrial conditions, entropy reshapes the catalyst picture and reliable mechanistic insight therefore requires approaches that explicitly capture temperature and pressure effects. Recent simulations at *operando* conditions indeed reveal a dynamic, ensemble-based reactivity in which "active sites" are not fixed motifs.

In the paper by Bonati et al. [38] the authors revisited the dissociation mechanism of molecular N_2 on the Fe(111) surface. Using MLP molecular dynamics simulations, the authors explicitly resolved how the Fe(111) surface dynamics evolves with temperature.
At 300K, the Fe(111) surface retains a stepped, rigid structure, and N_2 adsorption and dissociation follow the multi-step pathway involving well-defined precursor states (γ, δ, α and α'), consistent with the low-T picture. These steps involve gradual weakening of the N≡N bond through electron

donation from the surface to the molecule, ending in the α' state where dissociation occurs within a seven-fold coordinated iron site ($\chi 7$).
However, as temperature increases beyond 500K the system enters in the Hütting regime, where the atomic mobility on the surface increases significantly below the melting point. The surface becomes highly mobile and undergoes pronounced structural fluctuations while retaining local bcc order. At 700K, the ordered step structure destabilizes, and a dynamic adlayer forms, with atoms migrating between layers and the surface exhibiting characteristics of a fluxional interface (see video in [1]). Crucially, these fluctuations continuously create and destroy $\chi 7$ sites. Their population decreases by about one third relative to low T, and their lifetimes exhibit a long-tailed, non-Gaussian distribution with mean values of only tens of picoseconds but a tail reaching the nanosecond regime, enough to sustain reactivity (Fig. 2.2a).
Free-energy profiles projected onto the N−N distance d and the N_2 total charge q show that the transition state (TS) region preceding the N−N cleavage, remains centered near $q \approx 2$ and $d \approx 1.7$Å at low and high T. However, the TS energy increases by $\sim$ 0.2 eV at high T (Fig. 2.2b). More importantly, the α' state is no longer metastable once the surface becomes mobile (Fig. 2.2c). Committor and environment-similarity analyses further reveal that the ensemble of configurations with high $q(N_2)$ and preceding N−N cleavage, span both $\chi 7$-like and defect-rich neighborhoods. Thus, substantial electron donation is the common prerequisite for N≡N cleavage, not a precise $\chi 7$ motif (Fig. 2.2d, see video in [2]). Furthermore, the authors demonstrated that a pronounced temperature dependence of the activation energy for N≡N cleavage: the effective kinetics at *operando* T deviate from low-T trends. Indeed, a non-Arrhenius behavior arises from barrier changes and the loss of the metastable α' precursor. Consequently, high-temperature behavior cannot be inferred by extrapolating low-temperature data and quantitative predictions must be obtained directly in the relevant *operando* T window. This perspective corroborates Spencer's dynamical picture of catalysis and supports the view that "active sites" are transient structures arising from an ensemble of thermally accessible configurations.

Building upon the results that catalysis on iron surfaces at *operando* temperatures emerges from dynamic and transient atomic arrangements rather than fixed active sites, Tripathi et al. [39] extended the paradigm by explicitly incorporating the effects of lateral interactions originating from the accumulation of adsorbed nitrogen atoms. Coverages up to 50% at 300K and 700K were examined in the work. At low temperatures, consistent with classical findings, nitrogen adatoms form ordered structures that inhibit catalytic activity by blocking $\chi 7$ sites [37]. However, at high temperature, this poisoning effect is alleviated by the interplay of two phenomena. Specifically, nitrogen induces the stabilization and clustering of Fe_3N_i ($i = 1, \ldots, 4$) triangular surface motifs (Fig. 2.3, see video in [3]). The clustering of N atoms introduce structural heterogeneity across the surface and spatially segregate the poison away from catalytically active regions. Crucially, the $\chi 7$ catalytic motifs, while fewer in number, develop longer lifetimes under these conditions due to a reduction in competitive N binding, preserving catalytic activity. Despite partial poisoning, the free-energy barrier for N_2 dissociation stays nearly constant at 700K across coverages, in sharp

[1] https://youtu.be/PojmmzHpXQM?si=12FfHnAt1f4b3rAj
[2] https://youtu.be/qIAVgzj7my8?si=tkv9ZXx2fFdAXMkX
[3] https://youtu.be/SqLkuXtxDIs

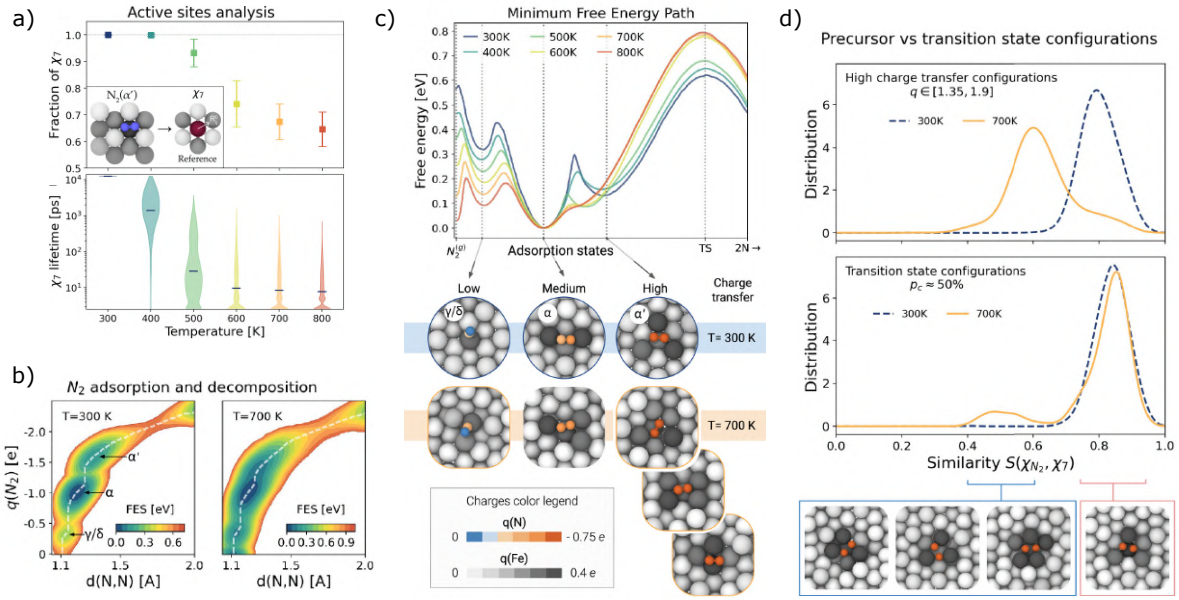

Figure 2.2: **N_2 adsorption on Fe(111) surface**. a) Distribution of $\chi 7$ sites. b) Free energy surfaces at 300K and 700K for the adsorption of N_2 as a function of the Bader charges $q(N_2)$ and the N−N distance. c) Minima free energy pathways at different temperatures and ensembles of configurations for each adsorption state. d) Distribution of the similarity between reference $\chi 7$ site and atomic configurations. Reprinted with permission from ref. [38]. Copyright 2025 National Academy of Sciences. Courtesy of Luigi Bonati.

contrast to the pronounced barrier increase observed at 300K. These findings underscore the critical importance of explicitly accounting for dynamic effects under *operando* conditions, as lateral interactions at high temperature can lead to fundamentally different behaviors compared to those observed in low-temperature or static scenarios.

2.2 Ammonia decomposition

In recent years, ammonia has attracted renewed interest not only as a key chemical feedstock but also as a promising energy vector in the transition toward carbon-neutral energy systems. Thanks to its high hydrogen density (17.6 wt%), ammonia provides an efficient solution for hydrogen storage and long-distance energy transport, helping to overcome the challenges associated with storing and handling molecular hydrogen. Furthermore, ammonia can be catalytically decomposed to release hydrogen on demand, making it a potential fuel for fuel cells and other clean energy technologies [40, 41, 42].
In order to enable ammonia to serve as a hydrogen carrier, it must undergo decomposition according to the following reaction:

$$2\,NH_3 \longrightarrow N_2 + 3\,H_2$$

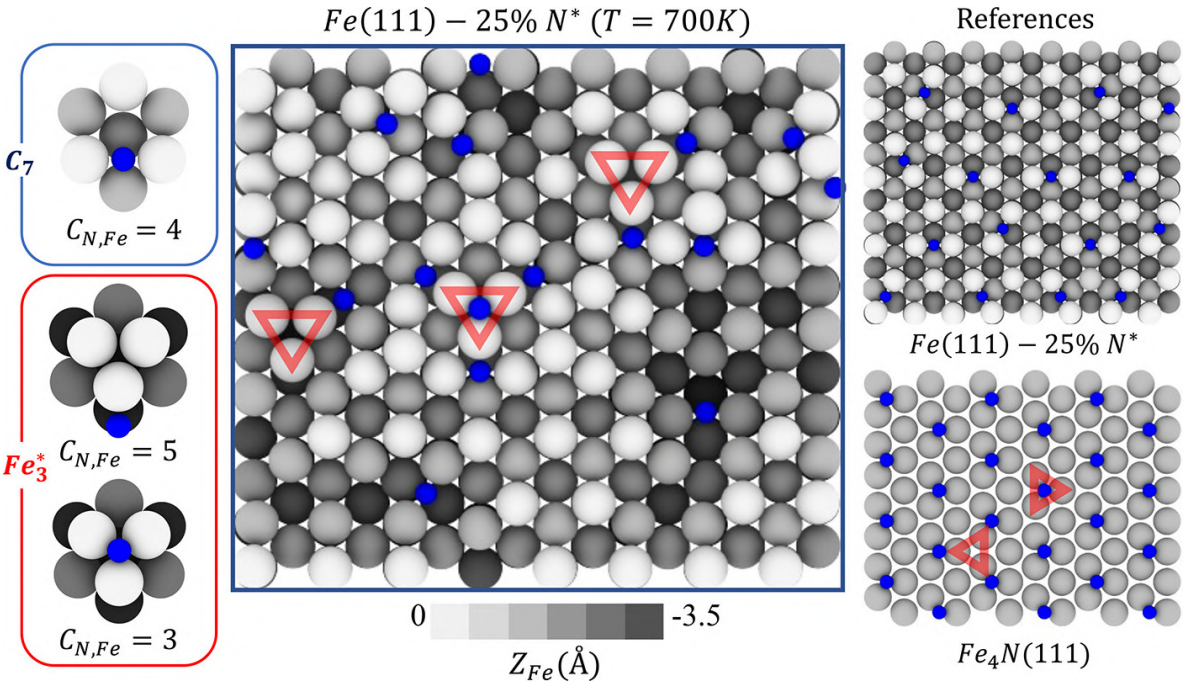

Figure 2.3: **Snapshot of the nitridated Fe(111) surface**. Snapshot at 700K of the Fe(111) surface at 25% coverage of N*. Fe atoms are color-coded according to their height along the [111] axis, the N* atoms are colored in blue. The Fe_3* sites, each surrounded by varying numbers of N* atoms, are indicated by red triangles. Different coordination environments of N* are illustrated on the left. In the top right panel is reported a reference Fe(111) with 25% N* coverage, obtained by optimizing the experimentally suggested structure at 300K, in which all N* occupy 4-fold coordinated (C7) sites. On the bottom right panel, a picture of the $Fe_4N(111)$ surface is presented, where the triangles analogous to those observed on the Fe(111) surface at 700K are highlighted. Reprinted with permission from ref. [39]. Copyright 2024 American Chemical Society. Courtesy of Luigi Bonati.

On metal surfaces, ammonia decomposition, also known as cracking, proceeds through three successive dehydrogenation steps, ultimately leading to nitrogen desorption via the recombination of surface-bound nitrogen atoms:

$$
\begin{aligned}
NH_3 + {}^* &\longrightarrow {}^*NH_3 \\
{}^*NH_3 &\longrightarrow {}^*NH_2 + {}^*H \\
{}^*NH_2 &\longrightarrow {}^*NH + {}^*H \\
{}^*NH &\longrightarrow {}^*N + {}^*H \\
2\,{}^*N &\longrightarrow N_2 + {}^* \\
2\,{}^*H &\longrightarrow H_2 + {}^*
\end{aligned}
$$

where * indicates the adsorption site. The temperature conditions are similar to those of the Haber-Bosch process, and in this case as well, the dynamic evolution of the surface, the emergence of a broad ensemble of reactive pathways and the effect of lateral interactions become

important.

The mechanism of NH_3 dehydrogenation and the diffusion of NH_x species (x=1-3) were investigated at 300K and 700K on the compact Fe(110) and the more open Fe(111) surfaces by Perego *et al.* [43] To isolate temperature and surface–structure effects, the study focused on a single NH_x adsorbate, so no lateral interactions were included. The two facets respond very differently to temperature, and this controls both diffusion of NH_x species and their reactivity.
On the compact, densely packed Fe(110) surface, the morphology remains essentially unchanged from 300 to 700K. The energy for defect formation is too high to occur spontaneously, and no Fe self-diffusion is observed even at high T. On this facet, diffusion and reactivity at 700K mirror the low-T picture (see video in [4]). Consistently, the free-energy profiles show no substantial temperature effect, and the kinetics follow simple Arrhenius-type behavior on this rigid surface (Fig. 2.4). By contrast, the more open Fe(111) surface exhibits significant thermal reconstruction at operando temperature. At 700K, above the Hüttig temperature, surface Fe atoms become mobile, continuously creating and annihilating defects and opening adsorption motifs absent on the ideal surface, such as defective bridges (db) and triangular sites (tri1, tri2). This fluxionality reshapes adsorption geometries and the diffusion of intermediates. In particular, NH_2, which at 300K diffuses almost freely among b12 bridges via b11, becomes markedly less mobile at 700K: it localizes on isolated db sites that provide an sp^3-like tetrahedral environment, suppressing its otherwise barrierless diffusion (see video in [5]). NH likewise populates several high-T environments (tri1, tri2, and defective fourfold sites) with significant charge transfer from the metal (see video in [6]). The net result is an increase in dehydrogenation barriers for NH_2 and NH at 700 K on Fe(111), a clear departure from Arrhenius behavior that static, low-T models cannot capture (Fig. 2.4). Furthermore, analysis of many reactive trajectories shows that there is no single transition state involved in dehydrogenation (see video in [7]). Indeed, each dehydrogenation step proceeds through a TS ensemble whose members differ in local Fe coordination.

The N removal following NH_3 decomposition on Fe(110) was later investigated by Purcel et al. [44] focusing specifically on the competition between nitrogen recombination to form N_2 and nitrogen migration into the bulk to form a stable nitride phase. At low nitrogen coverage (0%), they found that recombination of two adsorbed nitrogen atoms to form N_2 occurs via a transition state located at a bridge site and involves a substantial free energy barrier of approximately 216kJ/mol (Fig. 2.5, right panel). This step is identified as the rate-limiting process in the ammonia decomposition mechanism on Fe(110). By contrast, the migration of a nitrogen atom from the surface into the subsurface octahedral interstitial site is energetically more favorable, requiring a barrier of around 170kJ/mol, with further diffusion deeper into the bulk facing an even lower barrier of only about 65kJ/mol. As nitrogen coverage increases to 20% and 40% of a monolayer, a notable shift in the energy barriers between these two processes is observed. While the barrier for recombination

[4] https://youtu.be/VVSIw8NOg4M
[5] https://youtu.be/GwQJnIvgobk
[6] https://youtu.be/DsHwb-n-Aoo
[7] https://youtu.be/J_5FwGMwg78

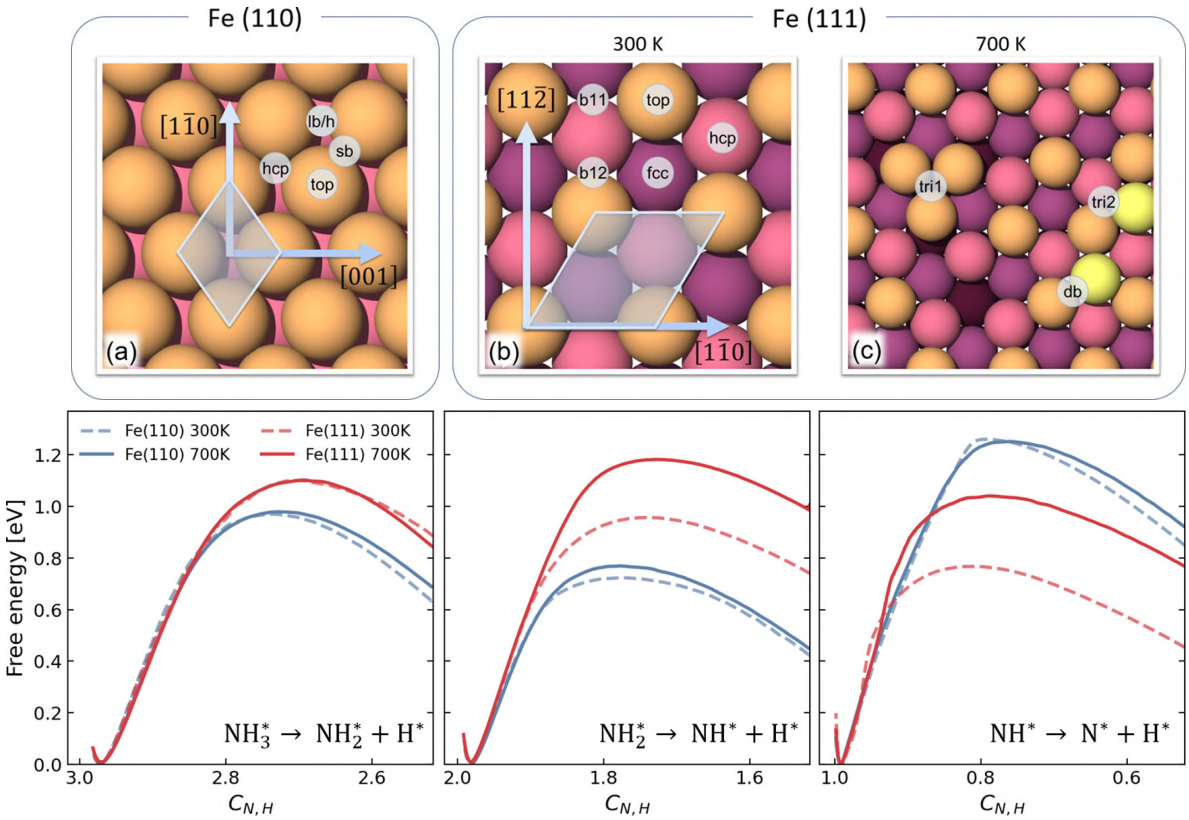

Figure 2.4: **Adsorption sites and free-energy profiles for NH_3 dehydrogenation**. (Top) Top view of adsorption sites on Fe(110) and Fe(111) surfaces: sb, short-bridge; lb, long-bridge; h, hollow; b11, bridge between top-layer atoms; b12, bridge between first and second layer; tri1, triangular reconstruction; tri2, high symmetry site between three consequential layers; db, defective b12. (Bottom) Free energy profiles for the three dehydrogenation steps on Fe(110) and Fe(111) surfaces as a function of the coordination number between N and H. Reprinted with permission from ref. [43]. Copyright 2024 American Chemical Society. Courtesy of Simone Perego.

into N_2 decreases slightly, from 216 to approximately 196 kJ/mol, the barrier for nitrogen migration is reduced more significantly, from 169 to 134kJ/mol (Fig. 2.5, right panel). This asymmetry in the energy reduction with increasing coverage strongly enhances the preference for nitrogen incorporation into the bulk Fe rather than desorption as N_2.

In order to limit nitrogen migration into the bulk and promote recombination, alternative materials must be explored. Experimentally, iron–cobalt alloys FeCo have shown improved catalytic performance compared to pure iron. A recent work by Perego and Bonati [46] elucidated the mechanism of ammonia cracking on FeCo(110) surface and rationalized atomistically its higher performance. For each of the three dehydrogenation steps, the authors determined the free energy profiles and discovered that each step involves multiple reactive pathways characterized by different local coordination environments (Fig. 2.6). Despite the variability of reactive envi-

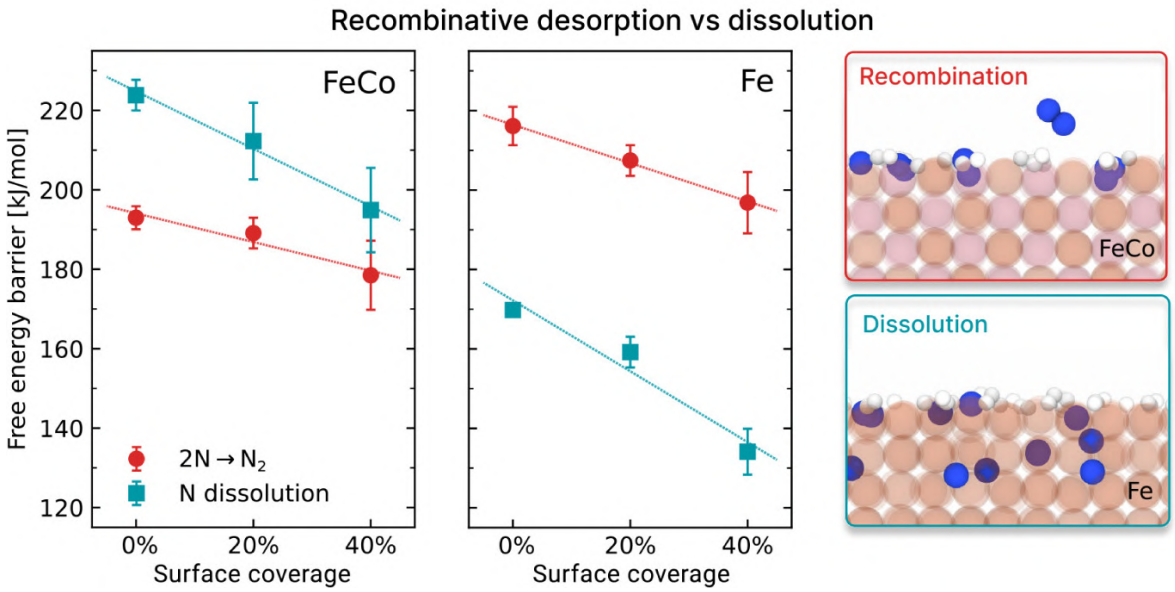

Figure 2.5: **Recombination and desorption of N* in Fe and FeCo**. The panels shows the trend of free energy barriers for N* recombination (red line) and migration (cyan line) in (left) FeCo(110) and (right) Fe(110). Reprinted with permission from ref. [45]. Copyright 2025 American Chemical Society. Courtesy of Simone Perego.

ronments, characterized by a range of N–Co coordination, the barrier heights remained nearly unchanged, demonstrating the degeneracy of accessible transition states under realistic thermal conditions. This mirrors observations made on pure iron surfaces and reinforces the conclusion that an ensemble of overcoordination environments is a prerequisite for reactivity.

In a following paper, Perego et al., investigated the recombination and bulk diffusion of nitrogen atoms at 700K on FeCo(110) [45]. The work also included later interactions by considering surface coverages of 20% and 40% of a monolayer, maintaining an N:H ratio of 1:3. The authors found that nitrogen recombination is favored over dissolution into the bulk (Fig. 2.5, left panel). This behavior is attributed to the stronger affinity of nitrogen for iron compared to cobalt, indicating weaker N-Co interactions. Additionally, nitrogen dissolution becomes progressively less favorable with increasing depth. In short, under operating conditions Fe drives bulk N uptake, whereas FeCo promotes N recombination on the surface, thereby supporting its enhanced catalytic stability during ammonia decomposition.

So far, the decomposition of ammonia was described on transition metal systems that at operando conditions provide transient active sites and ensembles of reactive pathways. Another class of materials investigated for ammonia cracking are ionic compounds, such as lithium imide (Li_2NH), lithium imide-amide mixture ($Li_{2-x}(NH_2)_x(NH)_{1-x}$) and lithium chromium nitride oxide ($Li_{14}Cr_2N_8O$). In the study by Yang et al. [47], MLP-MD simulations revealed that the catalytic activity of Li_2NH in

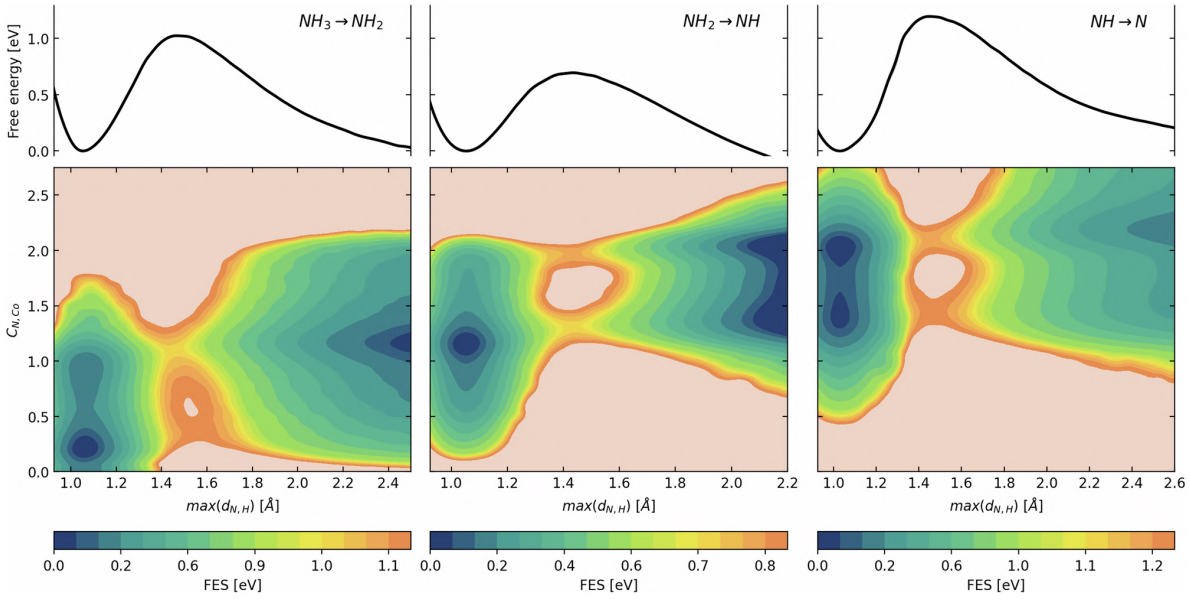

Figure 2.6: **Free energy profiles for dehydrogenation steps on FeCo(110) surface**. Free energy profile of dehydrogenation steps projected along the maximum distance between N and H atoms $max(d_{N,H})$ (top), and on the 2D plane defined by $max(d_{N,H})$ and the coordination number $C_{N,Co}$ (bottom). Note that for each reactions, there is an ensemble of transition states, characterized by a range of N−Co coordination. Reprinted with permission from ref. [45] under a Creative Commons Attribution 4.0 International License. Courtesy of Simone Perego.

ammonia decomposition arises from an emergent, dynamically activated interfacial state rather than from any fixed atomic configuration. At elevated temperatures (T>350K), bulk Li_2NH enters a superionic phase, wherein Li^+ cations exhibit high mobility through interstitial hopping between tetrahedral and octahedral sites, while the NH^{2-} imide anions oscillate around equilibrium positions (see video in [8]).

Before exposure to ammonia, the (111) surface remains structurally stable even under superionic conditions. However, when two NH_3 molecules interact with the surface, a stoichiometric reaction occurs:

$$2NH^{2-} + 2NH_3 \longrightarrow 4NH_2^-$$

This reaction transforms doubly charged imide anions into singly charged amides, leading to a sudden increase in charge disorder and a breakdown of the original surface order (see video in [9]). The result is a transition to a highly dynamic regime where both amides and imides display diffusive behavior, with amides diffusing more rapidly. Furthermore, Grotthuss-like proton shuttling between anions appears on nanosecond timescales. This transformation is spatially confined to the top two atomic layers of the surface, which begin to behave like an ionic liquid, offering a fluxional and adaptable reaction medium.

[8] https://youtu.be/8gsvvpXOqrE
[9] https://youtu.be/jdP3HTSjoQI

A central early step is the formation of the diazanediide anion ($[HN{-}NH^{2-}]$):

$$NH^{2-} + NH^{2-} \longrightarrow [HN{-}NH]^{2-} + 2\,e^{-}$$

The $[HN{-}NH^{2-}]$ is stabilized by a cloud of mobile Li^{+} cations, whereas the two electrons can reside either in a diffuse surface state or in a localized F-center, both favored by the interfacial electrostatics. A complementary route involves hydride formation:

$$NH^{2-} + {NH_2}^{-} \longrightarrow [HN{-}NH]^{2-} + H^{-}$$

The diazanediide, along with the resulting hydride (H^{-}) and electrons, initiates further reactions ending in the formation and desorption of N_2 and H_2.
Free-energy calculations make the interfacial activation explicit: on the activated (111) surface (namely, after beeing exposed to NH_3) the barrier to form $[HN{-}NH]^{2-}$ is lowered by approximately 60kJ/mol relative to the pristine surface, and the reaction is favored in the top two layers rather than deeper layers (Fig. 2.7a). This findings are consistent with the view of catalysis being a collective phenomenon that takes place in a dynamically activated environment.

In a following paper, Mambretti et al. [48] explored the catalytic behavior of non-stoichiometric lithium imide compounds, $Li_{2-x}(NH_2)_x(NH)_{1-x}$ toward ammonia decomposition. Using ML-accelerated MD simulations combined with enhanced sampling techniques, the authors demonstrated that increasing the concentration of amide groups (x) significantly enhances the structural disorder of the catalyst surface, even before ammonia exposure. This enhanced disorder originates from amplified fluctuations in the topmost atomic layers, including spontaneous promotion of ${NH_2}^{-}$ anions into an adlayer and associated interlayer atomic exchanges. Notably, the system exhibits dynamic characteristics that facilitate proton transfer via a Grotthuss-like mechanism, further supported by lithium ion mobility that helps stabilize negatively charged intermediates. These proton exchanges occur both before and after ammonia activation and increase in frequency with rising x, promoting a highly responsive interfacial environment. The free energy barrier for the formation of $[HN{-}NH]^{2-}$ from two imides is shown to decrease with increasing amide concentration, dropping from 238kJ/mol at x=0 to 204 kJ/mol at x=0.2, thereby making the key N−N bond formation step more favorable (Fig. 2.7b). Importantly, this barrier reduction is observed not only in the outermost adlayer but also within the second subsurface layer. Also in this case, the two electrons released are accommodated either in a diffuse surface state or in a localized F-center state, while nearby Li^{+} provides electrostatic stabilization. Overall, the enhanced structural disorder strengthens the dynamical, fluxional picture of catalysis on lithium imide–amide mixtures and rationalizes their experimentally observed higher activity relative to pristine Li_2NH.

Ultimately, the lithium chromium nitride oxide $Li_{14}Cr_2N_8O$ has been studied as a candidate for ammonia decomposition [49]. This compound adopts an ordered antifluorite-derived structure ($Na_{14}Mn_2O$-type, space group P3), which remains structurally stable up to 900K in the absence of reactants. Upon exposure to ammonia, however, superficial nitrides undergo protonation, leading to the formation of imide species and a dynamic surface adlayer composed of amides. This adlayer is stabilized by mobile Li^{+} ions, which coordinate with the negatively charged amides. As

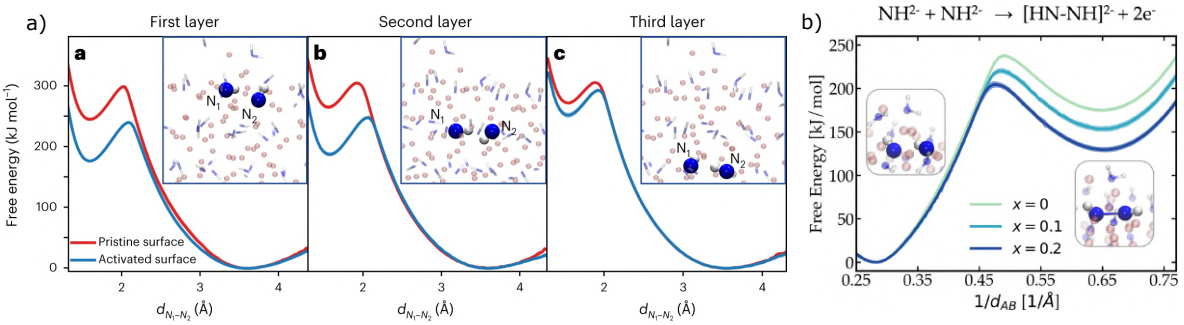

Figure 2.7: **Free-energy profiles for $[HN{-}NH]^{2-}$ formation**. a) $[HN{-}NH]^{2-}$ formation in Li_2NH for different superficial layers. The free-energy profiles are reported as function of the distance between nitrogens of two reacting imides. Reprinted with permission from ref. [47]. Copyright 2023, The Author(s), under exclusive licence to Springer Nature Limited. Courtesy of Manyi Yang. b) $[HN{-}NH]^{2-}$ formation in $Li_{2-x}(NH_2)_x(NH)_{1-x}$ at different x values. The free-energy profiles are reported as function of the inverse distance between nitrogens N_A and N_B of a given pair of imides. Reprinted with permission from ref. [48]. Copyright 2024 American Chemical Society. Courtesy of Francesco Mambretti.

amide coverage increases, lithium ion mobility is enhanced, contributing to significant structural disorder and a progressive loss of rigidity throughout the material. At full hydrogenation, even chromium atoms are displaced from their original lattice positions. These transformations are supported by operando X-ray diffraction data, which reveal the emergence of chromium nitride (CrN) phases at approximately 573K and a gradual decomposition of the pristine $Li_{14}Cr_2N_8O$ framework as the temperature approaches 900K. The absence of new lithium-containing crystalline phases suggests that lithium remains in a highly disordered, amorphous state following decomposition. From a mechanistic standpoint, the rate-limiting step involves the coupling of two imide species to form diazanediide $[HN{-}NH^{2-}]$, with a substantial activation barrier of ~288kJ/mol, consistent with the experimentally measured apparent activation energy of 362.8 kJ/mol (Fig. 2.8a). The electrons released in this process are transferred to nearby chromium atoms, which act as redox centers (Fig. 2.8b). Subsequent reaction steps include hydride (H^-) formation and hydrogen recombination to yield molecular H_2, with estimated barriers of 140kJ/mol and 100kJ/mol, respectively.

2.3 Electrocatalysis

Electrocatalysis occurs at electrochemical interfaces, where a solid electrode is in contact with a liquid electrolyte. These interfaces are inherently complex, with many coupled degrees of freedom spanning multiple length and time scales. Even seemingly "elementary" reaction steps happen within a dynamic environment that involves charge transfer, solvent reorganization, and electrode surface reconstruction under reaction conditions. Modeling such electrified interfaces involves two fundamental challenges: representing the electrolyte and applying a constant electrode potential [50].

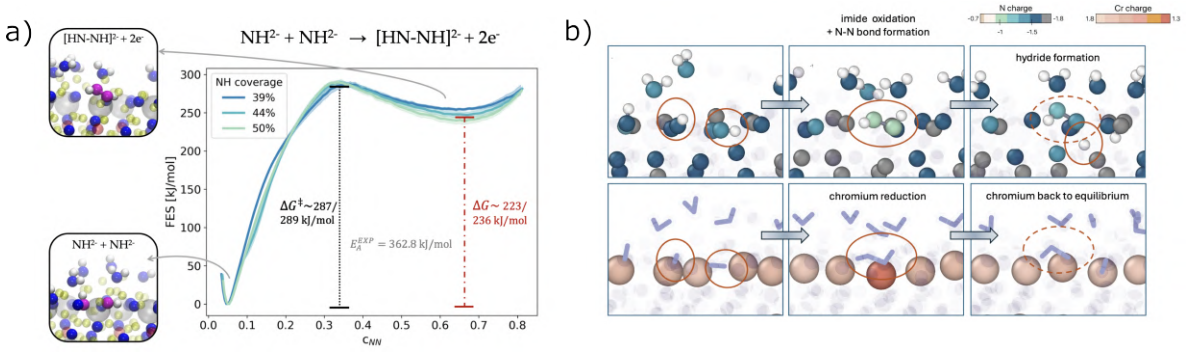

Figure 2.8: **$[HN{-}NH]^{2-}$ formation in $Li_{14}Cr_2N_8O$**. a) Free energy profiles for $[HN{-}NH]^{2-}$ formation along the coordination number between two NH^{2-}. b) Representative snapshots for $[HN{-}NH]^{2-}$ formation while monitoring the change in Bader charges for N (upper panel) and Cr (lower panel) atoms. Reprinted with permission from ref. [49]. Courtesy of Francesco Mambretti.

A key reason for properly modeling the liquid electrolyte phase is its active role in electrochemical reactivity [51, 52]. The electrolyte solvates adsorbates through directional interactions such as hydrogen bonding, and its structure and fluctuations near the surface modulate reaction thermodynamics and mechanisms, including the stabilization of intermediates and facilitation of proton and electron transfer steps. Several studies illustrate the essential role of explicit solvent in capturing such effects. For example, Farnesi Camellone and Marx used AIMD to study an Au_{11} nanocluster supported on reduced TiO_2(110) in contact with liquid water [53]. They showed that solvation induces a significant charge separation within the cluster, with top-layer Au atoms becoming partially oxidized ($Au^{+\delta}$) and inner-layer atoms becoming partially reduced ($Au^{-\delta}$). This charge redistribution is absent in vacuum and it is stabilized by directional hydrogen bonds. In fact, water solvates differently charged Au atoms as either cationic or anionic centers. Such solvent-driven effects create novel active sites and dynamically modulate the electronic structure of the interface, suggesting that water is not merely a passive medium but an active component that shapes catalytic reactivity.

Another compelling example is provided by Tocci and Michaelides, who performed AIMD simulations of the H_2O–ZnO($10\bar{1}0$) interface to explore the role of solvation in proton transport [54]. Comparing a single water monolayer, mimicking ultra-high vacuum conditions, with a more realistic liquid multilayer, they observed a more than tenfold increase in proton hopping rate in the presence of the extended liquid film. This enhancement was not due to increased water dissociation, which remained comparable in both cases, but rather to dynamic fluctuations in the hydrogen-bond network that transiently shorten O–O distances and lower the free-energy barriers for proton transfer. Their results revealed that proton mobility is highly sensitive to the local solvation environment and that only explicit solvent models can capture such Grotthuss-like transfer mechanisms.

While solvation is critical, modeling electrocatalysis also requires proper control of the electrode

potential [50, 52]. A widely adopted strategy is the Computational Hydrogen Electrode (CHE) approach introduced by Nørskov [55], which references the chemical potential of protons and electrons to that of hydrogen gas under standard conditions. Potential effects are included a posteriori by correcting the reaction free energies based on the number of proton–electron pairs transferred. Despite its utility for estimating adsorption trends and thermodynamic overpotentials, CHE presents key limitations. It lacks an explicit treatment of electrode potential during DFT calculations, neglecting field effects, charge redistribution, and solvent response. It also assumes concerted proton–electron transfer and typically omits activation barriers, limiting its ability to capture potential-dependent kinetics or stepwise mechanisms [56]. To overcome these issues, grand-canonical DFT frameworks have been developed [57, 58, 59]. These allow the number of electrons to vary at fixed electrochemical potential by adjusting the Fermi level and compensating net charge through background charges, continuum solvation, or explicit electrolytes. While more computationally intensive, these methods offer a more realistic, operando-consistent description of electrified interfaces and their reactivity [60].
While these kind of approach will be further discussed later in this thesis (see § 3.3), the effects of applied potentials are herein illustrated.

Electrode surfaces can reconstruct under operando conditions of finite applied potential and reactant coverage, creating new active sites [61, 62]. Using grand-canonical Monte Carlo (GCMC), environmental KMC (EKMC), and DFT, Zhang et al. tracked the reconstruction of Cu(100) across -0.2 to -0.8V vs RHE and show how these changes reshape CO_2-reduction pathways to C−1 and C−2 products [63]. Not only do H and CO coverages depends on potential, but they actively drive defect formation, producing adatoms, vacancies, and square four-atom Cu clusters (Fig. 2.9). Under applied potential and in the presence of adsorbates, these defects have lower formation energies than the clean surface, making them easy to form. The authors demonstrated also that selectivity is site-dependent: vacancies favor CH_4, adatoms promote H_2 and HCOOH, and four-atom clusters favor CO and C−2 products (e.g., ethylene, ethanol).

Lian et al. employed large-scale MLP-driven MD to explore the structural evolution of oxide-derived copper (OD-Cu) under electrochemical conditions [64]. Their simulations revealed that oxygen retention in OD-Cu is strongly dependent on pH, applied potential, and particle size, with smaller particles and alkaline environments favoring higher residual oxygen content. Under operando conditions, the copper surface undergoes significant reconstruction, giving rise to a heterogeneous distribution of under-coordinated sites—including atop, bridge, and hollow-like geometries—that are absent on ideal crystalline facets. These low-coordination environments are proposed to serve as active sites for CO_2 reduction, providing a microscopic explanation for the high catalytic activity observed experimentally in OD-Cu systems (Fig. 2.10) .

Ngoipala et al. investigated how hydride formation dynamically reconstructs palladium electrode surfaces under electrochemical conditions [65]. Using a combination of in situ electrochemical scanning tunneling microscopy (EC-STM), inductively coupled plasma mass spectrometry (ICP-MS), and ML-accelerated MD, they revealed that Pd surfaces undergo facet-specific structural changes upon hydrogen absorption. Subsurface hydrogen insertion induces lattice expansion and misfit

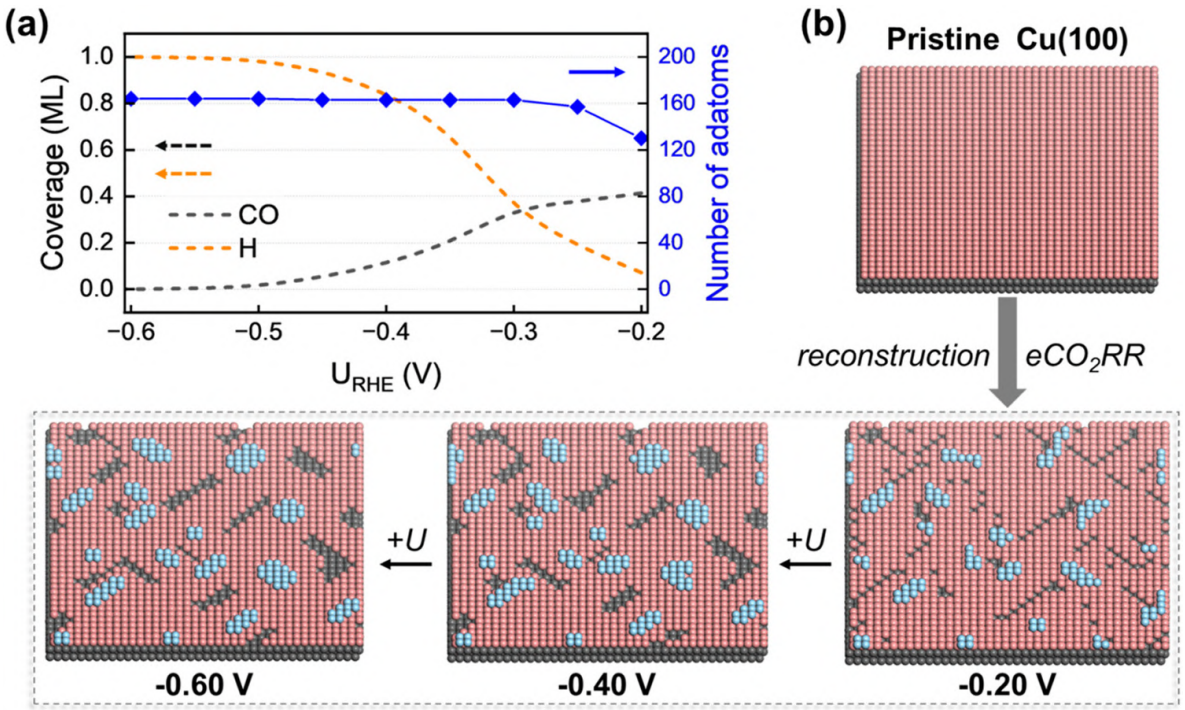

Figure 2.9: **Adsorbate coverage and electrochemical potential effects on Cu(100) surface**. a) Number of adatoms and adsorbate coverages as functions of electrochemical potential. b) Reconstructed surface structures under various electrochemical potentials. Blue atoms represent Cu atoms ejected from the surface. Pink atoms represent Cu atoms on the surface. Gray atoms represent bulk Cu atoms. Reprinted with permission from ref. [63]. Copyright 2025 American Chemical Society.

strain, which relaxes through the formation of extended surface defects. On Pd(111), adatoms and roughened step edges appear; Pd(110) develops missing-row and ad-row patterns; and Pd(100) rapidly reconstructs into {111}-like domains (Fig. 2.11). These reconstructions originate from strain relaxation and hydride phase transitions, highlighting that the active surface state of Pd during HER and HOR is not static but continuously evolves under operating conditions.

2.4 Liquid catalysts

In the previous sections, we discussed how the formation of fluxional active phases on catalyst surfaces is often at the heart of catalytic activity. If the key to effective catalysis lies in providing a highly dynamic and adaptable environment, then liquid catalysts naturally emerge as a compelling option [66]. Thanks to their intrinsic atomic mobility and diffusion, liquid catalysts can sustain a constantly evolving environment, where multiple transient active regions form and dissolve, enabling a broad ensemble of reactive pathways [66, 67].
A class of liquid catalysts that has recently garnered attention is that of liquid metals and liquid metal alloys. These materials have shown promise in a variety of catalytic applications, such as

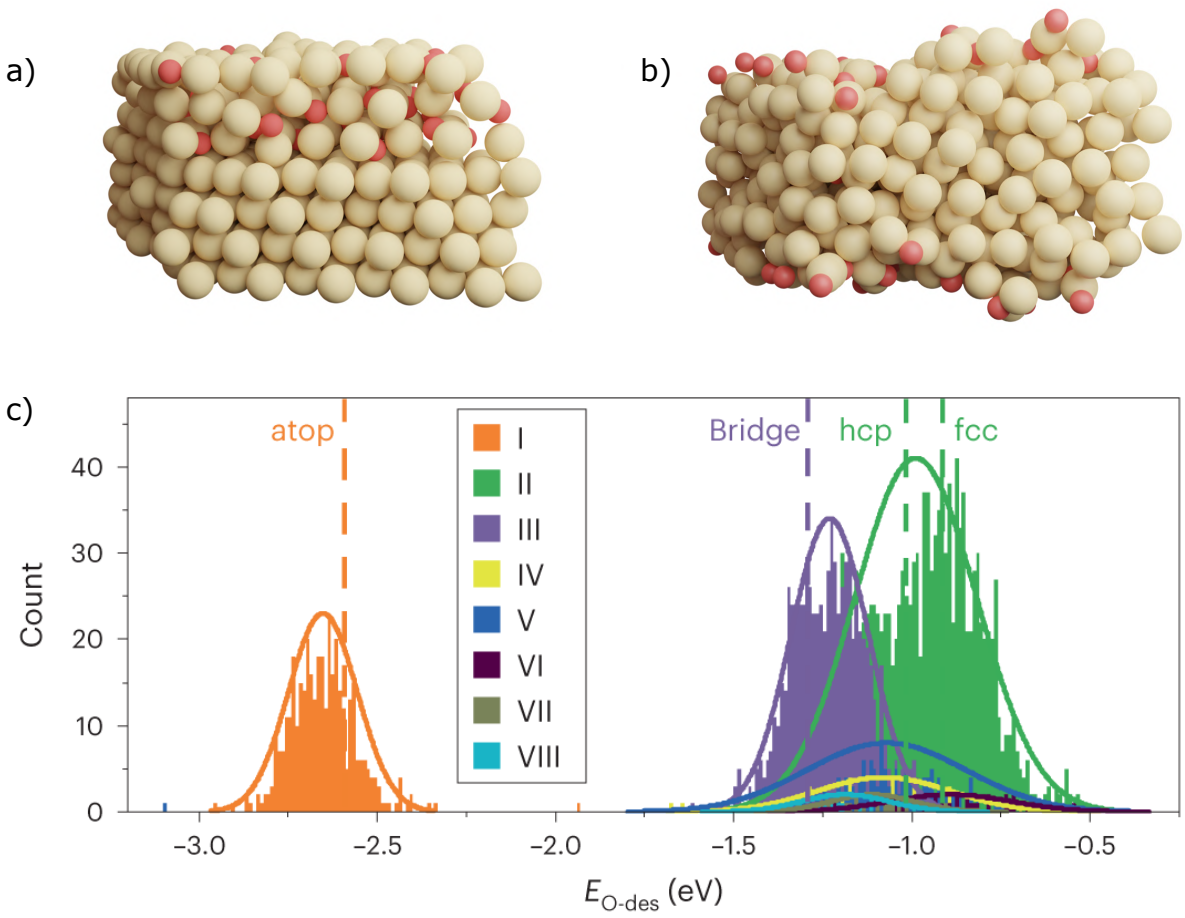

Figure 2.10: **Reconstruction and active sites creation on OD-Cu system**. a) Initial frame of an NVT trajectory. b) Last frame of an NVT trajectory. c) Distribution of oxygen desorption energy for different types of active sites at pH=0 and U=0V vs SHE. Reprinted with permission from ref. [64] under the Creative Commons Attribution 4.0 International License.

methane pyrolysis [68, 69], ammonia synthesis [70], and diamond growth [71].
Particularly compelling are those alloys that are liquid at or near room temperature, as they open new possibilities for carrying out reactions under milder operating conditions compared to traditional solid catalysts [72]. Moreover, their low viscosity and high thermal and electrical conductivity make them excellent media for efficient mass and heat transfer, which are crucial in sustaining catalytic activity under reactive conditions.
Liquid metals have also the ability to dissolve a wide variety of metallic elements in their elemental state, enabling the tuning of the composition and fine control over both electronic and local atomic features of the catalyst, offering a powerful tool for modulating catalytic efficiency [67].
The behavior of bulk liquid metals requires careful consideration and a deep understanding in order to be effectively harnessed for catalytic applications [72, 73]. However, characterizing these systems remains a significant challenge, as most analytical techniques have been developed for molecular or covalent systems and are not well suited to metallic liquids. Their high density, elec-

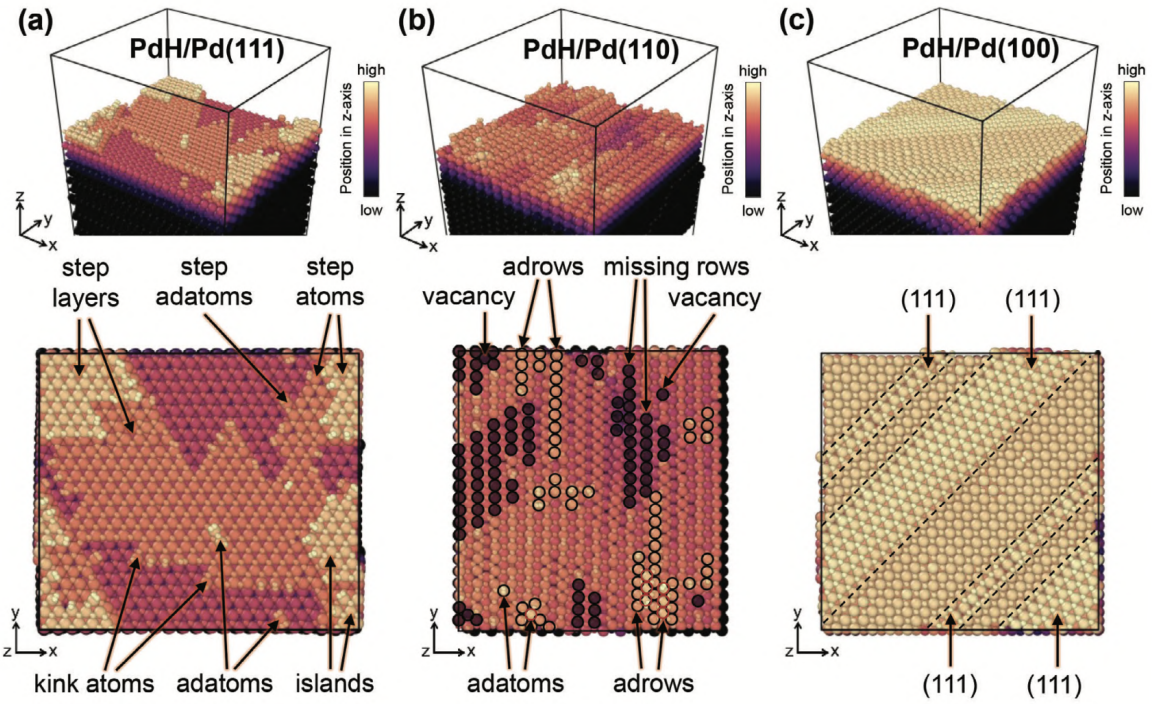

Figure 2.11: **Reconstruction of Pd surfaces**. Reconstruction of PdH/Pd(111), PdH/Pd(110), and PdH/Pd(100) surfaces by formation of adatoms, islands, and two step layers. Reprinted with permission from ref. [65]. Copyright 1999-2025 John Wiley & Sons, Inc.

trical conductivity, and opacity across much of the electromagnetic spectrum limit the applicability of conventional spectroscopic and imaging tools. In this context, computational simulations offer a powerful alternative, enabling direct insight into the atomic-scale structure, dynamics, and reactivity of liquid catalysts systems under operating conditions.
In the following, we present and discuss the results of several notable works that have employed liquid catalytic in various chemical transformations.

The catalytic ability of $GaSn_{0.029}Ni_{0.023}$ to produce propylene from hydrocarbons was investigated by Tang et al. [74]. This Ga-rich ternary alloy demonstrated high selectivity towards propylene formation, achieving 90.5% selectivity at 150°C when exposed to decane, and 94.5% when using canola oil as the feedstock. The presence of all three elements is essential for catalytic activity. Control experiments with binary or single-metal alloys such as Ga, $GaNi_{0.023}$, and $GaSn_{0.029}$ showed either no activity or significantly reduced performance. The observed synergistic effect can be rationalized through several electronic and structural factors. Bader charge analysis revealed that Sn and Ni acquire partial negative charges, enhancing their ability to interact with the hydrocarbon substrate. Owing to its higher polarizability, Sn preferentially interacts with the C−1 carbon of the decane chain, which subsequently positions C−2 in close proximity to a Ga atom. This arrangement facilitates hydrogen abstraction from C−2, with the extracted hydrogen

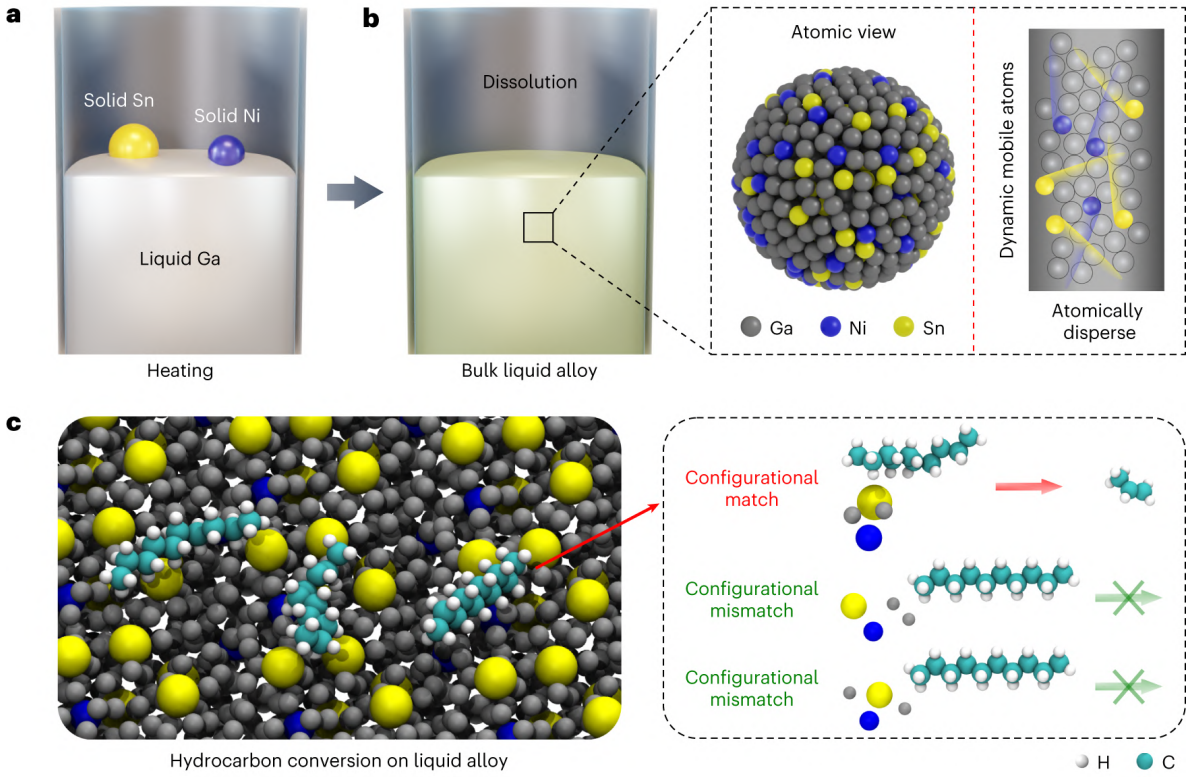

Figure 2.12: **Schematics and computational simulations describing the $GaSn_{0.029}Ni_{0.023}$ catalyst.** Schematics demonstrating the preparation a) and the atomic mobility b) of liquid $GaSn_{0.029}Ni_{0.023}$ catalyst. c) Schematics illustrating that one precise configurational alignment between the hydrocarbon and dispersed atoms in a Ga matrix initiated the selective reaction. Reprinted with permission from ref. [74]. Copyright 2023, The Author(s), under exclusive licence to Springer Nature Limited.

stabilized by two Ga atoms and one Ni atom. The hydrocarbon chain arrangement at the surface, positions C–4 directly above the dissociated hydrogen, favoring the formation of propylene over other alkenes. Upon alignment, the dissociated hydrogen is transferred to C–4, breaking the C–C bond between C–3 and C–4. Concurrently, the interaction between C–2 and Sn weakens, promoting the formation of a double bond between C–2 and C–3, thus completing the selective formation and desorption of propylene (Fig. 2.12).

The high cost of gallium significantly limits the scalability of Ga-based catalytic systems. As a lower-cost alternative, tin (Sn) was investigated as a metal matrix due to its ability to dissolve various elements and form alloys with melting points below 250°C. A notable application of Sn-based alloys for H_2 production from hydrocarbon was demonstrated by Tang et al. [75]. Among the compositions tested, the alloy $SnIn_{0.1034}Cu_{0.0094}$ exhibited high efficiency and selectivity of ~98% at 260°C. AIMD simulations revealed that indium (In) preferentially occupies low-density regions within the Sn matrix, while Cu tends to remain just below the interfacial layer. Cu atoms were

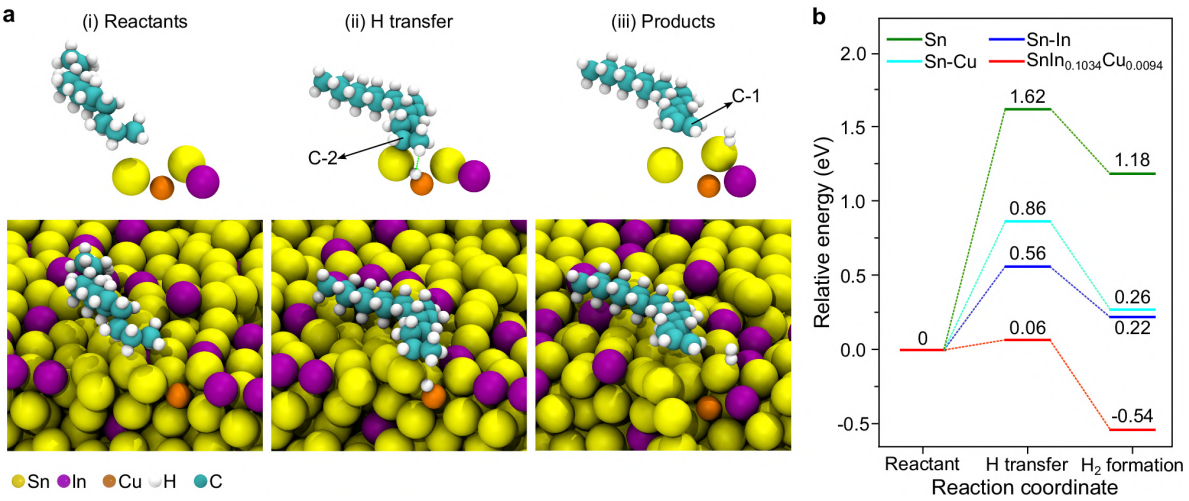

Figure 2.13: **Reaction mechanism for H_2 production**. a) Snapshots for H_2 formation from hexadecane. b) Energy barriers for H_2 formation from hexadecane by using different materials. Reprinted with permission from ref. [75]. Copyright 2025, The Author(s).

found to be highly mobile, transitioning between the bulk and the surface, and importantly, no Cu aggregation was observed. Hydrogen generation was only observed when both In and Cu were present, confirming a synergistic effect necessary to activate the catalytic process. Furthermore, no catalytic activity was detected below the alloy's melting point, indicating that the liquid state is essential for reactivity. To elucidate the reaction mechanism, hexadecane was used as a probe molecule. Upon hydrocarbon exposure, Cu atoms transiently migrate to the surface and abstract a hydrogen atom from the C−2 position. The resulting carbon radical is stabilized by a neighboring Sn atom. The Cu-bound hydrogen then reacts with another hydrogen at the C−1 position, forming molecular H_2 and 1-hexadecene as products. In this process, In appears to play a dual role: stabilizing surface hydrogen species and facilitating the presence and mobility of Cu at the interface (Fig. 2.13).

The flexibility of Ga in catalysis has been investigated also in ammonia synthesis by Zuraiqi et al. [70] using Ga−Cu alloy. Among various considered compositions, the alloy containing 2 wt.% Cu exhibited the highest activity, achieving an NH_3 production rate of $9500 \mu mol g^{-1} h^{-1}$ at 400°C and 4bar, outperforming the activity of pure Ga and pure Cu at the same operating conditions. This indicate a clear synergistic effect between the two metals. Moreover, the apparent activation barrier was determined to be ∼25kJ/mol, about ∼1/4 lower than that measured for a reference Ru-based catalyst. By using AIMD simulations, the authors investigated the behavior of Cu and Ga atoms in the liquid alloy. They found that Cu atoms remain atomically dispersed within the Ga matrix and behave predominantly as isolated single atoms, rather than forming clusters. Copper atoms exhibit high mobility, freely migrating between the bulk and the liquid surface without showing a strong preference for either region. Remarkably, Cu plays a critical role in facilitating hydrogen dissociation, a step that is energetically unfavorable on pure Ga surfaces. The presence of Cu enables hydrogen activation and its subsequent reaction with nitrogen species. Copper also

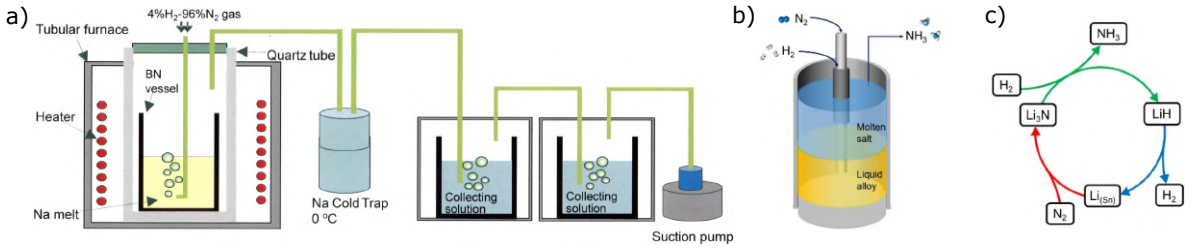

Figure 2.14: **NH_3 synthesis in liquid Na and Li−Sn alloy**. a) Reactor proposed in ref. [76]. Reprinted with permission from [76]. Copyright 2017, The Author(s). b) Schematic representation of the reactor proposed in ref. [77] c) Catalytic cycle for the dual system proposed in ref. [77]. Reprinted with permission from ref. [77]. Copyright 1999-2026 John Wiley & Sons, Inc.

promotes the formation of Ga−N−Cu intermediates, which adopt distorted geometries that are more reactive toward hydrogenation than pure Ga−N configurations. The simulations support an associative reaction mechanism, in which molecular N_2 adsorbs on the surface and undergoes stepwise hydrogenation, ultimately forming NH_3, which then desorbs from the interface.

Other attempts to utilize liquid metals for ammonia synthesis include work by Kawamura and Taniguchi, who employed molten sodium (Na) [76], and by Tang et al., who developed a liquid catalytic system based on a Li−Sn alloy and molten LiCl−KCl salt as part of a chemical looping protocol [77]. In the study by Kawamura and Taniguchi [76], a mixed gas of H_2 and N_2 (4–96%) was bubbled into molten sodium at atmospheric pressure (Fig. 2.14a). At 500°C, they reported a maximum ammonia synthesis rate of 4.17μmol/s, attributing the activity to the ability of liquid sodium to dissociate molecular nitrogen, thus promoting its reaction with hydrogen. However, this system was hindered by catalyst poisoning due to the formation of NaH, which reduced catalytic efficiency. Tang et al. [77] addressed the challenge using a chemical looping approach that avoids poisoning through a multiphase liquid system (Fig. 2.14b,c). In their setup, molecular N_2 is dissociated in a liquid Li−Sn alloy to form Li_3N, which is then transferred into the molten LiCl−KCl salt phase. There, Li_3N is decomposed and N^{3-} reacts with hydrogen to produce NH_3. The resulting LiH produced by Li_3N consumption is decomposed at the alloy–salt interface, regenerating metal Li in the Sn matrix and completing the loop. This process enabled an ammonia yield of approximately 0.03μg/s for alloys containing 60% of Li. Control experiments using pure Sn and eutectic LiCl−KCl salt showed negligible ammonia production, indicating that Sn alone is catalytically inactive for N_2 fixation and primarily acts as a solvent for Li.

As illustrated by the examples discussed, the combination of different elements and their mutual synergy gives rise to catalytic pathways that are inaccessible to the individual components alone. This behavior is a clear manifestation of emergent phenomena, where catalytic activity arises from the fluxional, compositionally dynamic environment of the liquid metal system, rather than from static, isolated atomic sites.

Chapter 3

Methods for modeling catalyst dynamics under operando conditions

This chapter is divided into two parts to reflect the evolution of atomistic simulation techniques in catalysis. The first part focuses on traditional, physics-based methods, including molecular dynamics, density functional theory (DFT), grand canonical DFT, and enhanced sampling techniques, which do not rely on data-driven models. These approaches have formed the foundation of theoretical catalysis, enabling a detailed understanding of reaction mechanisms and material behavior under various conditions. The second part introduces machine learning (ML) approaches that have recently emerged as powerful tools to accelerate and extend atomistic simulations. This includes the use of neural networks to construct interatomic potentials, as well as the integration of ML techniques into enhanced sampling methods. Together, these data-driven approaches allow for unprecedented simulation efficiency and access to larger spatial and temporal scales.

Physics-based methods

3.1 Molecular dynamics

In the introduction, we have discussed several examples demonstrating that catalysis emerges as a collective and dynamic phenomenon. To investigate such a behavior at the atomic scale, the computational method of choice is molecular dynamics (MD), which explicitly simulates the evolution of a system in time by integrating the equations of motion for all particles. This section outlines the fundamental principles of MD and discusses the methods available for computing the forces acting on atoms, which lie at the core of MD simulations [78, 79, 80].

3.1.1 Fundamentals of molecular dynamics

At the heart of MD is the classical formulation based on Newton's equations of motion. Given a system of N atoms with positions r_i the time evolution is determined by:

$$m_i \ddot{r}_i = \mathbf{F}_i = -\nabla_{r_i} \mathcal{U}(r^N) \tag{3.1}$$

where m_i is the mass of atom i , $\ddot{r}_i$ its acceleration, and $\mathcal{U}(r^N)$ is the interatomic potential energy which depends on the complete set of 3N atomic coordinates ($r^N = (r_1, r_2, \ldots, r_N)$). The definition of the model for the interaction between system particles is the first ingredient for the MD simulation and fundamentally determines the accuracy and applicability of the simulation (see § 3.1.2). Once the potential energy function is specified, the equations of motion in Eq. 3.1 are integrated numerically using finite-difference algorithms. These integrators propagate the atomic positions and velocities over discrete time steps Δt, typically on the order of femtoseconds, to produce a time-resolved trajectory of the system. A common numerical integration algorithm employed is the Velocity Verlet scheme [81]. In this scheme, the position $r_i(t + \Delta t)$ and velocity $\dot{r}_i(t + \Delta t)$ are updated as follows:

$$r_i(t + \Delta t) = r_i(t) + \dot{r}_i(t)\Delta t + \frac{1}{2}\ddot{r}_i(t)\Delta t^2 \tag{3.2}$$

$$\dot{r}_i(t + \Delta t) = \dot{r}_i(t) + \frac{1}{2}\left[\ddot{r}_i(t) + \ddot{r}_i(t + \Delta t)\right]\Delta t \tag{3.3}$$

The error term for positions is $O(\Delta t^4)$ while for velocities is $O(\Delta t^3)$. An important feature of this algorithm is that it is exactly time reversible and symplectic, meaning that it preserves the volume in phase space. Moreover, it requires just one force evaluation per step, making it computationally efficient [78, 79].

The third essential ingredient is the choice of a statistical ensemble, which determines which thermodynamic quantities, such as temperature, pressure, or chemical potential, are held constant during the simulation. The most natural ensemble for molecular dynamics is the microcanonical ensemble (NVE), where the number of particles (N), volume (V), and total energy (E) are conserved. This is because, in the absence of external perturbations, the system's Hamiltonian remains constant over time. The Velocity Verlet algorithm inherently generates trajectories consistent with the microcanonical ensemble. However, by introducing suitable modifications to the Hamiltonian, such as coupling to thermostats [82, 83] or barostats [84, 85], it is possible to simulate other ensembles, like the canonical (NVT) or isothermal–isobaric (NPT) ensembles. These extensions allow for the control of temperature and pressure, enabling the study of systems under more realistic thermodynamic conditions.

Molecular dynamics trajectories provide access not only to the microscopic evolution of the system but also to macroscopic observables, which can be obtained as statistical averages of suitable microscopic quantities. In statistical mechanics, any measurable property A_{obs} can be expressed as the ensemble average of a function $A(q, p)$ of positions q and momenta p, following the standard notation of statistical mechanics:

$$A_{\text{obs}} = \langle A \rangle = \int dq\, dp\, \rho(q, p) A(q, p) \tag{3.4}$$

where $\rho(q, p)$ is the equilibrium distribution of the chosen ensemble. Under the assumption of *ergodicity* [86, 87], the time evolution of the system explores phase space in such a way that time averages along a sufficiently long trajectory are equivalent to ensemble averages [88]:

$$A_{\text{obs}} = \langle A \rangle = \int dq dp \rho(q, p) A(q, p) = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau dt A(\{r_i(t)\}, \{\dot{r}_i(t)\}) \tag{3.5}$$

In practice, the continuous time integral is approximated by a discrete average over a trajectory sampled at finite intervals:

$$A_{\text{obs}} \approx \frac{1}{n_k} \sum_{k=1}^{n_k} A_k \tag{3.6}$$

where n_k is the number of configurations collected during the simulation, and A_k is the value of the observable at the k-th frame [78]. Therefore, an average over a molecular dynamics trajectory provides an estimate of the macroscopic quantity measurable in experiments.
It is important to note that, although the ergodic theorem was originally formulated for Hamiltonian systems in the microcanonical ensemble, the same principle can be extended to other ensembles, such as NVT or NPT, provided that the dynamics is ergodic with respect to the corresponding equilibrium distribution. When ergodicity holds, time averages computed along MD trajectories remain valid estimators of ensemble averages regardless of the specific thermodynamic ensemble being simulated.

3.1.2 Description of the potential energy term

As anticipted, one of the fundamental points in MD simulations is determining a model for the potential energy term $\mathcal{U}(r^N)$. Different levels of approximation can be employed for this purpose, each offering a trade-off between computational cost and accuracy [89]. In general, more accurate representations of interatomic interactions demand higher computational resources, resulting in slower simulations. However, in many cases simplified models and approximate potentials are sufficient to reproduce the relevant dynamics with reasonable accuracy. The choice of the appropriate level of theory thus depends on the nature of the system, the properties of interest, and the available computational resources.

Ab initio molecular dynamics

In *ab initio* molecular dynamics (AIMD), the interatomic forces are computed by solving the electronic structure problem at each time step. This consists in solving the time-independent Schrödinger equation for a system of N interacting electrons:

$$\mathcal{H}_e \psi_k = E_k \psi_k \tag{3.7}$$

where $\mathcal{H}_e$ is the electronic Hamiltonian, ψ_k the many-electron wavefunction, and E_k the corresponding eigenvalue. However, exact solutions exist only for a few simple cases and various approximate methods have been developed over the years. Among these, density functional theory (see § 3.2) has become by far the most widely used approach for computing the electronic properties of a system [90, 91].

Semiempirical methods

Semiempirical approaches provide a practical solution for reducing the computational cost associated with fully *ab initio* quantum chemical methods, particularly when dealing with large molecular systems or extended simulation times. These techniques maintain a foundation in quantum

mechanics but introduce significant simplifications by omitting or simplifying electronic integrals, thereby accelerating computations. To counterbalance the loss of accuracy due to these simplifications, this class of methods incorporates empirical parameters that are fine-tuned using experimental observations or results from high-level theoretical calculations. Despite their efficiency, the performance of semi-empirical models is strongly influenced by the suitability of the parameter set for the system under analysis, which might limit their transferability [92].

Empirical force fields

At the most approximate level, empirical force fields represent the potential energy of a system through predefined analytical functions with parameters fitted to experimental measurements or quantum mechanical reference data. The simplest of these models are pair potentials, where the energy depends only on the distances between atomic pairs. Common examples include the Lennard–Jones [93] and Morse [94] potentials. Greater accuracy can be achieved by including angular-dependent terms (three-body interactions), such as in the Stillinger–Weber potential developed for covalently bonded systems [95]. A further level of sophistication is provided by many-body potentials, where the energy associated with each atom depends on the surrounding atomic environment; the Embedded Atom Model (EAM) is a well-known example widely used for metallic systems [96]. Over the past decades, hundreds of empirical force fields have been developed for specific applications. Prominent examples include CHARMM [97], AMBER [98], and OPLS [99], which are extensively used in simulations of biomolecules, polymers, and soft matter. These models enable the simulation of very large systems (up to millions of atoms) and long timescales (up to microseconds or beyond), making them suitable for studying macroscopic properties. However, this efficiency comes at the expense of accuracy, especially for processes involving bond breaking or formation like those in catalytic systems. To address these cases, reactive force fields such as ReaxFF have been developed [100], which dynamically update bonding interactions during the simulation. Despite their versatility, a major limitation of empirical force fields remains their limited transferability: their reliability decreases significantly when applied to conditions or chemical environments different from those used in their parameterization.

3.2 Density functional theory

A fundamental component of most AIMD simulations is the density functional theory (DFT), which provides a practical and widely adopted approach to solve the quantum mechanical problem underlying interatomic interactions.

As early as the 1920s, Enrico Fermi proposed that the energy of an electronic system could be expressed as a functional of the electron density $\rho(\mathbf{r})$, an object that depends only on three spatial variables [101, 102]. This idea was rigorously formalized in 1964 by Hohenberg and Kohn, who proved that the ground-state energy of an interacting electron system is uniquely determined by its ground-state electron density $\rho(\mathbf{r})$ [103].

The first Hohenberg-Kohn theorem states that there exists an energy functional $E[\rho]$ such that the ground-state energy and density of the system can be obtained by minimizing this functional.

The second Hohenberg-Kohn theorem provides a variational principle: for any trial density ρ_{trial} that satisfies the necessary conditions $\rho \geq 0$ and $\int d\mathbf{r}\rho(\mathbf{r}) = N_e$ (with N_e the number of electrons), the energy evaluated at that density is always greater than or equal to the exact ground-state energy:

$$E[\rho_{\text{trial}}] \geq E[\rho_{\text{exact}}] \tag{3.8}$$

However, Hohenberg and Kohn didn't provide an explicit form for the functional $E[\rho]$.
In 1965, Kohn and Sham introduced a practical approach to circumvent this problem by introducing an auxiliary system of non-interacting electrons moving in an effective potential $V_s(\mathbf{r})$ [104]. This fictitious system is constructed such that its electron density $\rho_s(\mathbf{r})$ exactly reproduces the electron density of the real interacting system $\rho_s(\mathbf{r}) = \rho(\mathbf{r})$. The Hamiltonian for the non-interacting system is:

$$\hat{\mathcal{H}}_s = \hat{T}_e + \hat{V}_{\text{eN}} + \hat{V}_s(\mathbf{r}) \tag{3.9}$$

where $\hat{T}_e$ is the kinetic energy operator and $\hat{V}_{\text{eN}}$ is the electron-nucleus potential. The wavefunction of this system is a single Slater determinant built from the Kohn–Sham orbitals $\{\phi_i\}$, and the electron density is given by:

$$\rho(\mathbf{r}) = \sum_{i<\text{occ}} |\phi_i(\mathbf{r})|^2 \tag{3.10}$$

The total energy functional is then decomposed as:

$$E_S[\rho] = T_S[\rho] + V_{\text{eN}}[\rho] + J[\rho] + E_{\text{XC}}[\rho] \tag{3.11}$$

where:

- $T_S[\rho]$ is kinetic energy of the non-interacting reference system;
- $V_{\text{eN}}[\rho]$ is the electron-nucleus interaction;
- $J[\rho]$ is classical Coulomb repulsion $J[\rho] = \frac{1}{2}\iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1-\mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2$;
- $E_{\text{XC}}[\rho]$ is exchange-correlation energy, which captures all the many-body effects missing from the non-interacting system, given by $E_{\text{XC}}[\rho] = T_e[\rho] - T_S[\rho] + V_{ee}[\rho] - J[\rho]$

To find the ground-state density, the Kohn–Sham orbitals $\{\phi_i\}$ must be obtained by solving the Kohn–Sham equations:

$$\hat{f}\phi_i(\mathbf{r}) = \varepsilon_i\phi_i(\mathbf{r}) \tag{3.12}$$

where the Kohn–Sham operator $\hat{f}$ is defined as:

$$\hat{f} = -\frac{1}{2}\nabla^2 + V_{\text{eN}}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + V_{\text{XC}}(\mathbf{r}) \tag{3.13}$$

Here, $V_{\text{XC}}(\mathbf{r}) = \delta E_{\text{XC}}[\rho]/\delta\rho(\mathbf{r})$ is the exchange-correlation potential, derived from the exchange-correlation energy functional. Because the exact form of E_{XC} is unknown, approximations must be used. A variety of approximate exchange–correlation functionals have been developed. A useful way to classify these functionals is through Jacob's Ladder [105], a conceptual framework in which

each “rung” represents a higher level of complexity and (ideally) accuracy. No single functional is uniformly best, so benchmarking a few candidates for the target material and observables is common before committing to production calculations. From a computational standpoint, standard Kohn–Sham DFT with semilocal functionals exhibits an approximately cubic scaling with the number of atoms N_e. This near-cubic scaling is the primary reason large DFT calculations become rapidly expensive as system size grows.

3.3 Gran canonical DFT

In the simulation of electrochemical interfaces using DFT, it is common to adopt a constant-charge approach, where the number of electrons N_e is fixed. However, this strategy does not reflect the experimental reality, in which the electrochemical potential, not the charge, is the externally controlled variable. Constant-potential simulations, based on the grand canonical ensemble (GCE), provide a more realistic description of electrochemical systems, but they are significantly more challenging from a computational point of view. Among the methods proposed to address this problem, Xia and Xiao introduced an algorithm designed to enable efficient and robust grand canonical DFT simulations of electrochemical interfaces [106]. The key idea is to iteratively adjust the number of electrons N_e at each SCF step so that the Fermi level E_{Fermi} exactly matches a target electrochemical potential μ, thus minimizing the grand canonical free energy of the system. The optimization of N_e is performed by Newton’s method:

$$N_{e,k+1} = N_{e,k} - H^{-1}(E_{\text{Fermi}} - \mu) \tag{3.14}$$

where $H = \partial E_{\text{Fermi}}/\partial N_e$. Since there is no analytical form for either H or E_{Fermi} , the authors proposed to approximate the $E_{\text{Fermi}} - N_e$ relationship using a polynomial fitting scheme [107]. By evaluating E_{Fermi} at several trial values of N_e, they construct a smooth fitted function whose derivative can be calculated analytically to provide a stable and accurate estimate of H. In this way, the optimization scheme becomes:

$$N_{e,k+1} = N_{e,k} - \text{lr} \cdot (H^{\text{fitted}})^{-1}(E_{\text{Fermi}}^{\text{fitted}}(N_e) - \mu) \tag{3.15}$$

where lr is a learning-rate parameter used to damp possible oscillations during the iteration. This algorithm is however specifically designed to be used in combination with implicit solvent models, like VASPSol. These models allow for a continuous variation of the number of electrons by providing a neutralizing background charge that maintains overall charge neutrality even when N_e changes. Hence, the algorithm proposed by Xia and Xiao cannot be directly applied in explicit solvent models, where the lack of a compensating dielectric medium would result in an unbalanced charge distribution and potentially divergent electrostatic interactions.
In order to expand the scope of this simulations, Lee introduced a hybrid explicit-implicit solvent model, termed SOLHYBRID, combined with a potential control algorithm named Target POTential (TPOT) [60]. In the SOLHYBRID model, explicit solvent molecules and ions are placed near the electrode, while the surrounding environment is treated using the continuum implicit solvent model VASPSol. The implicit solvent region is defined by a shape function $\zeta(\mathbf{r})$, which depends

on the local pseudo-electron density $n(\mathbf{r})$ as:

$$\zeta(\mathbf{r}) = \frac{1}{2}\,\mathrm{erfc}\left(\frac{\log[n(\mathbf{r})/n_c]}{\sigma\sqrt{2}}\right) \tag{3.16}$$

where n_c is a cutoff density and σ controls the smoothness of the transition. The electron density $n(\mathbf{r})$ considered in this shape function is given by:

$$n(\mathbf{r}) = n_{\mathrm{core}}(\mathbf{r}) + (1-\alpha_{\mathrm{pad}})n_{\mathrm{val}}(\mathbf{r}) + \alpha_{\mathrm{pad}}n_{\mathrm{pad}}(\mathbf{r}) \tag{3.17}$$

where $n_{\mathrm{pad}}(\mathbf{r})$ is a padding density and α_{pad} is a tunable mixing parameter. To maintain a constant electrode potential Φ, the TPOT module dynamically adjusts the number of electrons N_e in the simulation cell. The electrode potential is defined as

$$\Phi = \frac{\Phi_{\mathrm{vacuum}} - \mu_e}{e} \tag{3.18}$$

where Φ_{vacuum} is the energy level at the point far from the electrode (i.e., the vacuum level) and μ_e the electronic chemical potential. TPOT updates the number of electrons only when the deviation from the target potential Φ_{target} exceeds a threshold $\delta\Phi$.

$$N_{e,k+1} = N_{e,k} - \frac{\Phi_k - \Phi_{\mathrm{target}}}{R_i} \tag{3.19}$$

with $R_i = \partial\Phi/\partial N_e$. Together, SOLHYBRID and TPOT allow GCE-MD simulations with explicit solvation layers and controlled electrochemical potential.

3.4 Enhanced sampling

One of the persistent challenges in MD simulations is the gap between the accessible simulation timescales and the much longer timescales over which many chemical and physical relevant transformations occur. Processes such as conformational transitions in biomolecules, phase transitions during crystallization, or reactive events often take place on timescales well beyond the microsecond regime, frequently extending into seconds or longer (Fig. 3.1).

From a computational perspective, the bottleneck is severe for first-principles simulations such as DFT-based AIMD, whose computational cost scales as $O(N_e^3)$ with the number of electrons N_e. This scaling limits typical AIMD trajectories to a few picoseconds for systems of $\sim 10^2$ atoms, making it impractical to observe rare events directly. Even when using classical empirical force fields, which are orders of magnitude faster than DFT, brute-force simulations remain prohibitively expensive for sampling slow processes. While state-of-the-art specialized hardware has pushed the boundaries, for example, Shaw et al. [109] reported a 0.1 ms long MD trajectory, such simulations remain exceptional and require dedicated supercomputing resources. For most applications, the accessible timescales for unbiased MD remain in the picosecond to nanosecond range, far from the timescales relevant for many transformations.

This mismatch between simulation capabilities and physical timescales limits the sampling of

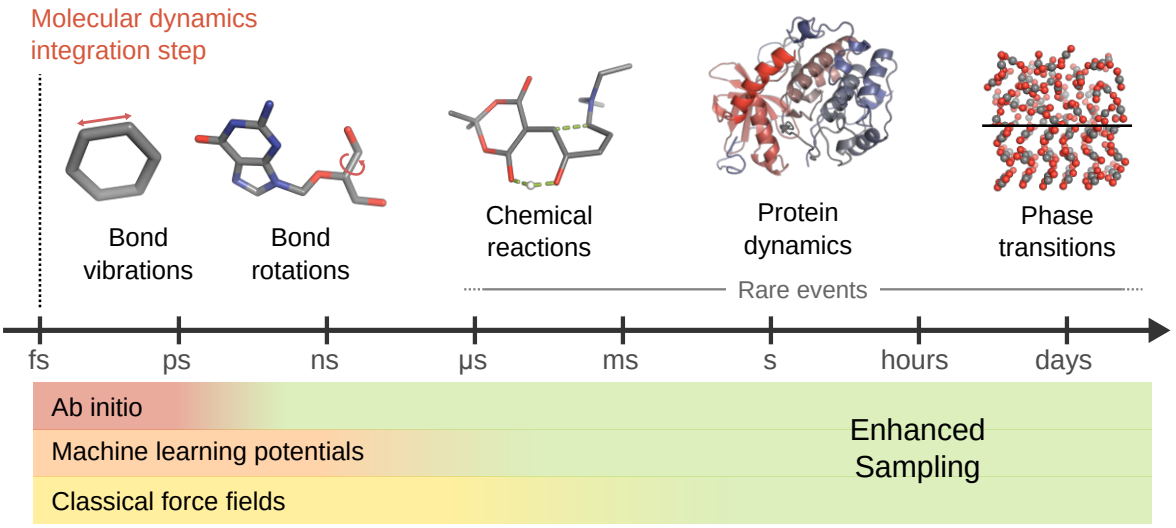

Figure 3.1: **The timescales in MD simulations.** MD simulations are typically performed with timesteps of order $\Delta t \sim fs$, whereas chemical reactions and physical transformations occur on much longer timescales. Such processes can only be accessed through the use of enhanced sampling techniques. Reprinted with permission from ref. [108]. Courtesy of Enrico Trizio.

low-probability regions of the phase space and motivates the development of enhanced sampling methods [108, 110].
In the canonical ensemble, the probability of each configuration in the phase space is related to the potential energy $\mathcal{U}(\mathbf{R})$ by the Boltzmann distribution [110]:

$$p(\mathbf{R}) = \frac{1}{Z} \exp\left(-\beta \mathcal{U}(\mathbf{R})\right) \tag{3.20}$$

where $\mathbf{R}$ represents the atomic coordinates, $\beta = 1/k_B T$ is the inverse temperature (k_B the Boltzmann constant), and $Z = \int d\mathbf{R} \exp\left(-\beta \mathcal{U}(\mathbf{R})\right)$ is the configurational partition function. Such a distribution is typically multimodal, characterized by multiple regions of high probability separated by regions of low probability. The high-probability regions correspond to metastable states, and transitions between these states constitute rare events. If an MD simulation is initiated within one metastable basin and the barrier to another basins is significantly larger than the available thermal energy, the system will remain trapped in that basin for the entire simulation time. As a consequence, only the local region of phase space around the starting configuration will be explored, and no transitions to other basins will occur. Even in the unlikely event that a transition to another basin takes place, the sampling will remain non-ergodic, as the system will not visit all relevant regions of phase space sufficient times. Under such conditions, no statistically meaningful conclusions can be drawn regarding any properties.
To effectively sample transitions between metastable states, one needs to alter the original probability distribution to enhance the exploration of the phase space. Enhanced sampling techniques allows this by modifying the probability distribution $p(\mathbf{R})$, while still being able to recover information about the original physical system.

A class of enhanced sampling techniques relies on identifying suitable *collective variables* and

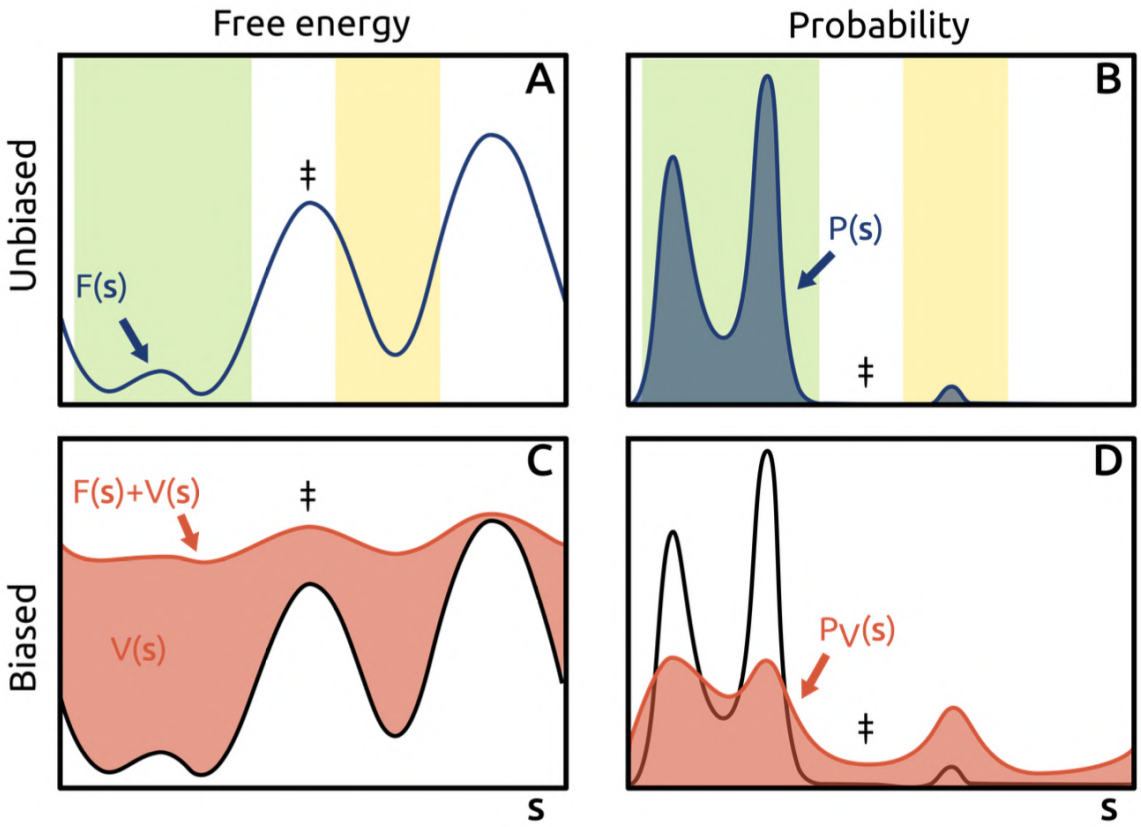

Figure 3.2: **Relation between free energy and probability distribution.** Schematic representation of the relationship between free energy (A, C) and probability distribution (B, D) along a collective variable s. In the unbiased case (top row), the free energy profile $F(s)$ shows pronounced minima corresponding to metastable states (highlighted by the green and yellow shaded regions), while the probability distribution is sharply peaked around these states and nearly vanishes in the transition state region (‡), where the free energy barriers are highest. When a suitable bias potential $V(s)$ is applied (bottom row), the effective free energy barriers are reduced. As a result, the probability distribution becomes broader and significantly populated even in the transition state region. Reprinted with permission from ref. [110]. Courtesy of Enrico Trizio.

introducing external bias potentials to enhance their fluctuations, thus facilitating transitions between metastable states. Collective variables (CVs) are low dimensional representation of the 3N dimensional Cartesian coordinate space. Typically defined as functions of atomic coordinates $s = s(\mathbf{R})$, CVs distinguish between different metastable states and capture slow degrees of freedom. The equilibrium distribution of the CV is defined as the marginal distribution of $p(\mathbf{R})$:

$$p(s) = \int d\mathbf{R}\delta\left[s - s(\mathbf{R})\right] p(\mathbf{R}) = \langle \delta\left[s - s(\mathbf{R})\right] \rangle \tag{3.21}$$

Following statistical mechanics, the *free energy surface* (FES) associated with the CV is then defined as:

$$F(s) = -\frac{1}{\beta} \log p(s) \tag{3.22}$$

If two metastable states A and B correspond to domains of the CV space, then the probability ratio can be obtained from the FES by integrating its exponential over the corresponding domains:

$$\Delta F(s) = F_A - F_B = -\frac{1}{\beta} \log \frac{p_A(s)}{p_B(s)} = -\frac{1}{\beta} \frac{\int_A ds \exp(-\beta F(s))}{\int_B ds \exp(-\beta F(s))} \tag{3.23}$$

Even if the difference between the FES of two metastable states is not affected by the choice of the CV, its profile and the free energy barrier depend on the chosen CV [111].
Once a suitable set of CVs is chosen, their fluctuations are ehnanced by introducing a bias potential $V(\mathbf{R}) = V(s(\mathbf{R}))$ that depends on s, and thus on $\mathbf{R}$ through $s(\mathbf{R})$. The idea to introduce a bias potential to enhance the sampling of the CV space can be traced back to umbrella sampling, which was first introduced in 1977 by Torrie and Valleau [112]. With this additional biasing potential, the CVs are distributed according to the following *biased distribution*:

$$p_V(\mathbf{R}) = \frac{1}{Z_V} \exp\left(-\beta \left(\mathcal{U}(\mathbf{R}) + V(\mathbf{R})\right)\right) \tag{3.24}$$

with the biased partition function $Z_V = \int d\mathbf{R} \exp\left(-\beta \left(\mathcal{U}(\mathbf{R}) + V(\mathbf{R})\right)\right)$. To retrieve the unbiased average of any observable $\mathbf{O}(\mathbf{R})$ from simulations performed under the biased potential, one uses the *reweighting* technique. Indeed, the unbiased and biased probability distributions are related through:

$$p(\mathbf{R}) \propto p_V(\mathbf{R}) \exp(\beta V(\mathbf{R})) \tag{3.25}$$

Hence, ensemble averages calculated under the original unbiased distribution can be obtained via reweighting:

$$\langle \mathbf{O}(\mathbf{R}) \rangle = \frac{\langle \mathbf{O}(\mathbf{R}) \exp(\beta V(\mathbf{R})) \rangle_V}{\langle \exp(\beta V(\mathbf{R})) \rangle_V} \tag{3.26}$$

where $\langle \cdot \rangle_V$ denotes averages over configurations generated with the biased potential. In practice, this relationship is applied using the ergodic theorem, which ensures that, for sufficiently long simulations, time averages converge to ensemble averages. Therefore, the reweighting formula in terms of simulation samples $\{\mathbf{R}_k\}$ is given by:

$$\langle \mathbf{O}(\mathbf{R}) \rangle \approx \frac{\sum_{k=1}^{n_k} \mathbf{O}(\mathbf{R}_k) \exp(\beta V(\mathbf{R}_k))}{\sum_{k=1}^{n_k} \exp(\beta V(\mathbf{R}_k))} = \sum_{k=1}^{n_k} \mathbf{O}(\mathbf{R}_k) w(\mathbf{R}_k) \tag{3.27}$$

with normalized weights defined as:

$$w(\mathbf{R}_k) = \frac{\exp(\beta V(\mathbf{R}_k))}{\sum_{k=1}^{n_k} \exp[\beta V(\mathbf{R}_k)]} \tag{3.28}$$

This reweighting procedure corresponds to correcting the weight of each sample to match the original unbiased distribution. Moreover, one can estimate the unbiased probability distribution $p(s)$ of the CV by expressing it as the ensemble average of a delta function:

$$p(s) = \langle \delta(s - s(\mathbf{R})) \rangle \approx \sum_{k=1}^{n_k} \delta(s - s(\mathbf{R}_k)) w(\mathbf{R}_k). \tag{3.29}$$

From a practical perspective, the delta function can be approximated by creating histograms weighted by $w(\mathbf{R}_k)$, allowing the direct estimation of the unbiased free energy profile $F(s)$.

3.4.1 On-the-fly Probability Enhanced Sampling

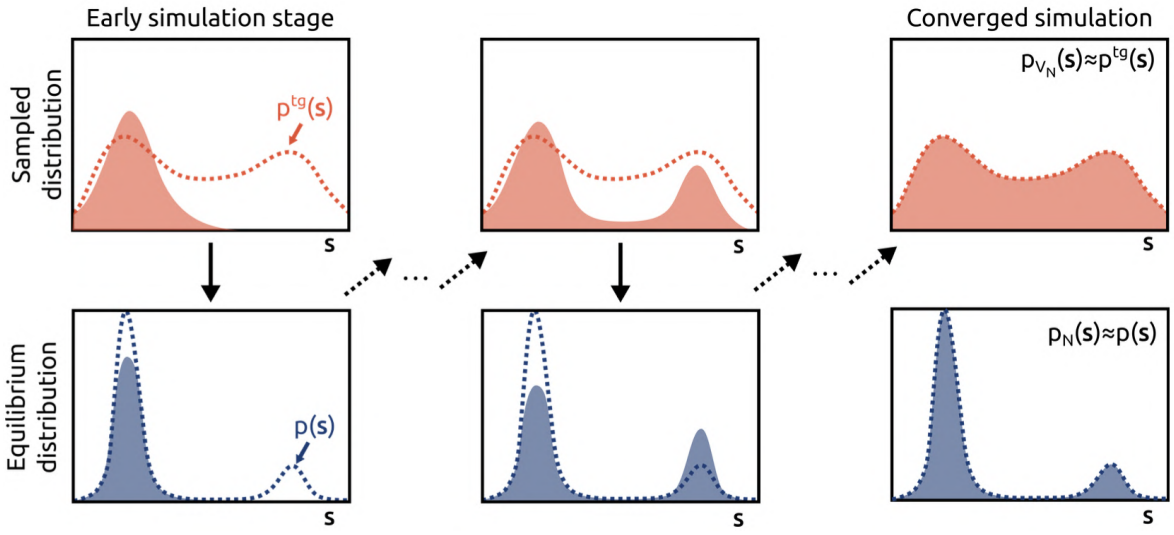

Figure 3.3: **Schematic representation of OPES.** At convergence, the sampled distribution matches the target distribution, and the estimated probability reproduces the unbiased equilibrium distribution. Reprinted with permission from ref. [110]. Courtesy of Enrico Trizio.

In the decades, many enhanced sampling schemes have been developed. In this work, we employed the On-the-fly Probability Enhanced Sampling (OPES) method to enhance the sampling efficiency of molecular dynamics simulations. OPES is a powerful evolution of metadynamics [113, 114] that was developed with the dual objective of accelerating the exploration of phase space while retaining the ability to extract accurate unbiased statistical information [115]. The fundamental idea behind OPES is to construct a bias potential V(s) that modifies the underlying sampling distribution in CV space s, such that the system explores a predefined target distribution $p^{tg}(s)$. Unlike metadynamics, where the bias is incrementally constructed from deposited Gaussians with the aim of flattening the free energy surface, OPES builds the bias based on an on-the-fly estimate of the probability distribution (Fig. 3.3). The bias potential takes the form:

$$V(s) = \frac{1}{\beta} \log \frac{p(s)}{p^{tg}(s)} \tag{3.30}$$

where $\beta = 1/k_B T$, $p(s)$ is the (unknown) equilibrium distribution, and $p^{\text{tg}}(s)$ is the desired target distribution. This formulation ensures that the bias is designed to transform the original distribution into the target one, while allowing for straightforward reweighting to recover unbiased ensemble averages. Since $p(s)$ is not known a priori, it is estimated during the simulation using a weighted kernel density estimator (KDE). OPES comes with different variants, that are hereafter described.

OPES-MetaD In OPES-MetaD, the target distribution is chosen to be the well-tempered distribution:

$$p^{\text{tg}}(s) = p^{\text{WT}}(s) \propto [p(s)]^{1/\gamma} \tag{3.31}$$

where $\gamma > 1$ is the *bias factor*, which controls the degree of barrier reduction. This target is analogous to the one sampled in Well-Tempered Metadynamics and serves to broaden the original distribution just enough to promote transitions between metastable states. The estimate of $p(s)$ at step n is obtained from the set of collected samples $\{s_k\}$ as:

$$p_n(s) = \frac{\sum_k^n w_k G(s, s_k)}{\sum_k^n w_k} \tag{3.32}$$

where $G(s, s_k)$ is a Gaussian kernel centered at sample s_k and w_k are statistical weights. To keep the number of kernels manageable, OPES employs a kernel-merging compression algorithm. The bias at step n then reads:

$$V_n(s) = \left(1 - \frac{1}{\gamma}\right) \frac{1}{\beta} \log\left(\frac{p_n(s)}{Z_n} + \varepsilon\right) \tag{3.33}$$

Here, $\varepsilon \ll 1$ is a regularization term to avoid divergences in the logarithm, and Z_n is a normalization factor calculated over the explored region Ω_n of the CV space:

$$Z_n = \frac{1}{\Omega_n} \int_{\Omega_n} p_n(s) ds \tag{3.34}$$

A key strength of OPES is that the bias builds up rapidly in an initial exploratory phase, then remains relatively stable, minimizing the transient regime and allowing for nearly the entire trajectory to be used in statistical reweighting. This feature also makes OPES a practical tool for assessing the quality of the chosen CVs: poor choices often result in slow or stalled dynamics, as the initial bias fails to drive transitions and does not improve significantly during the run. In practice, OPES-MetaD requires the tuning of only a few parameters: the pace (i.e., the frequency at which the bias is updated), the initial kernel width (which is typically estimated adaptively in the early stages of the simulation), and the barrier parameter $\Delta E \sim \gamma/\beta$, which defines the expected height of the highest free energy barrier to be overcome and the broadening of the target probability distribution.

OPES-Explore To improve robustness in cases where the CVs are suboptimal or when the focus is on rapid exploration rather than precise convergence, OPES-Explore may be employed. OPES-Explore shares the same target distribution $p^{\mathrm{WT}}(s)$ as OPES-Metad but constructs the bias from the biased (i.e., sampled) distribution, not the reweighted one. The well-tempered distribution is directly approximated by a simple kernel density estimator:

$$p_n^{\mathrm{WT}}(s) = \frac{1}{n}\sum_{k=1}^{n} G(s, s_k) \tag{3.35}$$

The bias is computed as:

$$V_n(s) = (\gamma - 1)\frac{1}{\beta}\log\left(\frac{p^{WT(s)}}{Z_n} + \varepsilon\right) \tag{3.36}$$

Because the bias is built from the sampled distribution rather than the reweighted one, it continues to evolve throughout the simulation. This dynamic behavior encourages more transitions, especially when CVs fail to capture all relevant degrees of freedom. Consequently, OPES-Explore tends to provide broader exploration, albeit at the cost of slower convergence and potentially less accurate free energy estimates in the early stages.

OPES-Flooding This variant of OPES was specifically developed to efficiently collect unbiased transition and to enable the study of transition states and kinetic properties. It combines the general OPES framework with concepts from conformational flooding and hyperdynamics, where the central idea is to enhance escape from a metastable state while preserving the unbiased character of the transition region. OPES-Flooding operates by constructing a bias potential $V(s)$ along a collective variable s that partially fills the metastable basin to increase the probability of leaving it [116, 117]. The method introduces an excluded region s_{exc}, a predefined interval in CV space corresponding to the transition state (TS) region, where no bias is applied. This ensures that the dynamics near the TS remain unaltered, enabling the generation of unbiased transition paths.
Formally, the method builds the same type of well-tempered bias as in OPES-Metad, but its deposition is suppressed when the CV enters the excluded region.

$$V(s) = \begin{cases} \left(1 - \frac{1}{\gamma}\right)\frac{1}{\beta}\log\left(\frac{p_n(s)}{Z_n} + \varepsilon\right), & \text{if } s \notin s_{\mathrm{exc}} \\ 0, & \text{if } s \in s_{\mathrm{exc}} \end{cases} \tag{3.37}$$

The unbiased time scale τ can be evaluated over the discrete MD trajectory:

$$\tau = \sum_{i}^{N_{\mathrm{tot}}} \Delta t \exp(\beta V(s_i)) \tag{3.38}$$

where Δt is the timestep of the simulation, s_i is the value of the CV at step i, and N_{tot} is total number of steps propagated to observe the transition. A single transition is not enough to properly estimate the kinetics, so multiple simulations ($n > 10$) with one-way transition must

be performed. The unbiased transition times are expected to be distributed according to a Poisson distribution. To assess the reliability of the computed kinetics, it is performed a two-sample Kolmogorov–Smirnov test between the empirical distribution of sampled transition times and the Poisson distribution with the same characteristic time [118]. Kinetics are considered reliable only when the p-value is larger than 0.05. The confidence interval of exponentially distributed data points can be estimated using the method proposed by Kaminsky [119]:

$$\left\{ \frac{2\sum_i^n \tau_i}{\chi^2_{2n}(0.975)}, \frac{2\sum_i^n \tau_i}{\chi^2_{2n}(0.025)} \right\} \tag{3.39}$$

3.4.2 Collective variables

As introduced earlier, in enhanced sampling methods, the identification of appropriate CVs is crucial to efficiently accelerate rare events and to accurately characterize the free energy landscape of the system under investigation. CVs are reduced-dimensional descriptors that capture the essential degrees of freedom associated with the slow transitions between metastable states. Formally, CVs are smooth and differentiable functions $s(\mathbf{R})$ of the atomic coordinates $\mathbf{R}$. They should fulfill several requirements: (1) they must provide a compact description of the process by projecting the complex, high-dimensional phase space onto a low-dimensional manifold; (2) good CVs should clearly distinguish between relevant metastable states; (3) ideally, CVs should encode the slow modes of the system that are responsible for the rare transitions; (4) be invariant under translational and rotational operations.
In practice, the choice of CVs is often guided by chemical intuition, aiming for descriptors that are not only effective in driving the sampling but also have a clear and straightforward physical interpretation. Tipical examples include interatomic distances (Eq. 3.40), angles (Eq. 3.41), coordination numbers (Eq. 3.42), etc...

$$d_{i,j} = |\mathbf{r}_i - \mathbf{r}_j| \tag{3.40}$$

$$\theta_{ijk} = \arccos\left(\frac{\mathbf{r}_{ji} - \mathbf{r}_{jk}}{|\mathbf{r}_{ji}||\mathbf{r}_{jk}|}\right) \tag{3.41}$$

$$s_{ij} = \frac{1 - \left(\frac{r_{ij}-d_0}{r_0}\right)^n}{1 - \left(\frac{r_{ij}-d_0}{r_0}\right)^m} \tag{3.42}$$

More sophisticated CVs can also be designed to address specific processes. For example, the problem of describing proton transfer between acid–base sites is not well captured by standard spherical coordination numbers as in Eq. 3.42. To overcome this limitation, Grifoni *et al.* introduced a set of CVs based on Voronoi tessellation, which provides a unique partition of the simulation box such that each proton can be unambiguously assigned to a specific Voronoi center [120, 121]. These CVs were successfully used to study acid-base equilibria in zeolites [26], and in bulk water [121, 122].
The crystallization problem can be studied using the CV proposed by Piaggi *et al.*, which measures

the similarity between local atomic environments and a chosen reference crystalline structure [123, 124]. This approach has also been employed to identify specific local environments, such as the χ_7 active sites on iron surfaces, as reported in ref. [38].
However, when the goal is to explore the phase space in a blind fashion, such as for reaction discovery, one cannot rely on prior knowledge of suitable CVs to escape from local minima. In such cases, a general-purpose CV should be employed. In this context, Raucci *et al.* introduced a CV based on the adjacency matrix of the atomic connectivity graph [125]. By computing the maximum eigenvalue of this matrix and using it as a CV, the method can automatically drive the exploration of different reaction pathways. This approach has been successfully applied to a variety of systems, including reactions of relevance to atmospheric chemistry [126].
There are many CVs, and the choice is clearly dependent on the system under study. A large set of CV definitions is available and well documented in the PLUMED library [127], which provides both standard descriptors (e.g. distances, angles, dihedral angles, and coordination numbers) and more specialized variables tailored to specific problems, including order parameters for crystallization, hydrogen-bond metrics, path collective variables, and local environment similarity measures. In addition, PLUMED offers tools to combine primitive descriptors into more complex functions or to implement user-defined CVs, allowing great flexibility in adapting the methodology to the needs of a particular system.

Data-driven methods

3.5 Machine learning

In this part of the thesis, we introduce the basic principles of machine learning (ML) and explore how these data-driven approaches can be integrated into atomistic simulations for catalysis [128, 129]. ML provides a powerful complement to the traditional physics-based methods presented earlier. In the following, we briefly introduce several ML-based approaches applied to atomistic simulations.
Broadly speaking, and without the ambition of being exhaustive, ML can be defined as the field of study that gives computers the ability to learn without being explicitly programmed.
An ML system typically consists of three key components: a model, a learning algorithm and a loss function. The model defines how to derive the output from the input data. In the simplest cases, the input–output relationship is represented by a parametric function $f_\theta(x) \rightarrow y$. The process of finding a "good" mapping from inputs to outputs is referred to as learning, or training. In many cases, training is achieved by minimizing the discrepancy between the model's predicted behavior and the actual observed output. This discrepancy is quantified by a loss function. In most applications, training is performed once, after which the learned parametric model can be used to make predictions on new input data. The most important concept is that of *generalization*: the model is expected to provide correct answer for data not seen during training.
Depending on the nature of the available data and the learning objective, ML can be divided into three broad categories:

Supervised Learning In supervised learning, the algorithm is trained on a dataset containing both inputs and their corresponding outputs, allowing it to learn an explicit mapping from input to output [130]. This is the most commonly used approach in atomistic simulations, where models are often trained to predict energies or forces from atomic configurations.

Unsupervised Learning In unsupervised learning, only input data is available, and the goal is to uncover hidden patterns or structures within the data [131], such as clustering similar atomic environments or reducing dimensionality [132].

Reinforcement Learning Reinforcement learning, by contrast, involves an agent that interacts with an environment and learns to make decisions by receiving feedback in the form of rewards or penalties [133]. While less common in atomistic simulations, it has shown promise in tasks like molecular design or exploring chemical reaction pathways.

3.5.1 Basic principles of supervised learning

To better understand the supervised learning approach, it is helpful to draw a parallel with the way traditional physics models operate. In physics, one typically starts with a well-established theoretical framework and known initial conditions, which one uses to predict the behavior of a system. In contrast, supervised learning begins with data: we have inputs (such as the environment or initial conditions) and corresponding outputs (such as observed behaviors), and the goal is to infer the underlying rules that map inputs to outputs, without explicitly defining them in advance. Given a dataset of N input-output pairs $(\mathbf{x}_i, y_i)$ (with $i = 0, \ldots, N-1$), in supervised learning we have the goal to learn a function $f(\mathbf{x}; \theta)$, parameterized by θ, that can accurately predict the output y from the input $\mathbf{x}$.
The collection of data $(\mathbf{x}_i, y_i)$ used to train the model is called *training set*. The objective of the learning (or training) process, is to minimize the discrepancy between the model's prediction $\hat{y}_i = f(\mathbf{x}_i, \theta)$ and the true output y_i. To quantify this discrepancy, a loss function is defined. A commonly used example is the mean squared error (MSE):

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^{N} (f(\mathbf{x}_i; \theta) - y_i)^2 \tag{3.43}$$

Minimizing the loss function allows the model to adjust its parameters to better fit the data.
We want to find the best set of parameters θ that minimize $\mathcal{L}$. Therefore, the problem that needs to be solved is:

$$\theta^* = \text{argmin}_\theta \mathcal{L}(\theta) \tag{3.44}$$

In most cases, this minimization is performed using iterative optimization algorithms, such as gradient descent:

$$\theta' \leftarrow \theta - \eta \nabla_\theta \mathcal{L}(\theta) \tag{3.45}$$

where $\eta > 0$ is the *learning rate*, a hyperparameter that controls the step size, and $\nabla_\theta \mathcal{L}(\theta)$ is the gradient of $\mathcal{L}(\theta)$ with respect to θ.
In practice, especially for large datasets, it is common to use stochastic gradient descent (SGD) or its variants (e.g., ADAM), which estimate the gradient based on a small batch of training examples rather than the entire dataset. This significantly reduces computational cost and often helps escape local minima or flat regions in the loss landscape. The training process consists of repeating these updates over multiple epochs, where in each epoch the model sees the entire training set (or a shuffled version of it).
Up to this point, we have not discussed the number of parameters in the model, which plays a fundamental role in determining its complexity. In general, the more parameters a model has, the higher its capacity to fit complex patterns in the data. However, if the number of parameters is too large relative to the number of training points, the model becomes overly flexible, fitting noise and fluctuations in the data, but its predictions on unseen data are typically poor. This results in overfitting. On contrary, if the number of parameter is too small, the model lacks flexibility and fails to capture the underlying structure of the data. This situation is described as underfitting. The number of parameters should therefore be carefully chosen to balance good fitting of the training data with good generalization to unseen data.

3.5.2 Neural networks

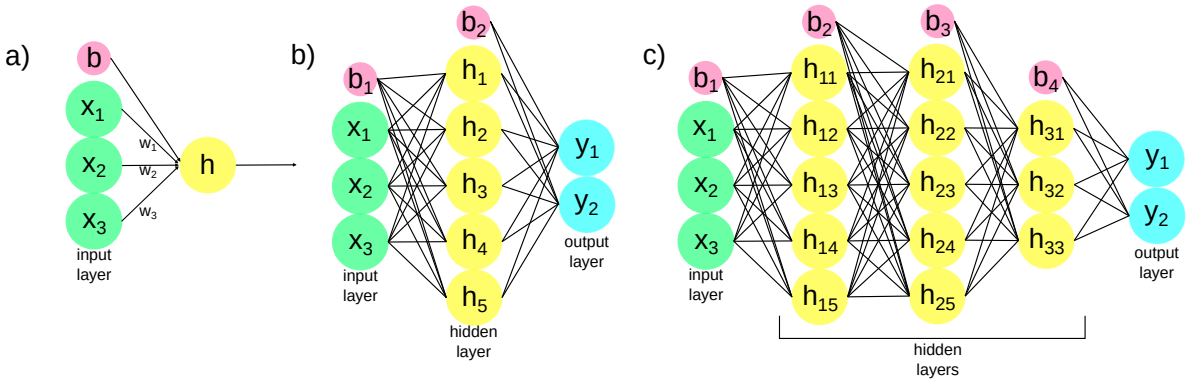

Figure 3.4: **Feedforward NN architectures**.

Neural networks (NNs) are among the most widely used and versatile models in modern machine learning [134, 135]. Their strength lies in their ability to approximate complex, non-linear functions in high-dimensional spaces. In this section, we introduce the basic structure and working principles of feedforward neural networks, the simplest and most commonly used type of NN. The purpose of a feedforward NN is to approximate an unknown function $f^*(x) \longrightarrow y$ by a parameterized model $f_\theta(x) \longrightarrow y$, where θ represents the set of learnable parameters. This approximation is achieved through a layered composition of functions, where each layer consists of a number of computational units called nodes, or neurons.
To illustrate the working principle of feedforward NNs, we can consider the simplest architecture that includes a single computational unit. An illustrative diagram to denote this single node is

depicted in Fig. 3.4a. This node takes as input x_1, x_2, x_3 and an intercept term b and outputs a value $h = a(z)$, where

$$z = \sum_{i=1}^{3} w_i x_i + b \tag{3.46}$$

w_1, w_2, w_3, b are the weight and intercept terms of the NN to be optimized during the learning process, and $a(\cdot)$ is a non-linear function. Common choices for the activation function $a(\cdot)$ are the sigmoid (Eq. 3.47), the hyberpolic tangent (Eq. 3.48), or ReLU function (Eq. 3.49).

$$\sigma(z) = \frac{1}{1 + \exp(-z)} \tag{3.47}$$

$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \tag{3.48}$$

$$\mathrm{ReLu}(z) = \max(0, z) \tag{3.49}$$

Each activation function has different non-linear characteristics, which affect both the model's expressiveness and its training behavior [136, 137]. For example, the sigmoid and tanh functions are smooth and provide a softer non-linearity. In contrast, the ReLU function introduces a stronger activation. In general, a node performs a two-step operation: a linear transformation of its input $z = \mathbf{w} \cdot \mathbf{x} + b$, and a non-linear activation $a(z)$ termed activation function. The same principle applies to more complex architectures, with arbitrary numbers of input, hidden, and output nodes. If we use an input vector $\mathbf{x}$ and increase the number of nodes in the hidden layer (Fig. 3.4b), namely we increase the width of the NN, we can write the output as

$$\mathrm{NN}(\mathbf{x}) = \mathbf{y} = \mathbf{w}_2 \cdot a(\mathbf{w}_1 \cdot \mathbf{x} + b_1) + b_2 \tag{3.50}$$

where $\mathbf{w}_1$, $\mathbf{w}_2$, b_1, b_2 are the weight matrices and intercepts associated with the hidden and output layers. According to the universal approximation theorem [138], a single hidden layer NN with a sufficiently large number of neurons can approximate any continuous function to arbitrary accuracy, under mild conditions. However, modern machine learning often uses NN with multiple hidden layers, the number of which determines the depth of the network (Fig. 3.4c). These architectures, called deep NN, have larger number of parameters to be optimized but are designed to increase expressiveness of the model to interpolate highly non-linear relationships [139]. A general deep feedforward network with L layers is defined recursively:

$$z^{(l+1)} = \mathbf{w}^{(l)} \cdot h^{(l)} + b^{(l)} \tag{3.51}$$

$$h^{(l)} = a(z^{(l)}) \tag{3.52}$$

with $l = 0, \ldots, L-1$, $z^{(0)} = \mathbf{x}$, the input vector, and $z^{(L-1)}$ the final output.
The parameters $\theta = \{\mathbf{w}, \mathbf{b}\}$ are optimized by minimizing a loss function $\mathcal{L}(\theta)$, which quantifies the discrepancy between model predictions and target values (Eq. 3.43). To train the NN, the parameters are initialized to small random values near zero. The random initialization serves the purpose of symmetry breaking, since if all the parameters start off at identical values, then all the hidden layer units will end up learning the same function of the input. The optimization

of parameters is commonly done via gradient-based optimization algorithms (Eq. 3.45), such as stochastic gradient descent (SGD) or ADAM. The gradient of the loss function with respect to the parameters $\nabla_\theta \mathcal{L}$ is computed by the backpropagation algorithm [140], which applies the chain rule to compute derivatives layer by layer, moving from the output back to the input. Each training iteration consists of:

1. **Forward propagation**: compute the network output $\mathrm{NN}(\mathbf{x}) = f_\theta(\mathbf{x})$;
2. **Loss evaluation**: compute the loss function $\mathcal{L}(y, \hat{y})$;
3. **Backpropagation**: compute $\nabla_\theta \mathcal{L}$ using the backpropagation algorithm;
4. **Parameters update**: update the parameters according to $\nabla_\theta \mathcal{L}$ with an optimization algorithm, $\theta \leftarrow \theta - \eta \nabla_\theta \mathcal{L}$ where η is the learning rate.

3.5.3 Machine learning potentials

As discussed before, AIMD is the elective tool for computing the temporal evolution of catalytic systems in which various chemical and physical transformations are involved. However, its cost limits the applicability to small systems and short simulation times. Thanks to the ability of NNs to perform as universal approximators, they can be used to interpolate the PES of a system and use this approximation to compute forces and run MD simulation at much cheaper cost. The idea is simply: given an atomic configuration, we would like to predict its energy. The first applications of neural networks to atomistic simulations date back to the '90s [141, 142]. These early models employed very simple architectures, yet they already enabled the study of realistic systems. However, such simplicity also introduced important limitations. First, the network architecture was fixed with respect to system size, meaning that the number of atoms had to remain constant. Second, fundamental physical symmetries were not incorporated into the model design.

Behler-Parrinello NNs Behler and Parrinello proposed in 2007 a conceptually simple yet powerful idea: decomposing the total energy of the system into a sum of atomic contributions [143]. Each atomic energy is predicted by an individual neural network. For a system of N atoms, the total energy is expressed as:

$$E_{\mathrm{tot}} = \sum_{i}^{N} \epsilon_i \tag{3.53}$$

To incorporate physical symmetries into neural network potentials, Behler and Parrinello introduced the concept of descriptors. Instead of using raw Cartesian coordinates as input, each atomic environment is encoded via a set of symmetry-preserving functions. In their original formulation, Behler and Parrinello proposed the Atom-Centered Symmetry Functions (ACSFs), a set of radial (Eq. 3.54) and angular (Eq. 3.55) functions designed to capture structural information about the atomic environment.

$$G^{\mathrm{rad}}_{i,\mu} = \sum_{j=1}^{N_{\mathrm{atoms}} \in R_c} \exp\left(-\eta (R_{ij} - R_s)^2\right) f_c(R_{ij}) \tag{3.54}$$

$$G_{i,\mu}^{\text{ang}} = 2^{1-\zeta} \sum_{i,j,k} \left(1 + \lambda \cos\theta_{ijk}\right)^{\zeta} \exp\left(R_{ij}^2 + R_{ik}^2 + R_{jk}^2\right) f_c(R_{ij}) f_c(R_{ik}) f_c(R_{jk}) \tag{3.55}$$

In the ACSF framework, f_c is a cutoff function which smoothly decays from one to zero as the interatomic distance R approaches the cutoff radius R_c. This ensures that contributions from atoms beyond the cutoff are negligible, and the descriptor remains localized. θ_{ijk} is the angle between the central atom i and neighbors j and k. The η, R_s, λ and ζ are the set of hyperparameters that needs to be chosen.
However, this formulation alone does not guarantee permutation invariance among atoms of the same species. To enforce this symmetry, the atomic energies for atoms of a given species are computed using neural networks that share the same weights. This not only ensures invariance under permutation of identical atoms, but also significantly reduces the number of parameters to optimize during training. The total energy of the system can thus be written as

$$E_{\text{tot}} = \sum_{i=1}^{N_{el}} \sum_{j=1}^{N_i} \epsilon_i^j \tag{3.56}$$

Since the descriptors are smooth differentiable functions, as well as the NN, the force acting on an atom can be obtained using chain rule:

$$F_\alpha = -\frac{\partial E}{\partial R_\alpha} = -\sum_{i}^{N_{\text{atoms}}} \frac{\partial \epsilon_i}{\partial R_\alpha} = -\sum_{i}^{N_{\text{atoms}}} \sum_{j}^{N_{\text{sym},i}} \frac{\partial \epsilon_i}{\partial G_{ij}} \frac{\partial G_{ij}}{\partial R_\alpha} \tag{3.57}$$

Forces are computed on-the-fly during training and used as an objective in the loss function, that is defined as follows:

$$\mathcal{L} = \frac{1}{N_{\text{struct}}} \left[\left(E_i^{NN} - E_i\right)^2 + \frac{\beta}{3N_{\text{atoms}}} \sum_{j=1}^{3N_{\text{atoms}}} \left(F_{ij}^{NN} - F_{ij}\right)^2 \right] \tag{3.58}$$

with β a parameter chosen to balance between the energy term and the forces term.

DeepMD-kit After the Behler-Parrinello NN, advances have been made. A popular ready-to-use package for MD simulations with MLP is the DeepMD-kit [144, 145].
In this framework, the atomic coordinates of each atom i are first transformed into an embedding matrix $\mathcal{R}^i$ that contains vectors $\mathbf{r}_{ji} = \mathbf{r}_j - \mathbf{r}_i$, with atom j beeing within a cutoff r_c from atom i. The embedding matrix is then transformed into reduced coordinates

$$\hat{x}_{ji} = \frac{s(r_{ji}) x_{ji}}{r_{ji}} \tag{3.59}$$

(the same for $\hat{y}_{ji}$ and $\hat{z}_{ji}$), where $s(r_{ji})$ is the following smoothing function:

$$s(r_{ji}) = \begin{cases} \frac{1}{r_{ji}} & r_{ji} < r_{cs} \\ \frac{1}{r_{ji}} \left[\frac{1}{2}\cos\left(\pi \frac{r_{ji} - r_{cs}}{r_c - r_{cs}}\right) + \frac{1}{2}\right] & r_{cs} < r_{ji} < r_c \\ 0 & r_{ji} > r_c \end{cases} \tag{3.60}$$

where r_{cs} is a smooth-cutoff parameter. The resulting matrix $\tilde{\mathcal{R}}$ of reduced coordinates is invariant under rotation and translation. To ensure permutational invariance, the $\tilde{\mathcal{R}}$ matrix is further transformed into a feature matrix $\mathcal{D}^i$:

$$\mathcal{D}^i = (\mathcal{G}^{i_1})^T \tilde{\mathcal{R}}^i (\tilde{\mathcal{R}}^i)^T (\mathcal{G}^{i_2}) \tag{3.61}$$

where $\mathcal{G}$ are embedding matrices given by

$$(\mathcal{G}^i)_{jk} = (G(s(r_{ji}))_k \tag{3.62}$$

Where G is the local embedding network mapping a single input to $M1$ outputs (resp $M2$). $\mathcal{D}^i$ is thus a $M1 \times M2$ matrix that is flattened into a vector to be given as input of the atomic energy network.
The energy and the forces are computed using a neural network with a Behler-Parrinello structure and the loss function commonly used is the following:

$$\mathcal{L} = \frac{1}{|\mathcal{B}|} \sum_{i \in \mathcal{B}} p_e |E_i - E_i^{NN}|^2 + p_f |F_i - F_i^{NN}|^2 \tag{3.63}$$

where p_e and p_f are prefactors chosen to balance between energy and force terms and $\mathcal{B}$ is the batch. The optimization algorithm generally used is the ADAM optimizer.

MACE: an equivariant message passing neural network Another important class of NNs is graph neural networks (GNNs) [146]. These architectures represent atomic configurations as graphs, where atoms are treated as nodes and interatomic interactions as edges. GNNs are increasingly used in the development of MLPs due to their greater flexibility and expressiveness [147]. Among this class of NN, MACE has emerged for its remarkable data efficiency and for excellent in-domain and out-of-domain predictions [148, 149]. MACE builds upon the framework of equivariant message passing neural networks (MPNNs). Each node i in layer t is described by a tuple

$$\sigma_i^{(t)} = (\mathbf{r}_i, z_i, h_i^{(t)}) \tag{3.64}$$

where $\mathbf{r}_i$ are the atomic coordinates, z_i the atomic number, and $h_i^{(t)}$ the feature tensor. The feature tensors are designed to be equivariant under 3D rotations:

$$h_{i,kLM}^{(t)} \left(Q \cdot \{r_j\}\right) = \sum_{M'} D_{MM'}^L(Q)\, h_{i,kLM'}^{(t)} \left(\{r_j\}\right) \tag{3.65}$$

where $Q \in O(3)$ is a rotation applied to the atomic coordinates, and $D^L(Q)$ is the Wigner D-matrix associated with angular momentum L.
The message passing process in MACE consists of three main phases repeated over multiple layers. In the message construction phase, atom i receives information from its neighbors $j \in \mathcal{N}(i)$

$$m_i^{(t)} = \bigoplus_{j \in \mathcal{N}(i)} M_t \left(\sigma_i^{(t)}, \sigma_j^{(t)}\right) \tag{3.66}$$

where $\bigoplus$ denotes a summation over neighbors, and M_t a learnable message function. In the update phase, the feature tensors are updated through another learnable function U_t:

$$h_i^{(t+1)} = U_t(\sigma_i^{(t)}, m_i^{(t)}) \tag{3.67}$$

Finally, in the readout phase, the energy contribution of each atom is obtained by applying a readout function R_t to the state at each layer:

$$E_i = \sum_t R_t\left(\sigma_i^{(t)}\right) \tag{3.68}$$

The strengths of MACE rely on the ability of the model to capture complex atomic environments with a minimal number of layers, significantly reducing the number of trainable parameters and the risk of overfitting. Furthermore, in conventional message-passing architectures, the receptive field of an atom, namely the region of the system that influences its predicted properties, grows with the number of NN layers. However, MACE achieves a large and expressive receptive field even with shallow architectures by constructing higher-order, many-body features through tensor products of equivariant atomic representations. This design allows MACE to incorporate complex local geometries and interactions involving multiple neighbors simultaneously. These characteristics, combined with remarkable computational efficiency, enable the training of MACE models on large and chemically diverse datasets, paving the way for the development of transferable and general-purpose interatomic potentials. As a result, MACE is well-suited for constructing foundational models, which can provide accurate predictions across a wide range of materials and chemical systems [150, 151].

3.5.4 Caveats

Most machine learning interatomic potentials (MLPs) rely on the locality approximation, where the total energy is decomposed into a sum of atomic contributions based on local environments within a cutoff radius between 6Å and 10Å. While this approach enables high accuracy and computational efficiency for many materials, it inherently neglects long-range interactions such as electrostatics, dispersion forces, and non-local charge transfer effects. As a result, systems where these interactions play a crucial role, such as polar materials, ionic compounds, π-conjugated systems, ect.. , may not be accurately described by these standard MLPs. Extending the cutoff radius or increasing the number of message-passing steps (in GNN-based models) can improve coverage, but comes with higher computational cost and increased complexity in fitting high-dimensional descriptor spaces. More recent generations of MLPs have addressed part of this issue by explicitly including long-range electrostatics using environment-dependent atomic charges and newer MLPs have been developing to account for global charge equilibration, allowing charges to depend on the full system rather than just the local environment [152, 153, 154]. These developments mark significant progress, but incorporating accurate and transferable long-range interactions remains an open challenge and an active area of research.

One of the fundamental limitations of MLPs is their tendency to perform better at interpolation than extrapolation. Since these models are trained on a finite set of reference data, their accuracy is confined to the regions of phase space well represented in the training set. If, during a MD

simulation, the system explores unseen or poorly sampled regions, the MLP may yield unphysical predictions of energy and forces, leading to unstable dynamics. To mitigate this issue, the training set must ideally include all relevant atomic environments that the system might encounter. However, in complex systems, such as those involving chemical reactions, phase transitions, or heterogeneous interfaces, *a priori* knowledge of the relevant region of phase space is often unavailable. This uncertainty can limit both the exploration of new mechanisms and the reliability of the simulations. Even so-called "universal" MLP models, trained on very large and diverse datasets, may fail when applied to systems that differ significantly from those represented in their training data. As such, careful validation and, when possible, active learning or on-the-fly retraining strategies are crucial to ensure the robustness and generalizability of MLPs in practical applications.

3.5.5 General procedure for training set construction

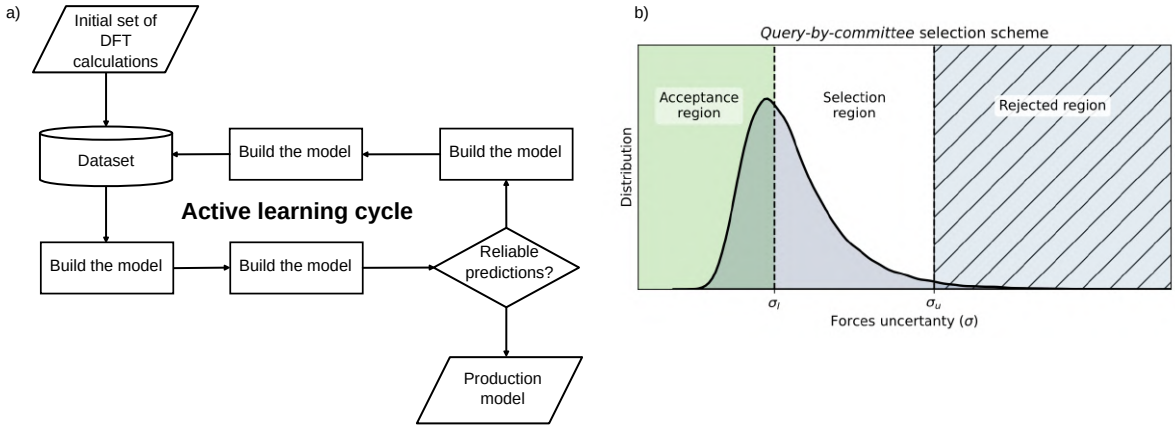

Figure 3.5: **Active learning scheme**. a) Workflow for the active learning scheme adopted in a general procedure for training set construction. b) Acceptance criteria for configuration selections based on the maximum uncertainty on forces computing over an ensemble of models.

Accurately modeling complex chemical reactions, such as heterogeneous catalytic processes, requires a detailed description of bond formation and cleavage, while accounting for the dynamic behavior of the system. AIMD is ideally suited for this task. However, its substantial computational cost limits its applicability to small systems and short simulation times.
Since the pioneering work of Behler and Parrinello, MLPs have emerged as a promising alternative, offering a balance between the accuracy of *ab initio* methods and the efficiency of classical force fields. These potentials are trained to reproduce energies and forces from a dataset of reference calculations. By leveraging the locality of the energy decomposition, they can be trained on small systems and then applied to much larger ones, enabling simulations over nanosecond timescales [155].
Despite their advantages, MLPs are inherently better at interpolation than extrapolation. This means their predictive accuracy is limited for configurations that lie outside the training set. As a

result, developing reliable MLPs for complex, reactive, multicomponent systems remains challenging. It requires a comprehensive training dataset that includes a diverse set of chemical species, ranging from reactants and products to transition states, across all relevant reaction pathways [156].
Sampling configurations near transition states, particularly those involving high energy barriers, is computationally demanding. Enhanced sampling techniques, such as metadynamics and OPES, are instrumental in overcoming this limitation. These methods accelerate the exploration of phase space, enabling efficient collection of reactive configurations and generation of uncorrelated structures that span the full reactive landscape, including all key transition regions.
One general procedure to build an effective training set can be briefly summarized as follows:

1. **Initial Geometry Collection**. The training procedure begins with the generation of initial atomic configurations, typically extracted from many short AIMD trajectories.

2. **DFT Labeling and Initial Training**. The collected structures are labeled using single-point DFT calculations to obtain accurate energies and atomic forces. This labeled dataset forms the basis for training an initial ML potential.

3. **Active Learning and Model Refinement**. To iteratively improve the model, an active learning strategy is employed based on the *query-by-committee* approach (Fig. 3.5). In this framework, M independent models are trained using the same dataset but with different random initializations (seeds). Usually $M = 4$ models are trained. One of these models is used to run MD simulations and generate new candidate configurations. The remaining models are used to estimate the model uncertainty by computing the standard deviation of the predicted forces for each atom:

$$\sigma = \max_i \left(\sqrt{\frac{1}{M} \sum_{\alpha=1}^{M} ||F_i^\alpha - \overline{F}_i||} \right) \tag{3.69}$$

   To select configurations, the distribution of σ values across the sampled dataset is analyzed, and two thresholds are defined: a lower bound σ_l and an upper bound σ_u. The lower threshold σ_l is typically set slightly above the peak of the uncertainty distribution, thereby defining an acceptance region. The upper threshold σ_u is introduced to filter out nonphysical configurations and is generally defined as $\sigma_u = \sigma_l + (0.2 \sim 0.3)\text{eV Å}^{-1}$. Configurations with uncertainty values within the interval $[\sigma_l, \sigma_u]$ are randomly selected for DFT relabeling.

4. **Ab Initio Relabeling**. The selected high-uncertainty configurations are then relabeled via new DFT single-point calculations. The resulting energies and forces are added to the training dataset to improve the model's coverage and accuracy.

5. **Dataset Expansion and Final Model Construction**. Steps 3 and 4 are repeated iteratively, gradually expanding the dataset and refining the ML potential. Through this active learning cycle, the model becomes increasingly accurate and generalizable, achieving an accuracy comparable to the reference DFT method.

Sampling the full configurational space in molecular systems, especially regions corresponding to transition states and reactive pathways, is challenging due to the presence of high free-energy barriers. These barriers make such configurations rare and difficult to observe in standard, unbiased molecular dynamics simulations. To overcome this limitation and ensure comprehensive coverage of the potential energy surface, enhanced sampling techniques are integrated into the procedure in steps 1 and 3. Methods like metadynamics, umbrella sampling, and OPES are particularly effective in this context. They accelerate the crossing of energy barriers, enabling the generation of rare, high-energy configurations that are crucial for capturing reactivity. At the same time, they reduce temporal correlations between successive configurations, thereby improving the diversity and quality of the dataset. Among these methods, OPES has proven especially effective in systematically exploring configuration space, allowing the collection of uncorrelated structures that span all relevant metastable regions.

3.5.6 Data-efficient protocols for training set construction

The previous section described the general procedure for constructing a training set, which begins with running multiple AIMD simulations. However, configurations are collected in a way that ensures they are separated by a certain time interval, typically on the order of a few picoseconds, in order to reduce correlations between data points and minimize redundancy in the training set. As a result, a substantial portion of the AIMD trajectory is discarded. Moreover, the active learning strategy based on the *query-by-committee* approach adds configurations solely through the maximum force deviation of a single atom, without explicitly accounting for the similarity of its local atomic environment. As a result, many structurally similar configurations may be selected, potentially unnecessarily enlarging the training set and ultimately reducing the data efficiency of the overall process.

In the case of reactive systems, such as those encountered in heterogeneous catalysis, the number of configurations required to train an accurate MLP can become exceedingly large. This is due to the intrinsic complexity of the potential energy surface and the necessity to sample a broad chemical space, encompassing various reactive intermediates, transition states, surfaces, and adsorbates. In such scenarios, the training set may easily grow to hundreds of thousands of configurations. Moreover, this number can increase significantly if one aims to include structural defects, which are crucial for realistically modeling catalytic systems under operating conditions. As a result, the number of configurations can grow explosively, highlighting the urgent need for more data-efficient sampling strategies that prioritize chemical and structural diversity while minimizing redundancy.

To enable the accurate and data-efficient training of MLPs for chemically disordered and large-scale materials, Dobbelaere et al. [157] introduced a cluster-based learning strategy. The key idea is to focus on local atomic environments and to identify underrepresented environments in the model's feature space, specifically, those the model is more uncertain about. Uncertainty is estimated by training a Gaussian Mixture Model (GMM) on the internal features of the MLP. Environments that are far from the high-density regions of the GMM are considered poorly represented. These are then extracted as small molecular clusters that preserve the local geometry and chemical interactions of the original environment. These clusters are then used for *ab initio*

calculations, making the process far more efficient than working with full periodic systems. Integrated into an active learning loop, this strategy allows the training set to be expanded only with the most informative and diverse structures, reducing redundancy and improving data efficiency. Another approach proposed by Perego and Bonati [43] uses Gaussian processes (GPs) to learn the potential energy surface on-the-fly and then to identify novel local environments. One of the core strengths of GPs is that they naturally provide a measure of uncertainty for every prediction [158]. The protocol consists in a Exploration and a Convergence step. In the Exploration step, the PES of reactants is learned on-the-fly using the sparse GP implementation in FLARE [159]. The GP model is used to perform uncertainty-aware OPES-flooding simulations to discover reactive pathways without prior knowledge of reaction mechanisms. When the GP's uncertainty exceeds a given threshold, new configurations are labeled with DFT and the GP model is subsequently retrained. While GPs are highly effective in the early stages of data collection, they exhibit important limitations when applied to large and complex datasets. Specifically, their computational cost scales cubically with the number of training points, which becomes prohibitive as the dataset grows. Moreover, GPs rely on kernel-based similarity measures and typically assume a stationary kernel, which can limit their ability to model highly non-linear relationships in high-dimensional spaces. These factors reduce their scalability and flexibility in PES learning compared to neural networks. To overcome these limitations, the Convergence step of the protocol consists in training a GNN model, specifically, MACE. GNN architectures are more flexible than GPs and have demonstrated remarkable data efficiency, enabling robust simulations as early as hundreds/thousands of training points. Once trained, the GNN is used to run longer MD simulations that sample a wide range of reactive events. During these GNN-driven simulations, a novel active learning scheme called *Data-Efficient Active Learning* (DEAL) is applied. In this scheme, a *query-by-committee* approach is first used to identify configurations with high uncertainty based on the disagreement between multiple independently trained GNNs. From this pool, configurations are pre-selected based on uncertainty and then filtered using the local uncertainty from a retrained GP to identify a non-redundant subset for DFT labeling. This two-step DEAL approach significantly reduces the number of required DFT calculations, only 5% of the *query-by-committee* suggestions are retained, while achieving uniform accuracy across the entire reactive landscape.

Recent "foundation" or universal MLPs have emerged thanks to large DFT datasets and GNNs such as MACE-MPo [151], CHGNet [160], M3GNet [161], and MatterSim [162], enable robust simulations across diverse systems. However, while these models often yield stable and qualitatively correct MD, they can lack quantitative reliability for predictive use, especially for targets outside the training distribution or when the underlying DFT level is insufficient. So, adaptation to specific systems is typically required. This naturally motivates transfer learning, where information in a pretrained model is reused and specialized on small, system-specific datasets, via strategies such as fine-tuning, or multi-head schemes. Very recently, franken was introduced as a lightweight transfer-learning framework for MLPs that reuses the node features of a pretrained GNN, e.g. universal models such as MACE-MP, as atomic fingerprints [163, 164]. These features are mapped to energies and forces through random Fourier features (RFF), a non-linear transformation that approximates a given kernel function as a scalar product in a suitable transformed space: $k(h, h') \approx \phi(h)^T \phi(h')$. The RFF layer uses fixed, randomly sampled parameters, leaving only

the linear weights to be optimized. This can be done in an exact way, without the need for gradient descent. This design merges the strengths of GNN representations (transferability, accuracy) with those of kernel methods (closed-form optimization, data efficiency), while avoiding their main drawbacks, which are expensive training and poor scaling with dataset size, respectively.

3.5.7 Machine learning and enhanced sampling

ML is not limited in fitting the PES but can be employed to design CVs [110, 108], expanding its scope in atomistic simulations. We recall that CVs are functions of the atomic positions and must be continuous and differentiable. In order to be effective, a CV should also perform dimensionality reduction by capturing the slow degrees of freedom of the system and distinguishing between different metastable states. Additionally, CVs should respect the relevant symmetries of the physical process. Traditionally, CVs are designed based on chemical and physical intuition. While such descriptors offer interpretability, they may fail to capture the complexity of highly dynamic environments, such as those encountered in heterogeneous catalysis or conformational changes in biomolecules, where the active site can undergo significant structural changes on the same timescale as the reaction.
In such cases, ML can assist in identifying or constructing a more effective set of CVs, tailored to the underlying dynamics of the system. ML-based approaches allow for a data-driven design of CVs, potentially revealing hidden patterns or slow modes not easily accessible through intuition. A typical ML framework for CV discovery includes four essential components:

Input representation In order to represent the system, a transformation of atomic coordinates, or a subset of, can be used as input for the model function. The transformation should ensure that the inputs are invariant or equivariant under certain symmetry group transformations. Alternatively, several physical descriptors can be used as input for the ML model.

Model Function The CV is defined as a parametric function of the input features. The model can be either linear or non-linear. Examples of linear models are principal component analysis (PCA) [165], linear discriminant analysis (LDA) [166, 167], and time-lagged independent component analysis (TICA) [168]. Non-linear model can be obtained with feed-forward NN, autoencorders, or GNN.

Data acquisition The quality and diversity of the data used for training are critical to the success of any ML-based CV [169]. To accurately capture the relevant physics, the dataset should include representative configurations from all metastable states and transition regions. Missing regions may lead to poor generalization and unreliable behavior in enhanced sampling applications. While data quantity matters, it is often the quality and coverage that determine the CV's effectiveness. To address this, an iterative exploration strategy is often adopted: in the initial stage, aggressive sampling methods can be used to broadly explore the phase space and identify relevant states. In a second stage, more focused, near-equilibrium simulations help refine the

data and improve its consistency. This combination allows for the construction of CVs that are both accurate and robust, providing a solid foundation for enhanced sampling.

Learning Objective In essence, we need to define what we want the CV to learn from the data. Different methods target different objectives, depending on the available data and the purpose of the simulation. From a high-level perspective, these objectives can be grouped into three increasingly specific categories: (1) *Dimensionality reduction*: the CV should compress the system's dynamics into a low-dimensional space while retaining the most relevant information; (2) *State discrimination*: the CV should assign distinct metastable states to separate regions of its space, avoiding overlaps; (3) *Dynamical relevance*: the CV should capture the system's slow modes, which govern rare transitions between metastable states. In enhanced sampling, the most desirable learning objective is capturing the system's slow modes, as this directly accelerates rare transitions. However, this often requires high-quality, reactive trajectories, which are typically unavailable at the start. This creates a chicken-and-egg problem: good CVs are needed to explore the relevant phase space, but such exploration is necessary to construct good CVs. To overcome this, simpler surrogate objectives, such as dimensionality reduction or state discrimination, are often used initially. While these may lead to suboptimal CVs, they can still be effective and serve as a basis for iterative refinement as better data become available.
A number of works can be found in literature where ML-based CVs were used to study complex molecular and reactive systems [170]. The implementation of ML-based CV is greatly facilitated by the use of the mlcolvar library.

Beyond their use in constructing CVs, ML techniques can also be employed to design bias potentials. A particularly elegant example is the Kolmogorov bias, introduced by Kang and Trizio [171, 172], a method aimed at promoting the sampling of the transition state (TS) region rather than the metastable basins. This approach is rooted in the committor function $q(\mathbf{x})$, which defines the probability that a trajectory starting from configuration x will reach the product state B before returning to the reactant state A. While computing the exact committor is generally intractable, it can be approximated using a neural network trained as a classifier, with boundary conditions $q(\mathbf{x}) = 0$ in state A and $q(\mathbf{x}) = 1$ in state B. A key observation is that the gradient norm $|\nabla q(\mathbf{x})|^2$ is sharply peaked near the TS region, where the system has equal probability of committing to either metastable state. Based on this, a bias potential can be defined as a function of $|\nabla q(\mathbf{x})|^2$, for instance:

$$V(\mathbf{x}) = \lambda \log\left(1 + |\nabla q(\mathbf{x})|^2\right) \tag{3.70}$$

where λ controls the strength of the bias. Unlike methods such as metadynamics, which fill the basins to accelerate transitions, the Kolmogorov bias lowers the free energy of the TS region, effectively transforming it into a local minimum in the biased landscape. This facilitates targeted sampling of configurations near the TS, which is particularly useful for studying reaction mechanisms.
This Kolmogorov-type bias has also been adapted for reaction discovery, where the aim is to escape an initial metastable basin A without prespecifying products or hand-crafted collective variables. The procedure is simple: from a short unbiased trajectory confined to A, a neural network

is trained to output a scalar field $I(\mathbf{x})$ that is approximately constant on the observed configurations. As a consequence, $|\nabla I(\mathbf{x})|^2 \approx 0$ within A, but becomes finite where the model must extrapolate—precisely at the frontier of the sampled region. One then applies a nonlocal Kolmogorov bias

$$V_K(\mathbf{x}) = -\frac{\lambda}{\beta} \log\left(|\nabla I(\mathbf{x})|^2 + \epsilon\right) + \frac{\lambda}{\beta} \log \epsilon \tag{3.71}$$

whose maximum is 0 inside A and which is negative beyond it, thereby lowering effective barriers and nudging trajectories toward unexplored configurations. Unlike basin-filling approaches (e.g., metadynamics, or OPES), this strategy leverages the limited generalization of neural networks as a feature: the gradient of I acts as an automatic detector of "where we have not yet been" producing a bias that is agnostic to the choice of descriptors. In practice, the procedure is iterated: explore, retrain I on the newly visited data, and reapply the bias. In this way, multiple metastable states can be discovered sequentially without prior knowledge of end states or reaction coordinates.

Chapter 4

A bulk-phase transformation drives ammonia synthesis on barium hydride

In chapter 2, it was introduced the broader conceptual framework that motivates our investigation: the shift from a static view of catalysis toward a dynamic picture, in which the catalyst undergoes structural and chemical transformations under *operando* conditions.
This dynamic picture is particularly relevant for materials like barium hydride, which has recently emerged as a promising transition metal-free catalyst for ammonia synthesis in chemical looping processes. In this chapter, the results of our simulations are presented [28].

4.1 Introduction

Barium hydride (BaH_2) has been recently experimentally investigated showing an ammonia production rate comparable to that of some of the more traditional transition metal catalysts [173, 174]. Our interest in this system has been driven by several factors. Firstly, there is a pressing demand for more sustainable and efficient alternatives to the well-established but environmentally impactful Haber-Bosch process [32, 175]. Secondly, it is rather surprising that a transition metal-free ionic compound can be an active catalyst. Even more remarkably, in order to obtain significant ammonia yield, BaH_2 has to be alternately exposed to N_2 and H_2 in a chemical looping process [176]. A microscopic understanding of this phenomenon is still lacking, but the very fact that it takes place, suggests that the exposure to N_2 changes in a dramatic way the catalyst structure, leading to a compound that stabilizes nitrogen fixation intermediates, and catalyzes the successive hydrogenation step.
Using state-of-the-art molecular dynamics simulations powered by enhanced sampling methods and driven by *ab initio* quality machine learning-based interatomic potentials, we provide a detailed microscopic explanation of the experimentally observed catalytic behavior. Our simulations show that the initial step in the catalytic cycle is facilitated by the spontaneous release of hydrogen molecules from the surface at the *operando* temperature, which leave behind anionic vacancies acting as reducing agent during the N_2 adsorption process. Furthermore, the temperature induces a high diffusivity of the hydride anions. Subsequently, when the catalyst is exposed

to N_2 in the first step of the looping process, a series of intermediate reactions occurs, resembling those hypothesized for the nitrogenase enzyme [177]. This results in the formation of the mixed compound, $BaH_{2-2x}(NH)_x$, in which the Ba^{2+} cations maintain their crystallographic positions whereas the hydrides are replaced by a mixture of hydrides and imides. This compound is stable at the *operando* temperature, and it is characterized by an high mobility of both hydride and imide anions. At the end of this first nitridation step, the compound is ready to be exposed to H_2, and we have observed that the imides are easily hydrogenated, leading to the formation of ammonia. The superionic bulk behaviour of the mixed compound is crucial for its catalytic activity and for the regeneration of the catalyst after the H_2 looping cycle. This scenario could be considered an extreme version of the one suggested by Schlögl, where the transformation involves not just the catalyst surface, but the entire system [9].

4.2 Results

As we have recently shown in other catalysts (see chapter 2), analyzing their bulk and surface dynamics is crucial to the understanding of the catalytic activity. For this reason, we start our investigation by examining the dynamics both in the bulk and on the (120) surface, which is its most stable surface [174].

4.2.1 Bulk dynamics

Below the temperature of 770K, BaH_2 crystallizes into an orthorhombic structure (space group Pnma) while at higher temperature it becomes hexagonal (space group $P6_3/mmc$) [178, 179, 180, 181, 182]. In the orthorhombic arrangement, barium cations (Ba^{2+}) form a slightly distorted hexagonal close-packed (hcp) lattice in which the H^- anions occupy six of the twelve tetrahedral sites and all six octahedral sites. In the literature, these nonequivalent H-sites are referred to as H1 and H2 but, in order to remark their different coordination, we refer to these anions as H_T and H_O, respectively. Each Ba^{2+} cation is coordinated by five H^- anions that form distorted squared pyramids of formula BaH_5 where four vertexes are occupied by H_T and the remaining vertex is occupied by one H_O. The H_T sites are shared among adjacent BaH_5 pyramids while H_O are located on the unshared vertexes (Fig. 4.1a-b). In the absence of defects, only half of the possible twelve tethraedral sites in the hcp unit cell are occupied by H_T anions. This leads to the presence of tethraedral voids that permit anisotropic oscillations of H_T from their crystallographic positions. Indeed, from our simulations of stoichiometric bulk, a difference in H_T and H_O dynamical behavior is observed. In particular, the distribution of atomic displacements of H_O is isotropic while those of H_T are elongated along the x-direction, reflecting the existence at the H_T sites of a geometric frustration in agreement with neutron scattering measurements [181] (Fig. 4.1c-d-e). Like other alkali and alkaline earth metal hydrides, BaH_2 has the tendency to dehydrogenate leaving behind anionic vacancies [183, 178, 174]. In such a defected system, H^- can now diffuse via a vacancy mechanism [181, 184] that involves jumps between H_T and H_O sites (Fig. 4.2). The electronic properties of the system are also modified. Stoichiometric BaH_2 has a band gap and behaves as a semiconductor. When anionic vacancies are present, electrons are called in to restore charge

neutrality replacing the missing H^-. Thus, this compound is more properly described as an electride [185] with formula $[Ba_xH_{2x-2y}]^{2y+}(e^-)_{2y}$. This system is characterized by the presence of an electronic state near the Fermi level in which the electrons occupy the overlapping Ba^{2+} *5d* orbitals (Fig. 4.3). As described in the following section, when this electronic state is localized on the surface, it plays a crucial role in activating N_2 toward the NH_3 formation.

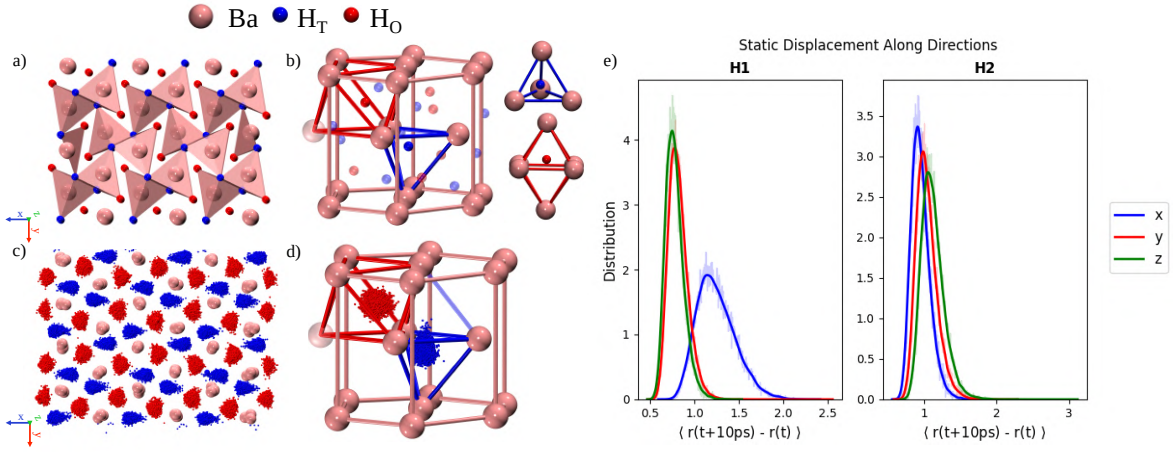

Figure 4.1: **Bulk structure of BaH_2**. a) Structure of orthorhombic BaH_2 at 0K. b) Distorted hcp unit cell of Ba^{2+} cations. Hydrides occupying tetrahedral (H_T) and octahedral (H_O) sites are shown in blue and red, respectively. c) Scatter plot of H_T and H_O atom positions in orthorhombic BaH_2 at 550 K. Configurations are reported every 1 ps during a 5 ns long trajectory. d) Anisotropic oscillations of H_T and $H_O$anions highlighted in the hcp unit cell. Configurations are reported every 1 ps during a 5 ns long trajectory. e) Static displacements of H1 (H_T) and H2 (H_O) during a 28 ns long simulation of stoichiometric bulk at 550K. The simulation cell contains 1152 atoms.

4.2.2 (120) surface dynamics and N_2 adsorption

The morphology of the (120) surface at 0K is shown in Fig. 4.4. This surface is characterized by a periodic repetition of a particular complex polyhedral arrangement of Ba^{2+}. This motif results from the truncation of octahedral and tetrahedral of the bulk building blocks. Within each of these cavities, four H^- are contained. We still classify these hydrides as H_T and H_O depending on the environment that they would have had in the perfect lattice. A crucial feature of this structure is a square pyramid oriented toward the bulk (Ba^{2+} 1-5 in Fig. 4.4), resulting from an octahedron missing one vertex, and hosting a H_O hydride. This central square pyramid shares two of its edges with two triangles (Ba^{2+} 4,5,6 and 3,4,7, respectively), each derived from tetrahedrons that have lost one vertex and accommodate a H_T ion. Additionally, the square pyramid shares one of its vertices with an irregular polyhedron (Ba^{2+} 4,6,7,8) formed from an octahedral unit in which two vertices are missing, and that also contains one H_O.
Experimentally, one finds that the catalytic activity increases by a factor of 20 from 550K to 700K [173, 174]. For this reason we examine the surface behaviour at these two temperatures. The morphology of the surface is maintained at 550K. In fact, only small fluctuations of Ba^{2+} around

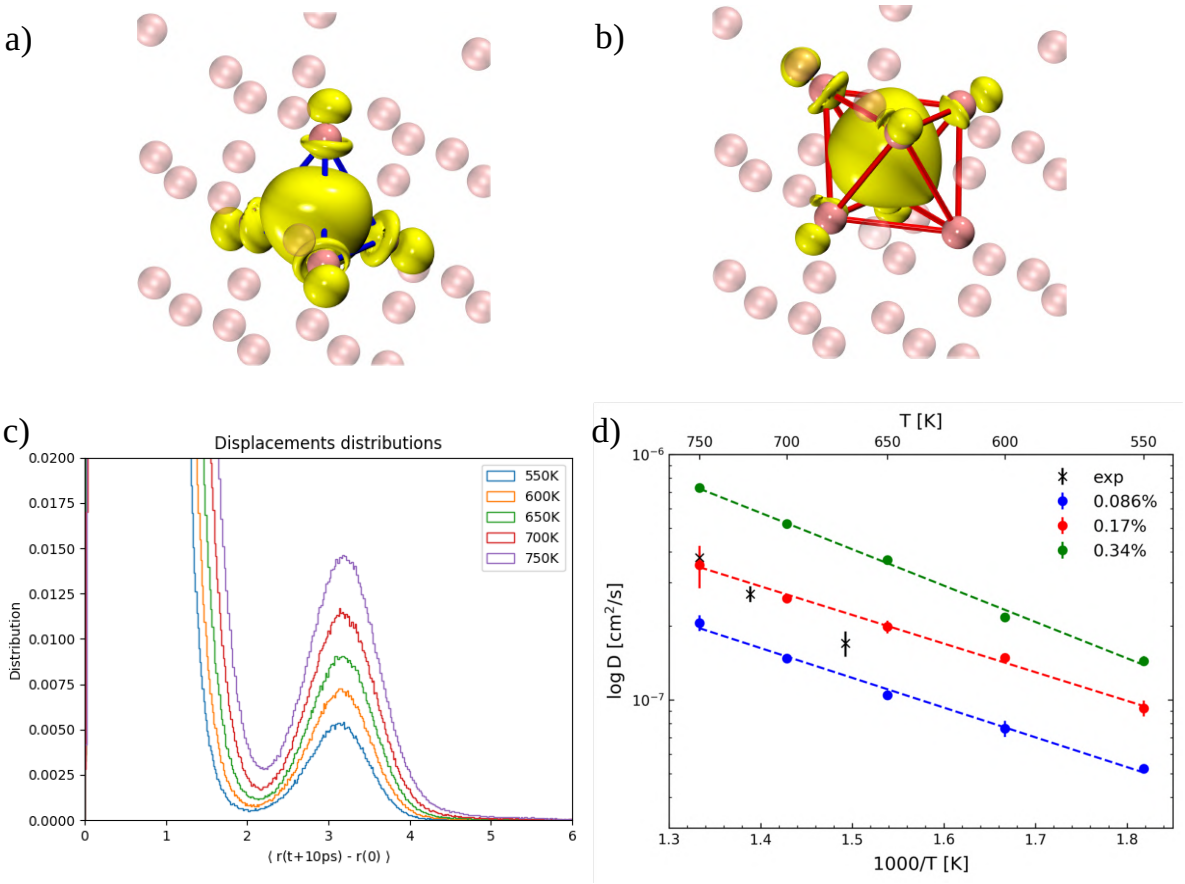

Figure 4.2: **Bulk dynamics.** (a) Charge density isosurface localized in a tetrahedral site after H_T removal, showing the overlap of Ba *5d* orbitals. (b) Charge density isosurface localized in an octahedral site after H_O removal, showing the overlap of Ba *5d* orbitals. (c) Distributions of H^- displacements at different temperatures. The simulation box contains 1150 atoms and has a concentration of H^- vacancies of 0.17%. This vacancy concentration has been chosen as a representative case. Each trajectory is $\sim$ 21 ns long. The displacements of H1 and H2 atoms were computed using OVITO [186]. (d) Diffusion coefficients at different temperatures and concentrations of H^- vacancies. The experimental values are taken from ref. [181].

the equilibrium positions are observed while H^- can diffuse with a diffusion coefficient finite but small. In contrast, at 700K some H^- and Ba^{2+} are promoted to the adlayer. In the adlayer, H^- anions are highly diffusive while Ba^{2+} cations oscillate around their putative hcp positions, occupying the missing vertexes of the tetrahedral and octahedral structures (Fig. 4.5).

As anticipated in the previous section, BaH_2 tends to release H_2 and form anionic vacancies. The enhanced H^- anions fluctuations at high temperature promote the release of hydrogen following the reaction

$$2\,H^- \longrightarrow H_2 + 2\,e^-$$

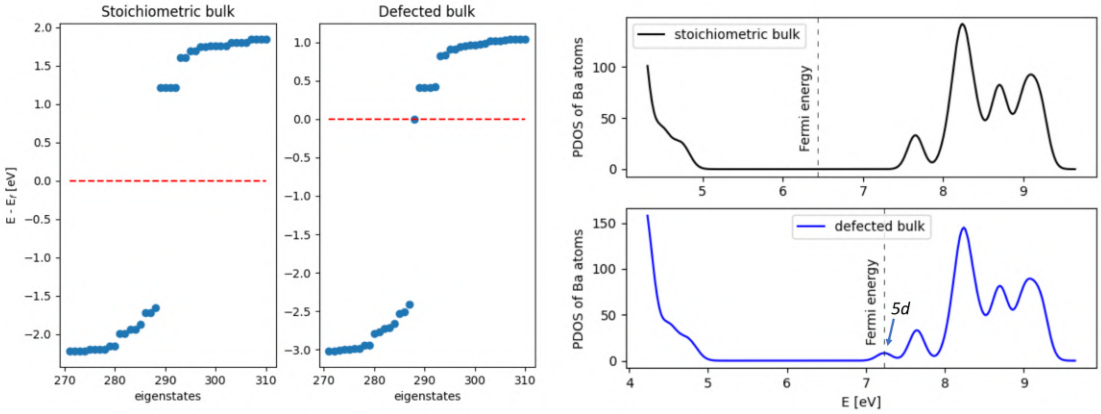

Figure 4.3: **Electronic properties of defective bulk.** Distribution of eigenstates for the stoichiometric and defected bulk. For the defected bulk, a single H^- is missing. Projected Density of States (PDOS) for Ba cations in both the stoichiometric and defected bulk. For the defected bulk, a single H^- is missing.

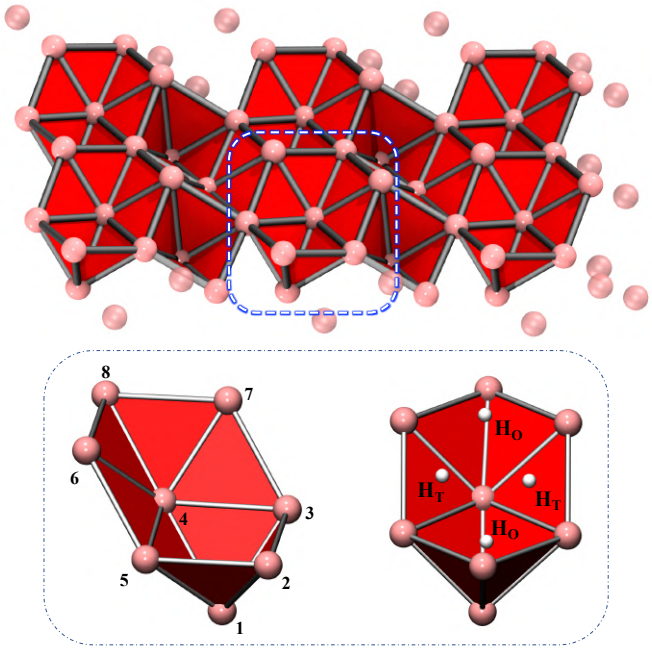

Figure 4.4: **(120) surface of BaH_2.** Upper panel: top view of the pristine (120) surface at 0K. The framed area highlights a representative cavity, while the red surfaces and bonds limit the polyhedra composing the cavities. Lower panel: arrangement of Ba^{2+} ions forming a complex polyhedral cavity. Ba and H atoms are colored pink and white, respectively.

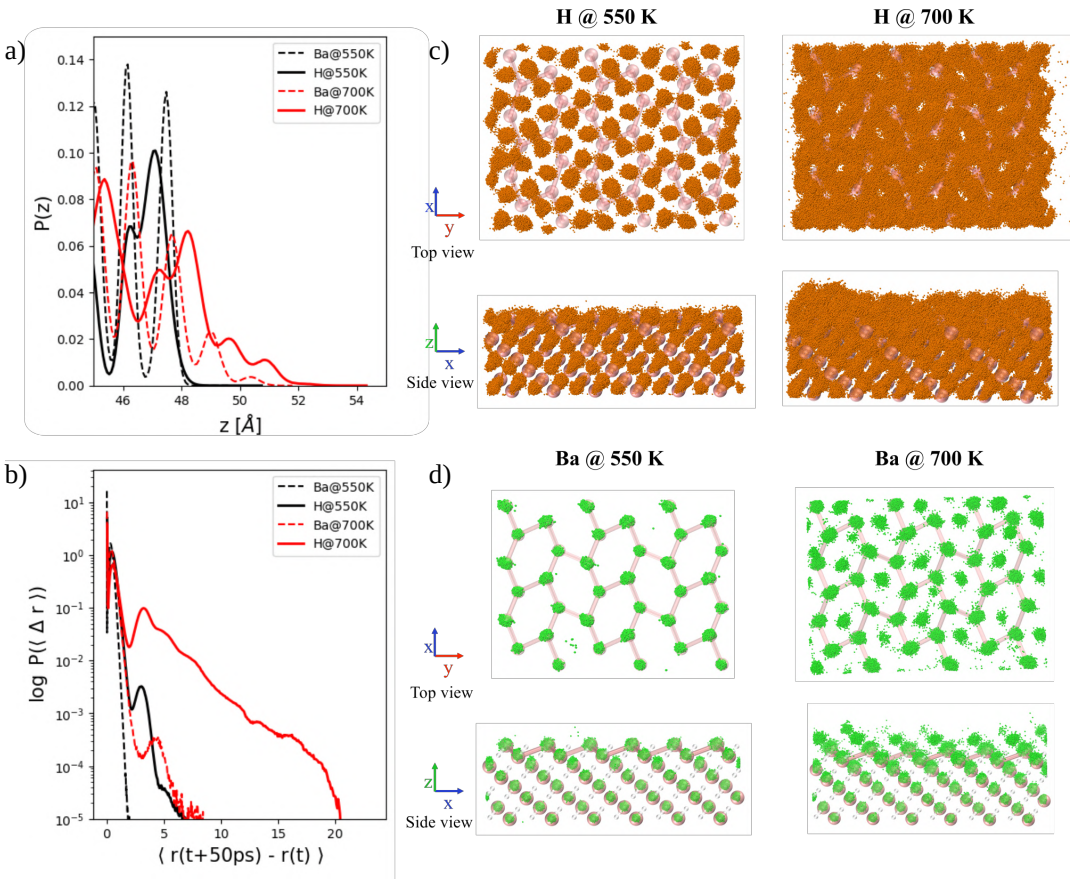

Figure 4.5: **Dynamics of the (120) surface**. (a) Distribution of Ba and H atoms along the z-direction at 550K and 700K. (b) Distribution of displacements for Ba and H at 550K and 700K. These are computed from a 300 ns long simulation. (c) Scatter plot of H^- atom positions at 550K and 700K. Configurations are reported every 10 ps during a 300 ns long trajectory. (d) Scatter plot of Ba^{2+} atom positions at 550K and 700K. Configurations are reported every 10 ps during a 300 ns long trajectory.

as depicted in Fig. 4.6. Interestingly, the two reacting hydrides are situated in the surface cavity. This reaction is facilitated by the Lewis acid character of the Ba^{2+} cations, that can use their $5d$ orbitals to accept the released electrons leading to a molecular orbital localized inside the cavity. If the atoms were in their perfect crystallographic positions, these electrons would be shared among the d orbitals of the Ba^{2+} ions surrounding the vacancy. However, due to thermal fluctuations and the high mobility of H^- , this ideal scenario is rarely observed, and the anionic electrons are typically localized in the d orbitals of just one or two barium cations.
Furthermore, these loosely bound electrons can be easily donated to electrophilic species [187, 188, 189], thereby acting as a hard Lewis base during the N_2 fixation process. The formed vacancy has notable reducing capability, similar to the anionic vacancies observed in related alkali and alkaline earth metal halides [190, 191, 192].

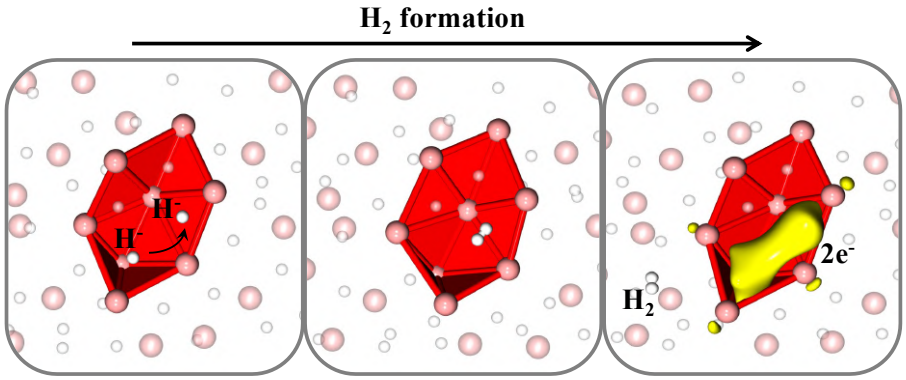

Figure 4.6: **H_2 release from the pristine (120) surface.** Schematic representation showing the release of H_2 from the (120) surface according to $2\,H^- \longrightarrow H_2 + 2\,e^-$. The red area highlights the cavity in which the reaction takes place, while the yellow isosurface indicates the charge density associated to the electride. Ba and H atoms are coloured pink and white, respectively.

The role of this state can be immediately understood if we initiate the reactive process by exposing the surface to a nitrogen molecule simulating the N_2-looping cycle. The N_2 molecule is not absorbed on the pristine (120) surface. Instead, if the defect state is present on the surface, it can interact with the empty LUMO orbital of N_2 leading to the formation of a diazenido anion (N_2^{2-}) (Fig. 4.7a). This anionic intermediate sits horizontally in the cavity previously occupied by the loosely bound electrons, and it is typically stabilized by five Ba^{2+} cations. When sitting in this conformation the N−N bond lengthens from 1.1Å to 1.2Å acquiring an extra charge of $q \sim 1.1$, as measured by the Bader charge decomposition (Fig. 4.7b). This finding explains why the looping process is necessary, and why exposing the surface to a mixture of N_2 and H_2 leads to a low ammonia yield. In fact, if one exposes BaH_2 to H_2, the important electron-rich defective states would be occupied by the hydrogen molecules, thereby poisoning the surface and making it catalytically inactive to N_2 fixation [174, 196]. The formation of N_2^{2-} is the first step in a process that leads first to the weakening of the N≡N triple bond and eventually, after a series of reduction steps, to its breaking.

In the first of these steps the diazanetriide anion (NNH^{3-}) is formed (Fig. 4.7a). This reaction can be interpreted by examining the frontier molecular orbitals. In fact, upon adsorption, the N_2^{2-} LUMO orbital is oriented in such a way that it can be attacked by a nucleophilic H^- present within the cavity. The high mobility of the hydrides facilitates this step, ensuring that the surface cavity is always occupied by reacting H^-. Additionally, the Ba^{2+} cations help also to stabilize this intermediate. A further lengthening of the N−N bond (1.4 Å) takes place after this process and the total Bader charge of the nitrogen atoms increases to a value of $q \sim 2.1$ (Fig. 4.7b).

We perform a detailed characterization of this process by calculating the free energy surface (FES) for nitrogen adsorption and diazanetriide formation at 550K and 700K, using the change in Bader charges [197, 198] of N atoms as a sensitive indicator of electronic structural variations during the reaction. With the increase in temperature, we observe a decrease in the barrier for the formation of NNH^{3-} from N_2^{2-} of about 4.3 $k_B T$ going from 550K to 700K. This is related to the greater

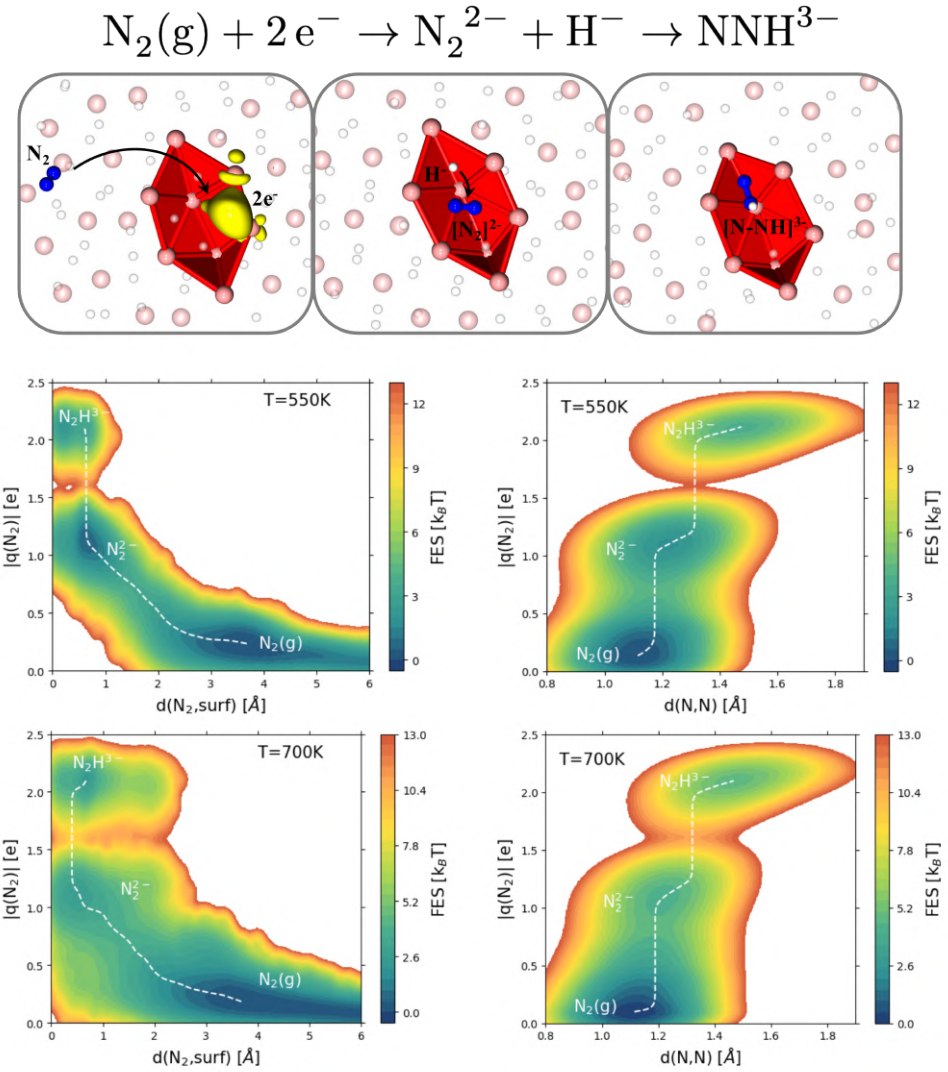

Figure 4.7: **N_2 adsorption on the activated (120) surface.** a) Schematic representation of the adsorption of N_2(g) in the cavity and the following reaction with superficial H^- to form NNH^{3-}. The red area highlights the cavity in which the reaction takes place, while the yellow isosurface indicates the charge density associated to the electride. b) Free energy surfaces (FESs) at 550K for the adsorption of N_2(g) and NNH^{3-} formation. In the left panel the FES is projected along the distance between N_2 and the surface (d(N_2, surf)) and the sum of N atoms Bader charges ($|\ q(N_2)\ |$). In right panel, the FES is reported as a function of the N−N bond distance (d(N_2)) and the sum of N atoms Bader charges. The minimum free energy pathway is shown as white dashed line and it is computed using MEPSA (Minimum Energy Path Surface Analysis) package [193]. Bader charges are predicted with a neural network, as detailed in the Method section. The d(N_2, surf) is computed using the Alpha-Shape method [194, 195], as implemented in OVITO [186]. Ba, H and N atoms are coloured pink, white, and blue, respectively.

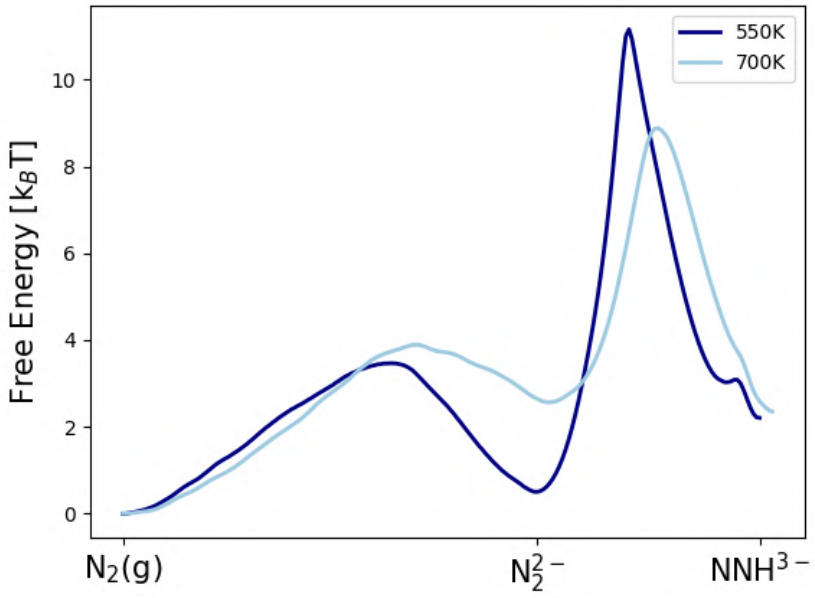

Figure 4.8: **Minima free energy pathways**. Minima free energy pathways at 550K and 700K in k_BT. These are computed using MEPSA (Minimum Energy Path Surface Analysis) package [193].

disorder of the surface which destabilize the N_2^{2-} state, thus lowering the barrier. More details can be found in Fig. 4.8.

4.3 Cleavage of the N−N bond and imide formation

The successive step is the cleavage of the N−N bond, leading to the formation of imide (NH^{2-}) anions. The diazanetriide HOMO orbital is oriented to facilitate the next reductive protonation reaction in which an H^- ion reacts with NNH^{3-}, according to

$$NNH^{3}- + H^- \longrightarrow HNNH^{2-} + 2\,e^-$$

[199] (see Fig. 4.9a). This reduction process is favored by the Ba^{2+} cations that also in this case act as Lewis acid accepting the released electrons in a localized state resulting from the overlap of their $5d$ orbitals. The formation of the diazanediide anion ($HNNH^{2-}$), effectively reduces the N−N bond order, which extends further its length to 1.5 Å. Interestingly, a competing pathway leading to NNH_2^{2-} (where two protons are bound to the same N atom) has not been observed in our simulations, emphasizing the preference for an alternating-like mechanism.
The newly formed diazanediide lies horizontally on the surface cavity stabilized by five barium cations. Following this step, $HNNH^{2-}$ can be further reduced by the anionic electrons, which can populate its antibonding σ^* orbital thus facilitating the cleavage of the N−N bond. This process

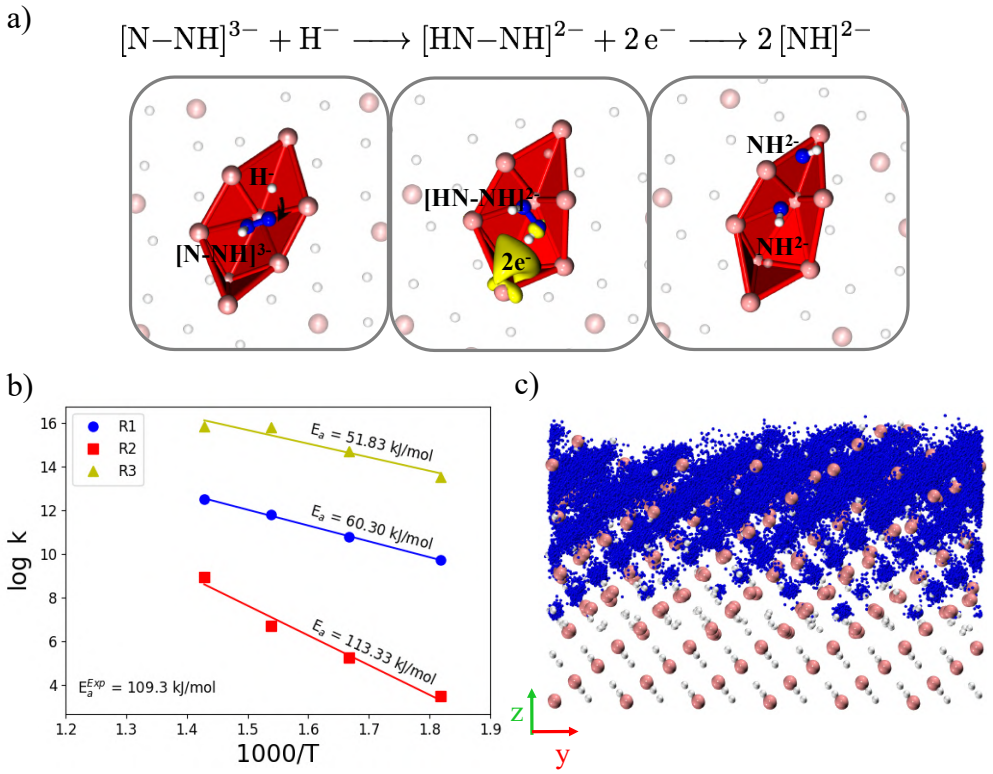

Figure 4.9: **N–N bond cleavage**. a) Schematic representation of the reduction of NNH^{3-} to form two NH^{2-}. The red area highlights the cavity in which the reaction takes place, while the yellow isosurface indicates the charge density associated to the electride. b) Arrhenius plot for the different steps involved in the nitridation process computed with the OPES flooding method. The activation energy (E_a) is obtained from the slope of the logarithm of the kinetic constant (k) versus 1000/T. The experimental value of the global activation energy (E_a^{Exp}) for the process is taken from ref. [173]. We refer to the following reactions: R1: $N_2(g) + 2\,e^- + H^- \longrightarrow NNH^{3-}$; R2: $NNH^{3-} + H^- \longrightarrow HNNH^{2-} + 2\,e^-$; R3: $HNNH^{2-} + 2\,e^- \longrightarrow 2\,NH^{2-}$. We combined N_2 absorption and NNH^{3-} formation into a single step to compute the kinetics, as the shortened lifetime of the N_2^{2-} intermediate at elevated temperatures makes it challenging to accurately determine the reaction rate for the second step ($N_2^{2-} + H^- \longrightarrow NNH^{3-}$). c) Scatter plot of N atom positions in the imidized surface at 700K. Configurations are reported every 1ps during a 60ns long trajectory. To simplify reading, a single instantaneous configuration is chosen to represent Ba and H atoms. Ba, H and N atoms are coloured pink, white, and blue, respectively.

results in the formation of two imide ions as depicted in (Fig. 4.9a).

$$HNNH^{2-} + 2\,e^- \longrightarrow 2NH^{2-}$$

By using the on-the-fly probability enhanced sampling flooding method ($OPES_f$), we estimated the kinetic rates of the three steps leading to the N-N bond cleavage at various temperatures.

The activation energy for each step, was then evaluated from the slope of the Arrhenius plot (Fig. 4.9b). The reduction of NNH^{3-} to $HNNH^{2-}$ is the rate-determining step with the highest activation energy, of about 110 kJ/mol. In contrast, the formation of NNH^{3-} and the dissociation of $HNNH^{2-}$ proceed with a lower and similar activation energy, of around 60 and 50 kJ/mol, respectively. Our activation energy is comparable with the one experimentally reported for the nitridation process [173, 174].
Thus far, it appears that we are dealing with a standard catalytic process involving a well-defined catalytic site. Its structure can be viewed as a complex exposed cavity permeable to H^-, and works as a sort of nanoreactor where the catalytic steps take place. Indeed, the undercoordinated hydrides present inside the cavity serve as reactants in various reactive events.
However, these steps alone cannot explain the working of this catalyst, which involves a significant change in chemical composition and dynamics, as hypothesized by Schlögl and discussed in the following section.

4.3.1 Formation and characterization of the hydride-imide compound after N_2 exposure

The NH^{2-} ions can diffuse across the surface and penetrate into the bulk (Fig. 4.9c), where they can be hosted in octahedral sites (Fig. 4.10a). As the exposure to N_2 continues, a gradual replacement of the hydrides by imides occurs, leading to the formation of a mixed compound with general formula $BaH_{2-2x}(NH)_x$. Hence, the reaction that occurs during the N_2 looping process can be written as

$$2\,BaH_2 + xN_2 \longrightarrow 2\,BaH_{2-2x}(NH)_x + xH_2$$

Experimental evidences indicate that the nitridation yield is $\sim$ 50% at 700K and it reaches a plateau of $\sim$ 65% at 750K [173]. For this reason, we have investigated the bulk properties of $BaH_{2-2x}(NH)_x$ at 700K, considering x values of 0.125, 0.25, and 0.50. Our findings indicate that the cell remains stable even after permitting the cell size to vary using the Parrinello-Rahman Lagrangian [200]. The most striking feature of this compound is its superionic behavior. Both H^- and NH^{2-} are highly mobile (Fig. 4.10b), with H^- exhibiting a diffusion coefficient of $\sim 10^{-5}$ cm^2/s, comparable to that of water at room temperature (Fig. 4.10c). The diffusion mechanism is similar to that observed in pristine BaH_2, with H^- moving from tetrahedral to octahedral sites.
This mixed phase is stabilized at the operando temperature by entropic effects associated to the high diffusivity of hydrides and imides. At lower temperature, the diffusivity decreases (Fig. 4.11) and the phase coexistence of barium hydride and imide is likely to be expected [201] in line with the X-ray diffraction patterns recorded at room temperature [173].
The superionic behavior plays a critical role in enhancing hydrogen release during the nitridation process. Experimental findings indicate an increase in hydrogen release during the N_2 looping cycle, exceeding that observed when increasing the temperature under Ar flow [173, 174]. Our simulations show that the marked increase in the H^- mobility in $BaH_{2-2x}(NH)_x$ significantly facilitates the formation of H_2 in agreement with experiments. In particular, we performed $OPES_f$ simulations (computational details are reported in the Methods section) observing that the transition time (τ) for hydrogen release decreases with the rising concentration of imides on the surface

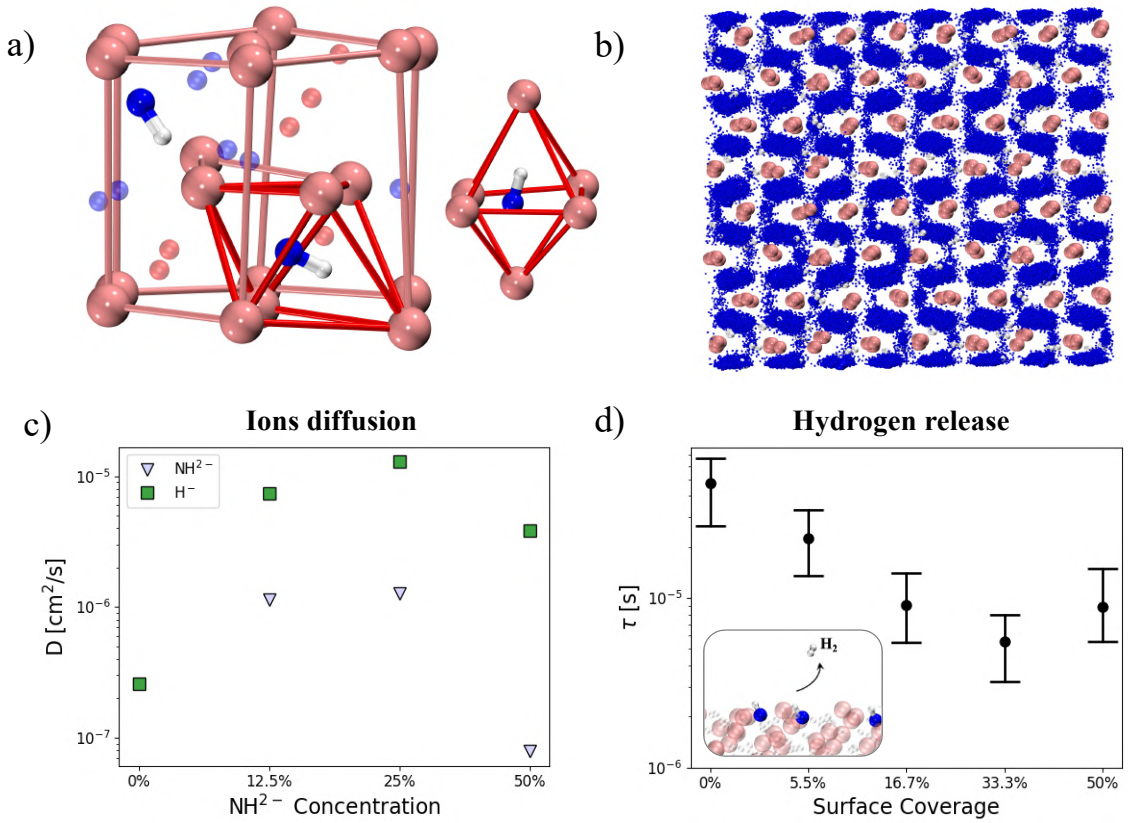

Figure 4.10: **Structure of the $BaH_{2-2x}(NH)_x$ mixture**. a) Hcp unit cell for $BaH_{2-2x}(NH)_x$. In the reported example, two octahedral sites are occupied by two NH^{2-} anions. The red bonds highlight the octahedral coordination for the NH^{2-} anion. b) Scatter plot of N atom positions in the bulk $BaH_{2-2x}(NH)_x$ at 700K with x= 0.125. Configurations are reported every 10ps during a 90ns long trajectory. c) Diffusion coefficients (D) of H^- and NH^{2-} at 700K for different concentration x of imides. The error bar falls within the size of the data point. We have performed 4 replicas for each concentration (Table 4.1). d) Transition times (τ) for H_2 release at different values of NH^{2-} surface coverage, computed using OPES flooding method. The uncertainties of τ were estimated by computing the 95% confidence intervals as reported in ref. [117]. We have performed 20 replicas for each value of surface coverage (Table 4.1). The inset shows a simplified representation of the H_2 formation and release from the imidized surface. Ba, H and N atoms are coloured pink, white, and blue, respectively.

(Fig. 4.10d). This indicates a faster kinetics, although this trend is not strictly linear at higher surface coverage. Furthermore, comparing the transition times for H_2 release and $HNNH^{2-}$ formation (the rate-determining step of the nitridation process, with a τ of 1.3e-4 s at 0% surface coverage), hydrogen release is consistently faster than the reductive protonation leading to $HNNH^{2-}$.

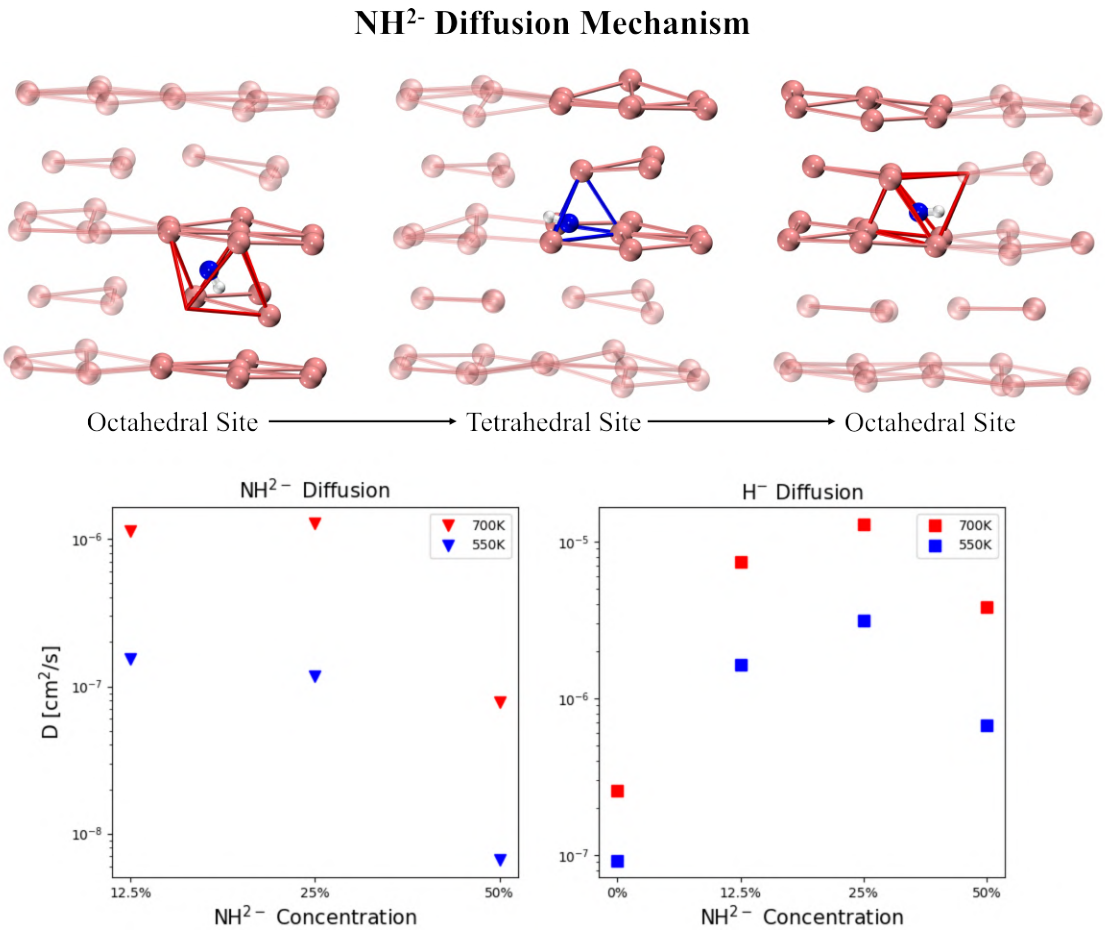

Figure 4.11: **Characterization of $BaH_{2-2x}(NH)_x$** (Top) Diffusion mechanism of NH^{2-} in the mixed compound $BaH_{2-2x}(NH)_x$. The H^- ions are omitted for clarity. (Bottom) Diffusion coefficients for H^- and NH^{2-} at 550K and 700K in the mixed compound $BaH_{2-2x}(NH)_x$ at different x values. The error bar falls within the size of the data point.

4.3.2 Hydrogenation and NH_3 release

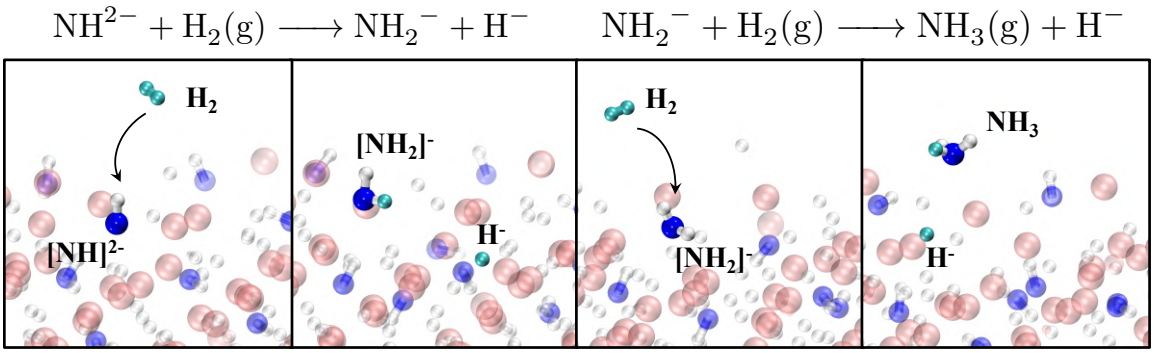

Figure 4.12: **Hydrogenation of $BaH_{2-2x}(NH)_x$ and ammonia release.** Schematic representation of the hydrogenation of NH^{2-} to NH_2^- and the following hydrogenation to NH_3. Ba, H and N atoms are coloured pink, white, and blue, respectively. The H_2 molecules introduced during the hydrogenation cycle are colored in cyan.

After the formation of this new solid state compound $BaH_{2-2x}(NH)_x$, we proceed to simulate the next stage in the looping process, which involves the exposure of the catalyst to H_2. In the initial phase of this step, imides on the catalyst surface are transformed into amides, while a hydride ion is simultaneously released according to:

$$NH^{2-} + H_2 \longrightarrow NH_2^- + H^-$$

At this stage the catalyst becomes a mixture of NH^{2-}, NH_2^- and H^- (Fig. 4.12).
The newly formed amides can be further hydrogenated leading to the formation of ammonia, following the reaction:

$$NH_2^- + H_2 \longrightarrow NH_3 + H^-$$

This process can continue until, in principle, all the NH_x groups are hydrogenated, and the original BaH_2 catalyst is reconstituted (Fig. 4.12)

$$BaH_{2-2x}(NH)_x + 2\,xH_2 \longrightarrow BaH_2 + xNH_3$$

The catalyst recovery is facilitated by the high mobility of the anions in $BaH_{2-2x}(NH)_x$.
Interestingly, our simulations did not show ammonia formation through the reductive protonation of imides, consistent with experimental evidence indicating no ammonia release during the nitridation process.

4.4 Discussion

A simplified outline of the catalytic process is reported in Fig. 4.13. A crucial feature of this process is the dynamic transformation of the catalyst during the catalytic cycle. This transformation

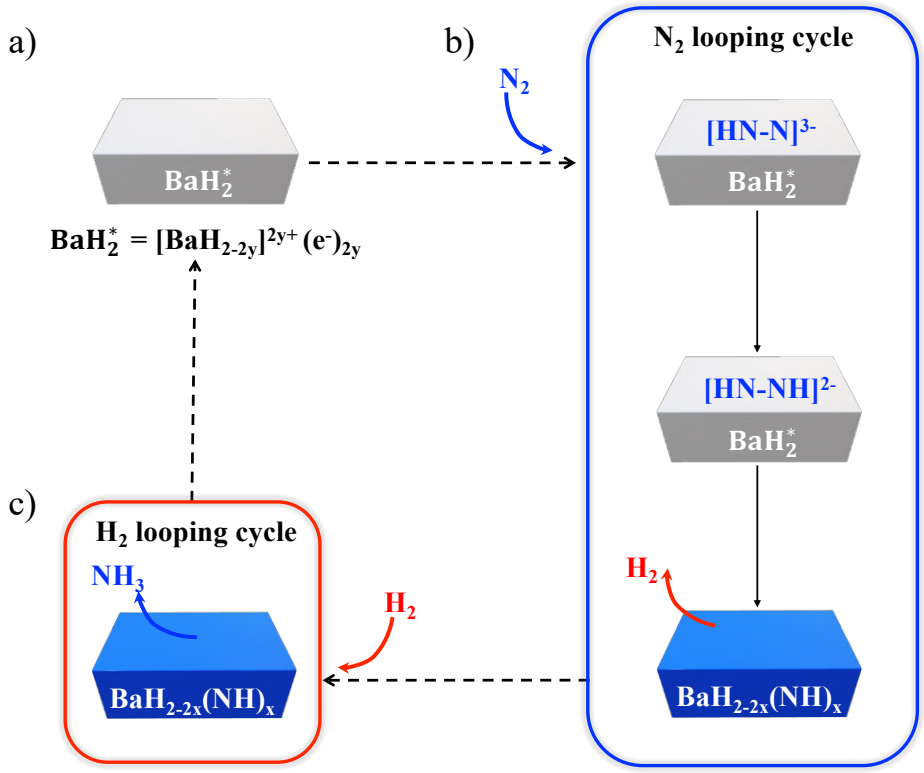

Figure 4.13: **Simplified scheme outlining the catalytic process.** Schematic representation of the ammonia synthesis process through nitridation of BaH_2 and hydrogenation of $BaH_{2-2x}NH_x$ to produce NH_3. a) Thermal activation of the catalyst facilitates the formation of electrides. b) N_2 looping cycle leads to nitrogen reduction and imides formation. In this step BaH_2 is trasformed into the mixed compound $BaH_{2-2x}NH_x$. c) H_2 looping step closes the catalytic process by hydrogenating the mixed compound $BaH_{2-2x}NH_x$, resulting in the formation of NH_3 and regeneration of the catalyst.

affects the whole system, and it is not confined to only the surface. Indeed, as catalyst BaH_2 transcends a traditional role of static support for electron exchange, and actively participates in the catalytic process by providing hydrides to both reactants and intermediates, thus moving the reactive process forward. The highly mobile hydrides can attack the absorbed N_2 molecules aiding in the weakening of the N≡N triple bond, until its final cleavage and eventual imide formation. These steps closely resemble the mechanism engineered by nature in the nitrogenase enzyme, where the N≡N bond is hydrogenated prior to breaking. While the exact mechanism in nitrogenase remains a subject of debate, in this system, hydrides are able to directly attack the adsorbed N_2 molecules. This behavior closely resembles that of certain homogeneous catalysts, where hydride transfer facilitates bond activation and reduction [202].
Furthermore, the system proximity to phase transitions is a key characteristic of this catalyst, which contributes to amplifying the system's overall fluctuations and creates a state of frustration crucial for its catalytic functions. Indeed, the approach to the orthorhombic-hexagonal phase transition, occurring at approximately 770K, results in a substantial increase in hydride mobility

on the pristine surface. These high fluctuations of surface hydrides are essential for the release of H_2 and the formation of the electride, thereby facilitating the catalyst's thermal activation. Additionally, the formation of the hydride-imide mixture, which is an intermediate state in the BaH_2 to BaNH phase transition, results in a superionic state characterized by large mobility of both imides and hydrides. This highly fluctuating environment is crucial for the stabilization of the imides following the N—N bond cleavage, and for the catalyst recovery after the hydrogenation step.
Finally, another key catalytic property of this system is its ability to facilitate electron exchange through anionic vacancies. This mechanism stands in contrast to that of traditional transition metal-based catalysts, which modulate the catalytic process by transferring electrons via their partially occupied *d* orbitals. We have recently reported the important role of F-centres in the ammonia decomposition on lithium imide, and it is significant that the ammonia synthesis on BaH_2 follows the same main steps but in the reverse direction. Furthermore, also recent experimental studies support the critical role of F-centres in ionic catalysts. Indeed, Ping Chen et al. have shed light on the catalytic behavior of lithium hydride (LiH) in ammonia synthesis when exposed to ultraviolet (UV) light [196]. It has been observed that LiH undergoes a photolytic reaction upon UV irradiation, resulting in the release of H_2 and the transfer of electrons generated from photon absorption to surface hydrogen vacancies, effectively creating F-centres, and thus activates LiH for nitrogen fixation. Similar intuitions for the reducing ability of surface defects were proposed for CaH_2, where the vacancies on the surface reduce the work function of the material, making it a plausible reducing agent [203].
Furthermore, the interplay between alkali and alkaline earth metal hydrides with transition metals can lead to an enhanced catalytic activity due to the facilitated electron exchange through the *d* orbitals of the transition metals. This synergistic effect is evidenced by the significant boost in ammonia synthesis observed in Ru/CaH_2, $Ru/Ca_2N{:}e^-$ [204] and Ni-BaH_2 catalysts [205]. Hypothetically, in these systems, the metal hydride serves as a second catalytic site that removes activated nitrogen from the transition metal, hence breaking the traditional scaling relationship between the nitrogen dissociation energy and the formation energy of reduced intermediates.
In conclusion, the ammonia synthesis on BaH_2 is another example that clearly underscores the importance of considering the catalyst dynamics and how it is influenced by the reaction conditions. Traditional approaches in computational catalysis typically model the surface as a static system at 0K, failing to capture this dynamic behavior, and highlighting the limitations of such static calculations in accurately representing the catalytic system under operando conditions. Only by adopting a dynamic microscopic view of catalysis, made possible by advances in machine learning-based molecular dynamics simulations, one can create new opportunities to improve the understanding and design of catalytic systems for industrial applications.

4.5 Methods

All the MD simulations were performed using a machine learning-based interatomic potential combined with enhanced sampling methods to sample reactive events.

Machine learning-based interatomic potential

The MLP was trained using the strategy illustrated in §3.5.5. The configurations of the training set were collected using an active learning approach augmented with the OPES method similar to that of refs. [47, 38] We built a training set composed of 224876 configurations encompassing bulk BaH_2, the pristine (120) surface, intermediates and transition states of all the reaction steps of the nitridation and hydrogenation processes. Furthermore, it also include configurations of the mixed $BaH_{2-2x}(NH)_x$ system across different values of x.

AIMD simulations. AIMD simulations were performed with CP2K (version 9.1) [206, 207] in order to build the initial training set. Simulations were performed in the canonical ensemble (NVT) with periodic boundary conditions, and a time step of 0.5 fs. The temperature was controlled using the stochastic velocity rescaling thermostat [83] with a coupling constant of 0.05 ps. Energy and forces were computed using the PBE exchange-correlation functional [208]. The Kohn–Sham orbitals were expanded in a triple-ξ plus polarization Gaussian basis sets and the plane wave expansion of the electronic density was truncated at an energy cut-off of 240 Ry. The core electrons were treated using the Goedecker–Teter–Hutter pseudopotentials [209, 210] optimized for PBE.

Labeling configurations via DFT calculations. The DFT energies and forces needed for training the neural network were calculated using the Quantum Espresso (QE) software (version 7.2) [211, 212, 213]. The PBE exchange-correlation functional together with the RRKJUS PBE pseudopotentials [214] taken from the QE pseudopotential library were used. The Kohn–Sham orbitals were expanded on a plane wave basis set up to a kinetic-energy cut-off of 60 Ry and a density cut-off at 240 Ry. The convergence threshold for self-consistency was set to 1.0×10^{-7} a.u.

Training of the ML-based interatomic potential. The Deep Potential Molecular Dynamics Smooth Edition scheme, as implemented in the DeePMD-kit software (version 2.1) [144], was used for training the deep potential models. An embedding and a fitting network with [30, 60, 120] and [240, 240, 240, 240] nodes per layer were used, respectively. A cut-off of 7Å was used to describe to local environment of each atom. The ADAM optimizer was used to train the network with an exponentially decaying learning rate from 1×10^{-3} to 5×10^{-8}. The pre-factors used in the in loss function for the energy and force terms are varied during the optimization process from 0.05 to 5 and from 1000 to 1, respectively. The batch size was set to 8 and the final models used for the production runs were trained for 2×10^6 steps.

Accuracy of the ML-based interatomic potential. The mean absolute errors (MAEs) of energies in the training and test set are 2.11 and 2.74 meV per atom, respectively. The MAEs of forces in the training and test set are 64.8 and 54.8 meV/Å, respectively (Fig. 4.14). Relevant frames along the reactive pathways and the transition states were included in the test set. We also carefully verified the accuracy of the forces predicted by the ML potential on the atoms involved in the reactions, confirming that the errors were minimal and consistent with the overall accuracy of the ML potential.

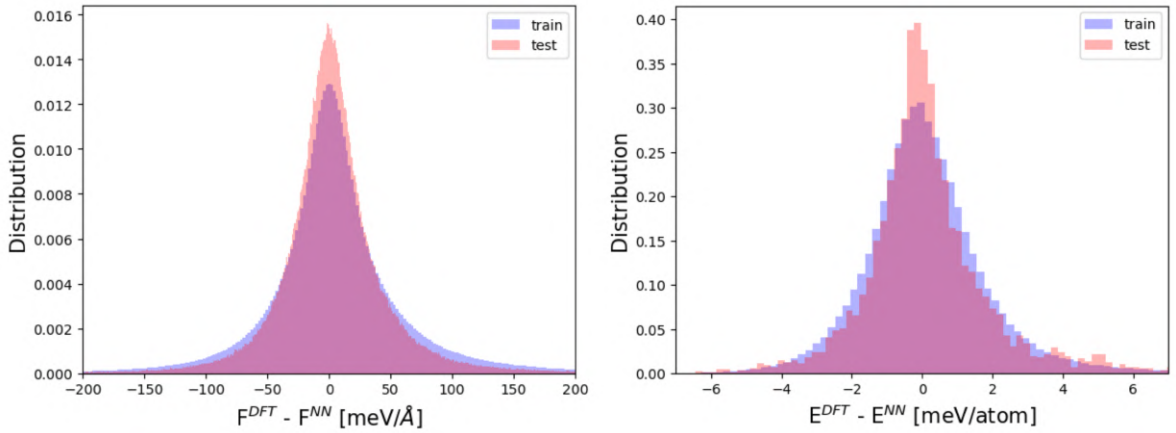

Figure 4.14: **Accuracy of the machine learning-based interatomic potential**. Histogram of the errors in the forces (top) and energies (bottom) in the train set and test set.

Molecular Dynamics simulations

The bulk dynamics of BaH_2 was studied in an orthorhombic supercell of edges 27.278, 25.376, 31.508 Å, containing 1152 atoms, which amounts to replicate 4×6×4 times the conventional unit cell. The ammonia synthesis on the (120) surface was modelled using a slab made of 756 atoms with edges 32.398, 23.648, 77.325 Å. In this setup, we left enough distance (about 20Å) between replicas along the z direction to minimize interactions. In all the simulations of the (120) surface, we fixed the atoms in the 2 bottom layers to their ideal crystal lattice positions to mimic the bulk environment. For reactive events where a N_2 or H_2 gaseous molecule is involved, we constrained the z-position of the molecule using a reflective wall [215]. In this way the molecule can move in a volume of $32.398 \times 23.648 \times 10$ Å^3.
Deep Potential molecular dynamics simulations were performed with Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) software (version 24Mar2022) [216], patched with DeePMD-kit and PLUMED (version 2.9) [127]. NVT simulations were performed with an integration time step of 0.5 fs. The temperature was controlled using stochastic velocity scaling thermostat [83] with a coupling constant of 50 fs.
MD simulations within the NPT ensemble were conducted using the Parrinello-Rahman formalism to ensure the stability of the $BaH_{2-2x}(NH)_x$ system. We used a pressure damping parameter of 250 fs and a scalar external pressure of 1 bar. All the simulations were performed on the deuterated system.

Enhanced Sampling Simulations

The PLUMED library [127] is used to calculate the collective variables and the OPES [115] bias potentials for both ab initio and deep potential molecular dynamics simulations. The formation of H_2 from the pristine surface was obtained by biasing the minimum of the distance between H

atoms $d = d(\mathrm{H}, \mathrm{H})$:

$$\min(d) = \frac{\beta}{\log \sum_i \exp\left(\frac{\beta}{d_i}\right)} \tag{4.1}$$

The biased simulations regarding the formation of the various hydrogenated nitrogen species (i.e. NNH^{3-}, $HNNH^{2-}$, NH^{2-}, NH_2^-, NH_3) were performed by biasing the maximum of the coordination number (C.N. = C.N.(N,H)) between N and H atoms:

$$\mathrm{C.N.} = \sum_{i \in \mathrm{H}} \frac{1 - \left(\frac{d_{\mathrm{N,H}}}{d_0}\right)^6}{1 - \left(\frac{d_{\mathrm{N,H}}}{d_0}\right)^{12}} \tag{4.2}$$

$$\max(\mathrm{C.N.}) = \beta \log \sum_i \exp\left(\frac{\mathrm{C.N.}_i}{\beta}\right) \tag{4.3}$$

where $d_{\mathrm{N,H}}$ is the distance between N and H atoms. Details regarding the bias settings are reported in Table 4.2.

Neural Network Model for Predicting Bader Charges

The training of the neural network for Bader charges inference is performed using a DeepDOS model [217] with an attention mechanism [218]. Bader charges were computed processing the DFT charge density obtained from QE with the code by Henkelman et al. [219]. The embedding network is composed of three layers with 25, 50 and 100 nodes, while the fitting network has three hidden layers with 240 nodes in each layer. The radius cutoff is chosen to decay from 0.5Å to 6.0Å. The number of layers in the attention scheme is set to 2, whereas the size of the embedding matrix to 12. Since we want to train only the scalar atomic properties, i.e. Bader charges per atom, we set the weight of the global loss to 0, whereas the weight of the atomic loss is set to 1. The training set and test set are composed by 23746 and 290 configurations, respectively. The atomic RMSE on the test set is 1.71e-2 e and the atomic MAE is 6.79e-3 e Fig. 4.15 and Fig. 4.16.

Temp [K]	Formula	Number of replicas	Total simulation time	Bias
Dynamics of stoichiometric BaH_2				
550	$Ba_{384}H_{768}$	1	28ns	No
600	$Ba_{384}H_{768}$	1	21ns	No
650	$Ba_{384}H_{768}$	1	21ns	No
700	$Ba_{384}H_{768}$	1	21ns	No
750	$Ba_{384}H_{768}$	1	21ns	No
H^- diffusion in bulk BaH_{2-x}				
500	$Ba_{384}H_{767}$	1	28ns	No
550	$Ba_{384}H_{767}$	1	28ns	No

600	$Ba_{384}H_{767}$	1	28ns	No
650	$Ba_{384}H_{767}$	1	28ns	No
700	$Ba_{384}H_{767}$	1	28ns	No
750	$Ba_{384}H_{767}$	1	28ns	No
500	$Ba_{384}H_{766}$	1	30ns	No
550	$Ba_{384}H_{766}$	1	41ns	No
600	$Ba_{384}H_{766}$	1	38ns	No
650	$Ba_{384}H_{766}$	1	41ns	No
700	$Ba_{384}H_{766}$	1	30ns	No
750	$Ba_{384}H_{766}$	1	41ns	No
500	$Ba_{384}H_{764}$	1	28ns	No
550	$Ba_{384}H_{764}$	1	28ns	No
600	$Ba_{384}H_{764}$	1	28ns	No
650	$Ba_{384}H_{764}$	1	28ns	No
700	$Ba_{384}H_{764}$	1	28ns	No
750	$Ba_{384}H_{764}$	1	28ns	No
Pristine (120) surface dynamics				
550	$Ba_{252}H_{504}$	4	300ns	No
700	$Ba_{252}H_{504}$	4	300ns	No
H_2 Formation from (120) surface				
500	$Ba_{252}H_{504}$	15	17ns	Yes
550	$Ba_{252}H_{504}$	15	9ns	Yes
600	$Ba_{252}H_{504}$	20	6ns	Yes
650	$Ba_{252}H_{504}$	20	6ns	Yes
700	$Ba_{252}H_{504}$	20	3ns	Yes
750	$Ba_{252}H_{504}$	20	4ns	Yes
N_2 adsorption on (120) surface				
550	$Ba_{252}H_{504}N_2$	1	5ns	No
550	$Ba_{252}H_{502}N_2$	1	5ns	No
NNH^{3-} Formation on (120) surface (FES)				
550	$Ba_{252}H_{502}N_2$	1	19ns	Yes
700	$Ba_{252}H_{502}N_2$	1	19ns	Yes
NNH^{3-} Formation on (120) surface				
550	$Ba_{252}H_{502}N_2$	21	8ns	Yes
600	$Ba_{252}H_{502}N_2$	15	3ns	Yes
650	$Ba_{252}H_{502}N_2$	15	2ns	Yes
700	$Ba_{252}H_{502}N_2$	19	5ns	Yes
$HNNH^{2-}$ Formation on (120) surface				

550	$Ba_{252}H_{502}N_2$	20	2ns	Yes
600	$Ba_{252}H_{502}N_2$	20	2ns	Yes
650	$Ba_{252}H_{502}N_2$	16	7ns	Yes
700	$Ba_{252}H_{502}N_2$	20	3ns	Yes
	NH^{2-} Formation on (120) surface			
550	$Ba_{252}H_{502}N_2$	19	15ns	Yes
600	$Ba_{252}H_{502}N_2$	20	23ns	Yes
650	$Ba_{252}H_{502}N_2$	20	11ns	Yes
700	$Ba_{252}H_{502}N_2$	19	10ns	Yes
	H^- and NH^{2-} Diffusion in bulk $BaH_{2-x}(NH)_x$			
700	$Ba_{384}H_{720}N_{48}$	4	98ns	No
700	$Ba_{384}H_{672}N_{96}$	4	60ns	No
700	$Ba_{384}H_{576}N_{192}$	4	60ns	No
550	$Ba_{384}H_{720}N_{48}$	4	60ns	No
550	$Ba_{384}H_{672}N_{96}$	4	60ns	No
550	$Ba_{384}H_{576}N_{192}$	4	60ns	No
	NH_2^- Formation on (120) surface			
550	$Ba_{252}H_{504}N_2$	5	3ns	Yes
700	$Ba_{252}H_{504}N_2$	5	3ns	Yes
	NH_3 Formation on (120) surface			
550	$Ba_{252}H_{506}N_2$	5	3ns	Yes
700	$Ba_{252}H_{506}N_2$	5	3ns	Yes

Table 4.1: **Performed simulations with our machine learning-based interatomic potential**. Details about enhanced sampling setup are reported in Table 4.2

Temp [K]	CV	PACE	BARRIER	EXCLUDED REGION
	H_2 Formation on (120) surface			
550	max(CN(H,H))	250	100kJ/mol	H(x-0.8)
600	max(CN(H,H))	250	100kJ/mol	H(x-0.8)
650	max(CN(H,H))	250	80kJ/mol	H(x-0.8)
700	max(CN(H,H))	250	70kJ/mol	H(x-0.8)
750	max(CN(H,H))	250	60kJ/mol	H(x-0.8)
	NNH^{3-} Formation (FES)			
550	min(d(N,H))	500	90kJ/mol	No

700	min(d(N,H))	500	70kJ/mol	No
		NNH^{3-} Formation (flooding)		
550	min(d(N,H))	200	55kJ/mol	H(1.8-x)
600	min(d(N,H))	200	55kJ/mol	H(1.8-x)
650	min(d(N,H))	200	55kJ/mol	H(1.8-x)
700	min(d(N,H))	200	55kJ/mol	H(1.8-x)
		$HNNH^{2-}$ Formation		
550	min(d(N,H))	200	80kJ/mol	H(2.0-x)
600	min(d(N,H))	200	80kJ/mol	H(2.0-x)
650	min(d(N,H))	200	80kJ/mol	H(2.0-x)
700	min(d(N,H))	200	80kJ/mol	H(2.0-x)
		2 NH^{2-} Formation		
550	d(N,N)	200	30kJ/mol	H(x-1.8)
600	d(N,N)	200	30kJ/mol	H(x-1.8)
650	d(N,N)	200	30kJ/mol	H(x-1.8)
700	d(N,N)	200	30kJ/mol	H(x-1.8)
		NH_2^- Formation		
550	d(N,H)	200	40kJ/mol	H(2.0-x)
700	d(N,H)	200	40kJ/mol	H(2.0-x)
		NH_3 Formation		
550	d(N,H)	200	40kJ/mol	H(2.0-x)
700	d(N,H)	200	40kJ/mol	H(2.0-x)

Table 4.2: Enhanced sampling details for biased simulations. The Heaviside step function, denoted by H, is used to delimit the bias deposition region.

4.5.1 Training and test set

The total number of configurations in the training set is 224876. The total number of configurations in the test set is 12720. Energy RMSE/Natoms and Forse RMSE are given in [eV/atom] and [eV/Å], respectively. Box sizes: * = [a = 13.56482 Å, b = 12.48153 Å, c = 15.68642 Å]; ● = [a = 10.799314 Å, b = 15.76516 Å, c = 49.381489 Å]; ◇ = [a = 21.59862 Å, b = 15.76516 Å, c = 52.02922 Å]; ▹ = [a = 20.34723 Å, b = 12.48153 Å, c = 15.68642 Å].

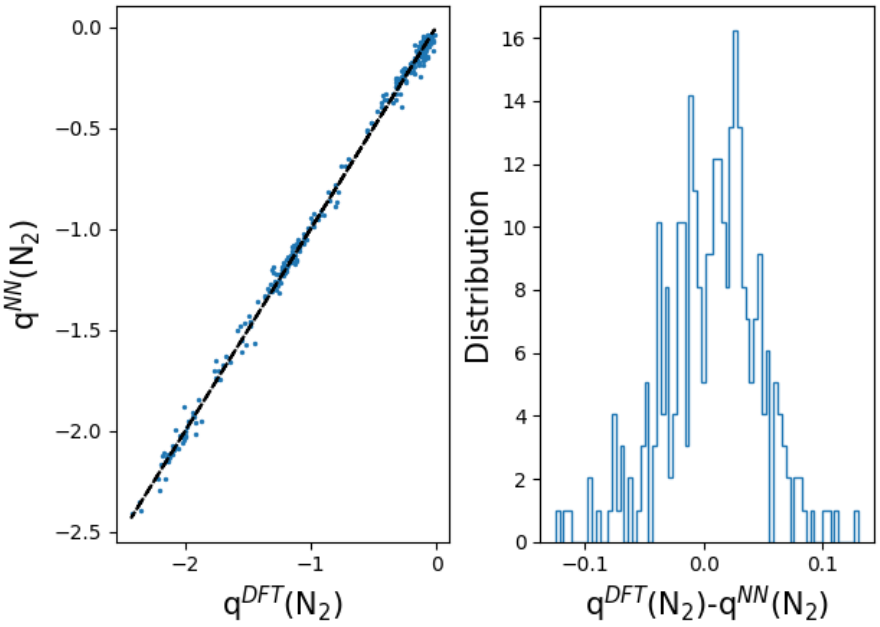

Figure 4.15: **Accuracy of the Neural Network trained for predicting Bader charges of N atoms**. Left panel: Comparison of the sum of the Bader charges of N atoms in the test set, as calculated at DFT level and using the final NN model. Right Panel: Error distribution for the sum of the Bader charges of N atoms.

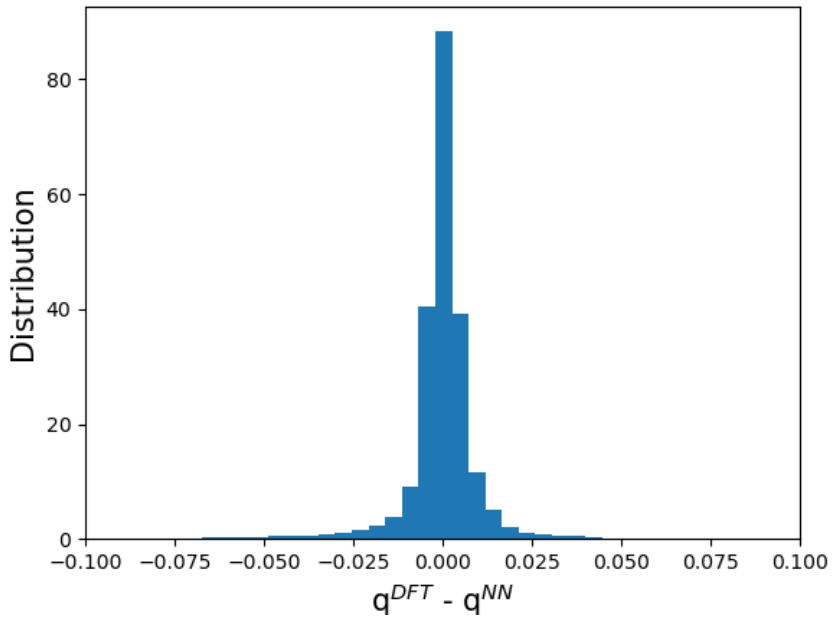

Figure 4.16: Accuracy of the Neural Network trained for predicting Bader charges for all atoms. Error distribution for Bader charges of Ba, H and N atoms.

	Training Set			Test Set		
System	Num. configs	Energy RMSE/Natoms	Force RMSE	Num. configs	Energy RMSE/Natoms	Force RMSE
			Bulk BaH_2 *			
$Ba_{48}H_{96}$	3250	5.25e-4	2.71e-2	300	4.08e-4	2.69e-3
			(120) surface ●			
$Ba_{32}H_{64}$	3499	9.30e-4	3.39e-2	300	4.53e-4	2.84e-2
$Ba_{32}H_{64}$ + 2 H_2	958	1.85e-3	3.96e-2	50	1.73e-3	3.99e-2
$Ba_{32}H_{64}$ + 2 N_2	652	1.43e-3	9.57e-2	50	1.50e-3	9.45e-2
$Ba_{32}H_{64}$ +N_2 +H_2	604	1.27e-3	7.39e-2	501	5.99e-3	1.42e-1
$Ba_{32}H_{62}$ +N_2 +H_2	3058	1.78e-3	5.75e-2	150	9.09e-3	1.60e-1
$Ba_{32}H_{64}$ + N_2	572	8.62e-4	8.34e-2	500	7.06e-4	6.10e-2
			H_2 Formation from (120) surface ●			
$Ba_{32}H_{64}$	19067	1.631245e-03	4.297567e-02	500	5.45e-4	2.54e-2
			N_2 adsorption and NNH^{3-} formation ●			
$Ba_{48}H_{94}N_2$	18665	1.12e-3	4.61e-2	642	1.09e-3	4.99e-2
$Ba_{48}H_{93}NNH$	21139	1.51e-3	5.25e-2	500	1.75e-3	4.39e-2
			$HNNH^{2-}$ formation ●			
$Ba_{32}H_{64}N_2$	1163	2.65e-3	5.04e-2	50	8.55e-4	4.10e-2
$Ba_{32}H_{60}HNNH$	4668	2.39e-3	6.13e-2	250	1.20e-3	4.55e-2
$Ba_{32}H_{60}HNNH$	23500	2.07e-3	5.58e-2	250	1.18e-3	5.02e-2
$Ba_{32}H_{64}HNNH$	14169	1.32e-3	4.88e-02	301	9.08e-4	3.39e-2
			Surface with imides and/or amides			
$Ba_{32}H_{64}(NH_2)$ ●	4300	1.022906e-03	3.917139e-02	150	1.88e-3	4.61e-2
$Ba_{32}H_{62}(NH_2)_2$ ●	9331	2.43e-3	4.62e-2	300	1.64e-3	3.86e-2
$Ba_{32}H_{62}(NH)(NH_2)$ ●	4058	2.28e-3	5.20e-2	300	2.58e-3	5.84e-2
$Ba_{32}H_{60}(NH)_2$ ●	3243	2.86e-3	6.11e-2	100	2.82e-3	5.56e-2
$Ba_{32}H_{61}(NH)(NH_2)$ ●	1964	3.01e-3	5.63e-2	300	4.06e-3	5.57e-2
$Ba_{32}H_{54}(NH)_4$ ●	5303	1.71e-3	6.25e-2	216	1.86e-3	6.26e-2
$Ba_{32}H_{56}(NH)_4$ ●	4709	1.88e-3	5.88e-2	500	5.45e-4	2.54e-2

$Ba_{32}H_{54}(NH)_4 + N_2$ ●	3859	1.71e-3	6.99e-2	200	1.80e-3	6.96e-2
$Ba_{32}H_{58}(NH)_2(NH_2)_2$ ●	1521	2.04e-3	6.57e-2	50	3.03e-3	5.67e-2
$Ba_{32}H_{57}(NH)_3(NH_2)$ ●	4491	1.81e-3	6.21e-2	50	1.51e-3	5.36e-2
$Ba_{80}H_{144}(NH)_8$ ◇	1930	9.08e-4	5.24e-2	100	1.10e-3	4.73e-2
$Ba_{80}H_{142}(NH)_8$ ◇	2419	1.09e-3	5.43e-2	100	1.06e-3	5.49e-2
$Ba_{80}H_{142}(NH)_8 + N_2$ ◇	3500	1.00e-3	5.84e-2	500	9.98e-4	5.76e-2
$Ba_{80}H_{140}(NH)_{10}$ ◇	3000	1.10e-3	6.05e-2	50	1.27e-3	6.15e-2
$Ba_{80}H_{136}(NH)_{12}$ ◇	4500	1.07-3	6.35e-2	500	1.06e-3	6.43e-2
$Ba_{80}H_{148}(NH)_4(NH_2)_4$ ◇	4000	8.25e-4	5.14e-2	500	7.79e-4	5.17e-2
			Bulk $BaH_{2-x}(NH)_x$			
$Ba_{48}H_{78}(NH)_6$ *	8500	1.21e-3	5.45e-2	500	1.22e-3	5.47e-2
$Ba_{48}H_{72}(NH)_{12}$ *	11000	1.58e-3	6.98e-2	1000	1.55e-3	6.97e-2
$Ba_{48}H_{60}(NH)_{24}$ *	13000	2.32e-3	9.30e-2	2000	2.23e-3	9.22e-2
$Ba_{96}H_{96}(NH)_{48}$ ▷	5500	1.48e-3	9.30e-2	459	1.51e-3	9.31e-2
			NH_3 formation ●			
$Ba_{32}H_{62}(NH_2)_2 + H_2$	859	1.43e-3	4.58e-2	50	1.39e-3	4.42e-2
$Ba_{32}H_{63}(NH_2) + NH_3$	2827	1.49e-3	4.73e-2	100	1.42e-3	4.79e-2
$Ba_{32}H_{64} + NH_3$	1372	1.47e-3	4.69e-2	51	1.64e-3	3.86e-2
$Ba_{32}H_{64} + 2\,NH_3$	1292	1.57e-3	5.50e-2	50	1.43e-3	5.22e-2
$Ba_{32}H_{59}(NH_2) + NH_3$	1598	1.05e-3	3.85e-2	100	4.06e-3	5.57e-2
$Ba_{32}H_{59}(NH_2) + NH_3 + H_2$	563	1.44e-3	4.25e-2	50	1.52e-3	4.18e-2
$Ba_{32}H_{64} + NH_3$	592	8.15e-4	3.40e-2	50	1.75e-3	4.39e-2
$Ba_{32}H_{62} + NH_3$	681	1.90e-3	4.84e-2	50	1.32e-2	7.51e-2

Table 4.3: Composition of the training and test sets.

Chapter 5

Electroreduction of CO_2 on electrified Ag/H_2O interface

In the previous chapter, we examined how *operando* temperature strongly influences the dynamic behavior of the catalyst, ultimately driving a reversible phase transformation under reactant exposure. This example represents an extreme case that clearly demonstrates how *operando* conditions can profoundly alter the catalyst structure and behavior. It highlights the limitations of static calculations and underscores the importance of accurately accounting for realistic reaction conditions in order to obtain an in-depth understanding of catalysis. In chapter 2, we also introduced the importance of explicitly modeling the solvent and applied potential in electrochemical simulations. In this chapter, we build upon these concepts by discussing the results presented in ref. [29], and highlight how the applied potential acts as a key *operando* variable that affects both the solvent dynamics and the reaction thermodynamics.

5.1 Introduction

Among the various factors, it is assumed that the electric potential plays a pivotal role in CO_2 reduction [220, 52]. As illustrated in Fig. 5.1, in a typical CO_2 reduction system, an external constant potential is applied to the cathode, causing negative charge accumulation on the surface of the catalyst [221]. The electric field formed at the catalyst-electrolyte interface alters the spatial distribution of ions and solvent molecules near the surface and directly influences the reaction mechanism [222, 223]. As reviewed in chapter § 3.3, various computational strategies have been proposed to model electrochemical systems, ranging from implicit solvent approximations [224, 225, 226] to grand-canonical DFT (GCE-DFT) methods [227, 228]. More recently, hybrid explicit-implicit solvent models have been developed within the framework of GCE-DFT to better capture interfacial solvent structure while maintaining constant-potential conditions [229, 60]. However, updating of the electron count at each self-consistent field step makes these simulations even more computationally demanding than standard AIMD [106]. As a result, the high cost of explicit ab initio methods continues to limit simulation timescale and system sizes, hindering an accurate modeling of the coupled influence of the electric potential and solvent dynamics on CO_2 electrore-

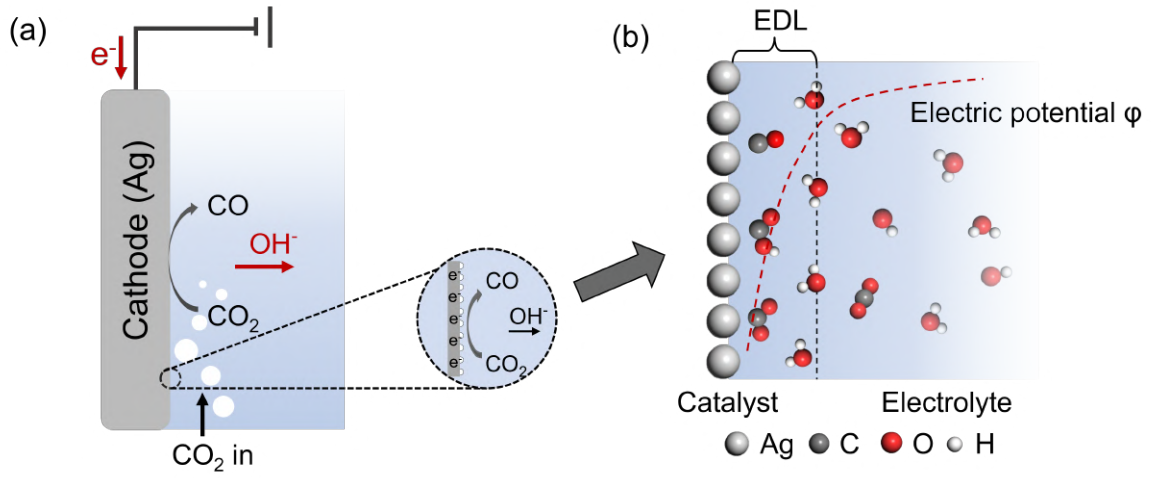

Figure 5.1: **Representation of the electrochemical interface for CO_2 reduction.** (a) Scheme of the cathode-side of a H-type cell for CO_2 reduction. (b) Scheme of the microscopic electrified catalyst-electrolyte interface for CO_2 reduction.

duction [221, 52]. Also in this case, MLP can be applied to address this limitations. In this work, we combined machine learning accelerated molecular dynamics simulations, explicit solvent model within the GCE-DFT framework to perform constant-potential simulations, and enhanced sampling methods to investigate the CO_2 electroreduction mechanism on the Ag(111) surface. We chose silver as the model catalyst due to its well-established selectivity and industrial prospect for the CO_2-to-CO conversion [230]. Silver exhibits a high CO Faraday efficiency that exceeds 90% and supports current densities above 200 mA/cm^2, making it a benchmark system for both fundamental studies and practical applications [231, 232]. We considered two electrochemical potentials: the potential of zero charge (PZC) and a constant potential (CP) corresponding to experimentally relevant working conditions. By constructing two MLPs, each trained to reproduce DFT-level energies and forces under PZC and CP conditions, we unraveled the full complexity of the CO_2 reduction process under operando electrochemical conditions. Our simulations reveal that the applied potential affects reaction barriers not simply through electronic effects, but through the dynamic reorganization of the solvent environment, which governs the accessibility and stability of reactive intermediates and transition states. This coupling especially influences the hydrogenation step and the corresponding ion-pair dissociation process, whose barriers and kinetics are particularly influenced by the solvent response to the electric bias.

5.2 Results

As anticipated, we compare the results under two distinct electrochemical potential conditions: the potential of zero charge (PZC) and a constant potential (CP) of –1.5 V vs. SHE. The latter closely approximates the experimental onset potential for CO_2 reduction on Ag [233, 234], while PZC represents a key reference state where no excess charge is present on the surface [235]. This comparison

provides insight into how the solvent environment and reaction kinetics respond to the applied potentials. To capture the differences between these conditions, we trained two separate sets of MLPs to model the Ag(111)/H_2O interface and the CO_2-to-CO reaction pathway under PZC and CP, respectively. A detailed description of the MLP training workflow is provided in the Methods section (Fig. 5.12).

5.2.1 Ag(111)/H_2O interface response to the applied potential

We first explore the response of the interfacial water structure and dynamics to the applied working potential. As shown in Fig. 5.2a and b, the applied potential has a significant impact on the structure of the interfacial water. At PZC condition, the O density profile (Fig. 5.2a, upper) exhibits two distinct peaks at 2.7 Å and 3.2 Å from the Ag surface, corresponding to two close adsorption layers of water oxygen atoms. The distribution of water dipole orientation reveals a peak at $\sim 55^\circ$ in Fig. 5.2b, indicating that the interfacial water molecules slightly tilt toward the electrolyte, with one O−H bond nearly parallel to the Ag surface at PZC condition. However, when a negative potential is applied to the interface, the peak in the O density profile at 2.7 Å disappears, while two peaks could be observed in the H density profile close to the Ag surface (Fig. 5.2a, down). The dipole orientation of water has a peak at $\sim 135^\circ$. These results suggest that the interfacial water molecules rotate to a "H-down" structure due to the electrostatic attraction between the negatively charged Ag substrate and the positively charged H atoms of water molecules (see Fig. 5.2c). Additional features appear in the density profile at approximately 6 Å and 9 Å. These peaks indicate the presence of partially ordered water layers. Beyond this region, the density gradually approaches 1 g/cm^3 under both potential conditions, consistent with the bulk-like water behavior. The two dimensional distribution of water dipole along the z-direction (Fig. 5.3) confirms that the influence of electric potential on water is confined to the interface and gradually diminishes as one moves away from it. In addition to the static structural properties of water near the interface, it is essential to investigate its dynamic behavior under the two electrochemical conditions. Following the water density profile along z, we identify three different regions: interfacial (0–5 Å), intermediate (5–10 Å), and bulk (10–15 Å). The averaged number of hydrogen bond donors of water molecules (N_{donor}) in these regions are shown as a violin plot in Fig. 5.2d. The mean value of N_{donor} within the interfacial region decreases from 1.62 to 1.38 as the potential shifts from PZC to CP conditions. This is consistent with the "H-down" structure of interfacial water at CP condition, where some H atoms orient toward and interact directly with the Ag surface, rather than participating in hydrogen bonding with neighboring water molecules. The distribution of the N_{donor} is nearly identical across the intermediate and bulk water regions. Moreover, we observed a slight enhancement in the mobility of interfacial water under CP condition, as evidenced by a broader distribution of displacement vectors (Fig. 5.4a). Further analysis indicates that the hydrogen bond lifetime of interfacial water at CP condition is 21.2% shorter than at PZC (Fig. 5.4b). Additionally, both the libration and H−O−H bending modes exhibit a noticeable red shift, indicating a weaker hydrogen bond environment near the electrified Ag surface [236, 237, 238] (Fig. 5.4c,d,e). The distinct properties of interfacial water at negative potential are consistent with previous studies on Pt/H_2O and Au/H_2O interfaces, investigated via in situ Raman spectroscopy and AIMD simulations [236, 239, 240]. These findings validate the reliability of our MLPs in modeling the behavior of

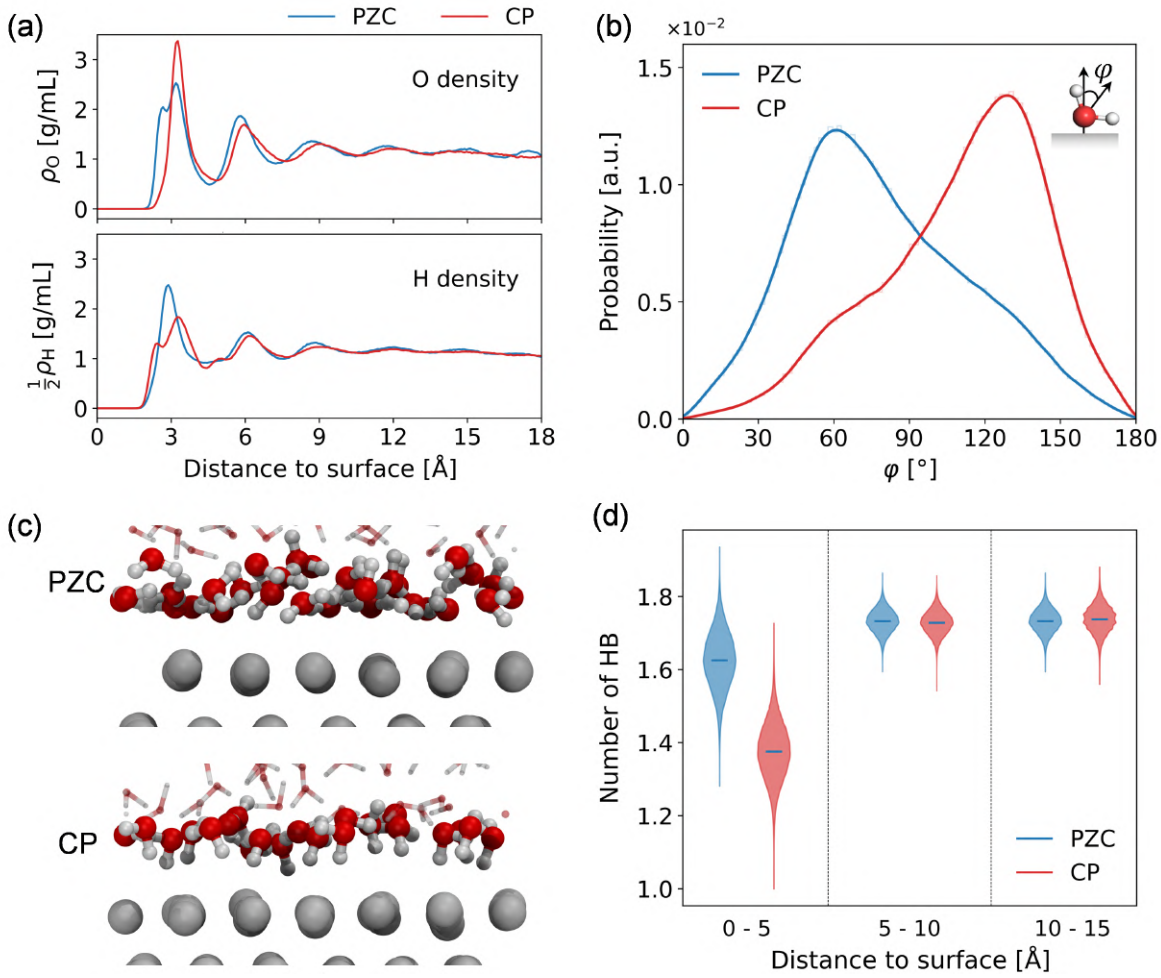

Figure 5.2: **Characterizations of Ag(111)/H_2O interface at PZC and CP conditions.** (a) Oxygen (upper) and hydrogen (lower) density profiles of water along the surface normal (z-direction), where the zero corresponds to the location of the Ag surface. (b) Probability distribution of the angle φ between the bisector of the water H-O-H angle and the z-direction computed for the interfacial water molecules (φ is shown in the inset). The interfacial water is taken as being within 5 Å from the Ag surface, as suggested in the density profiles in (a). (c) Representative snapshots of Ag(111)/H_2O interface at PZC and CP conditions. The Ag, O, H atoms are colored in light gray, red and white, respectively, the water molecules outside the interfacial region are displayed in transparent stick style for clear illustration. (d) Violin plot showing the distribution of averaged hydrogen bond donor number of water in the interfacial region (0 - 5 Å), intermediate region (5 - 10 Å), and bulk region (10 - 15 Å). The blue horizontal lines represent the mean value of N_{donor}. All results were obtained from unbiased trajectories of a 6×6 Ag(111) slab with 180 water molecules, for a total simulation time of 2.5 ns.

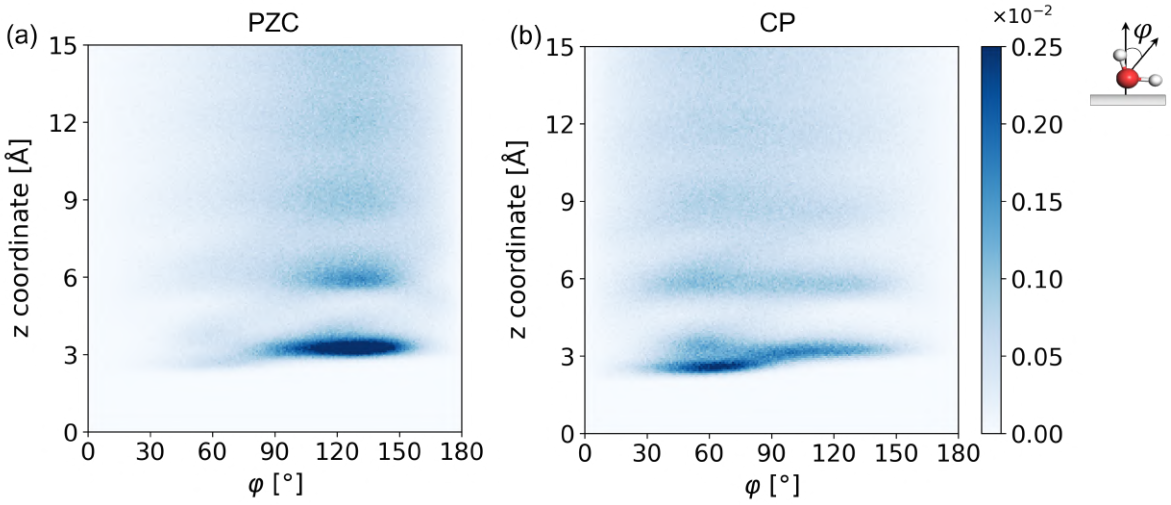

Figure 5.3: **Distribution of water**. Two-dimensional distribution of the water orientation under (a) PZC and (b) CP conditions. The histogram is shown as a function of the angle φ and the z-coordinate of water from the Ag surface. The definition of φ is illustrated in the inset.

the Ag(111)/H_2O interface under varying potential conditions and provide a solid foundation for further exploration of CO_2 reduction reactions at the Ag(111)/H_2O interface.

5.2.2 CO_2 adsorption process

The CO_2 reduction to CO is widely recognized to proceed through four elementary steps [241, 242]: (1) CO_2 adsorption, (2) $^*CO_2$ hydrogenation leading to the formation of $^*COOH$, (3) splitting of this intermediate, followed by (4) the final *CO desorption.

$$(1) \quad CO_2 + e^- + {}^* \longrightarrow {}^*CO_2{}^-$$
$$(2) \quad {}^*CO_2{}^- + H_2O \longrightarrow {}^*COOH^- + OH^-$$
$$(3) \quad {}^*COOH^- + e^- \longrightarrow {}^*CO + OH^-$$
$$(4) \quad {}^*CO \longrightarrow CO + {}^*$$

Here, we adopt a step-by-step strategy to systematically investigate how the potential-driven changes in interfacial water structure and dynamics regulate the above steps of CO_2 reduction.
We first analyzed the CO_2 absorption process, and converged the free energy surface as a function of the z-distance (d) of CO_2 from the surface and the O=C=O bond angle (Fig. 5.5a), using the on-the-fly probability enhanced sampling (OPES) method [115]. Two metastable states were identified both in the absence and in the presence of the applied potential: a solvated CO_2 state in which the solute is to be found at the interfacial region (CO_2(aq)) and one in which it is chemisorbed ($^*CO_2$). The shape of the FES is very similar in the two scenarios. The CO_2 molecule, which is linear in the solvated phase, bends to an angle of approximately 115 to 130° upon adsorption. The equilibrium distance between the $^*CO_2$ center of mass and the surface is approximately 2.1 Å.
The FESs projected along the distance d from the surface are shown in Fig. 5.5b. The adsorption

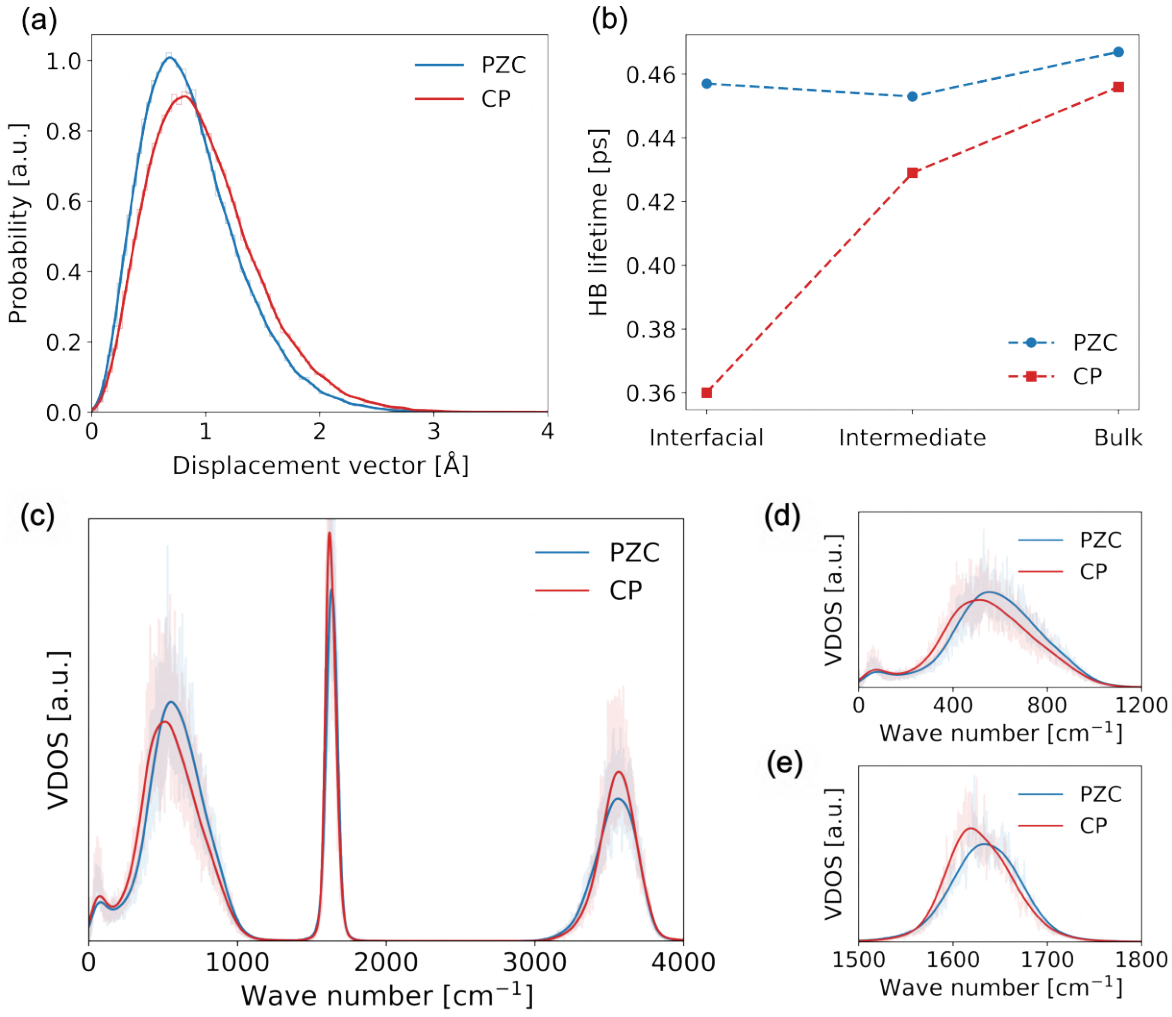

Figure 5.4: **Effects of potential on water proprerties**. (a) Probability distribution of the displacement vectors d of the O atoms of interfacial water molecules under PZC and CP conditions. The displacement vectors are calculated using a Δt of 0.5 ps and the modulus of a displacement vector for O atoms at time t is defined as: $d = \sqrt{(x(t) - x(t-\Delta t))^2 + (y(t) - y(t-\Delta t))^2 + (z(t) - z(t-\Delta t))^2}$. The right shift of the peak and the tail of the probability distribution indicates larger displacement range of interfacial water under CP condition. (b) Lifetime of hydrogen bond in different regions of the Ag(111)/H_2O interface under PZC and CP conditions. (c) The computational VDOS of interfacial water molecules within 5 Å of the electrode under PZC and CP conditions. VDOS was computed following ref. [**?**]. Expanded views of (d) libration modes (400–800 cm^{-1}) and (e) H-O-H bending modes (1600–1700^{-1}) are highlighted to show spectral differences between two potential conditions.

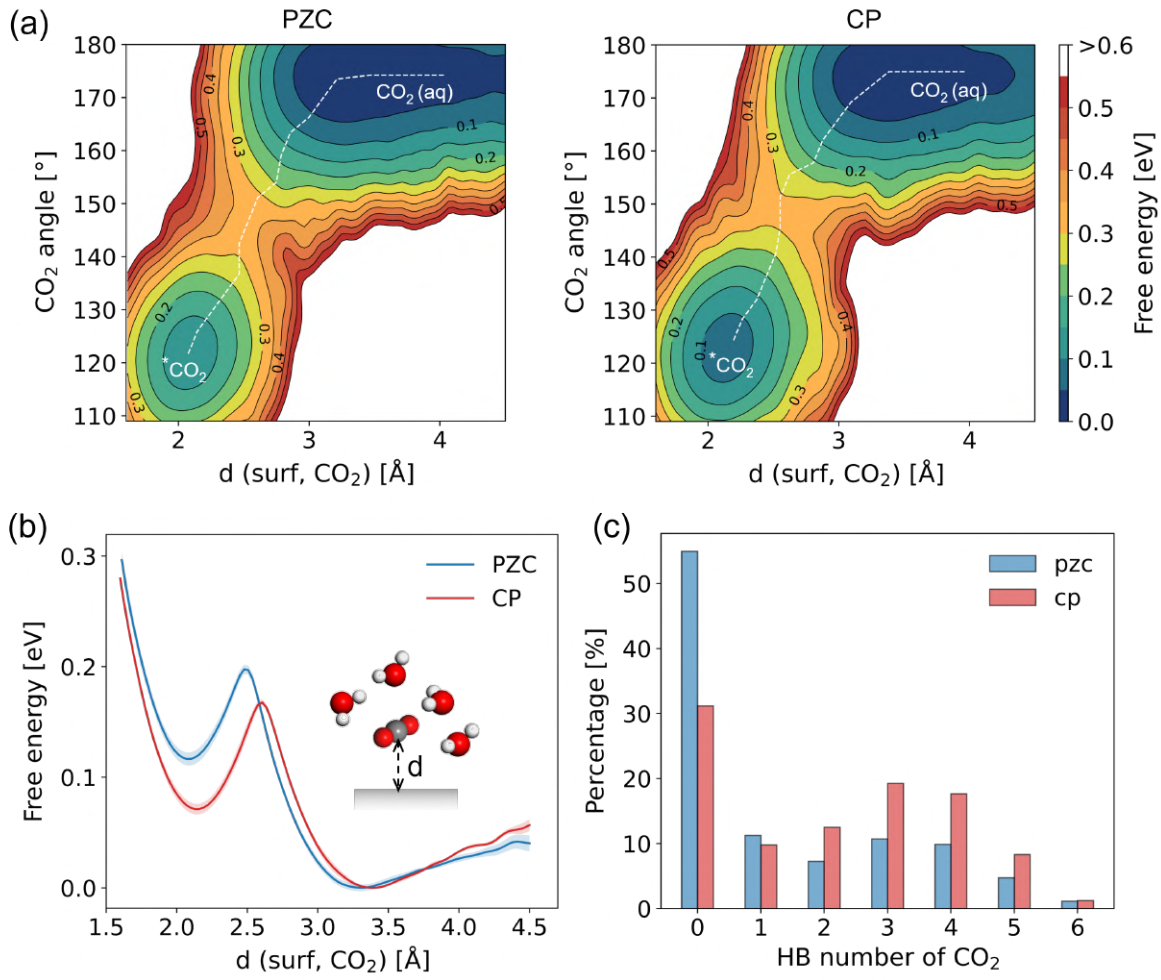

Figure 5.5: **Potential-dependent energetics and solvation structure in CO_2 adsorption process.** (a) Two-dimensional FES as a function of the vertical distance of CO_2 from the surface (d) and O=C=O angle of CO_2 at PZC (left) and CP (right) conditions. The minimum free energy pathway from CO_2(aq) to $^*CO_2$ is depicted with white dashed line and it is computed using MEPSA (Minimum Energy Path Surface Analysis) package[**?**]. (b) One-dimensional projection of the FES along the vertical distance of CO_2 from the surface, the statistics errors are calculated using the weighted-block average techniques as discussed in ref.[**?**] and displayed with shaded colors in the plot. (c) Histogram showing the distribution of the number of hydrogen bonds formed with CO_2 during the intermediate stage of the CO_2 adsorption process (2.25 Å < d < 2.75 Å).

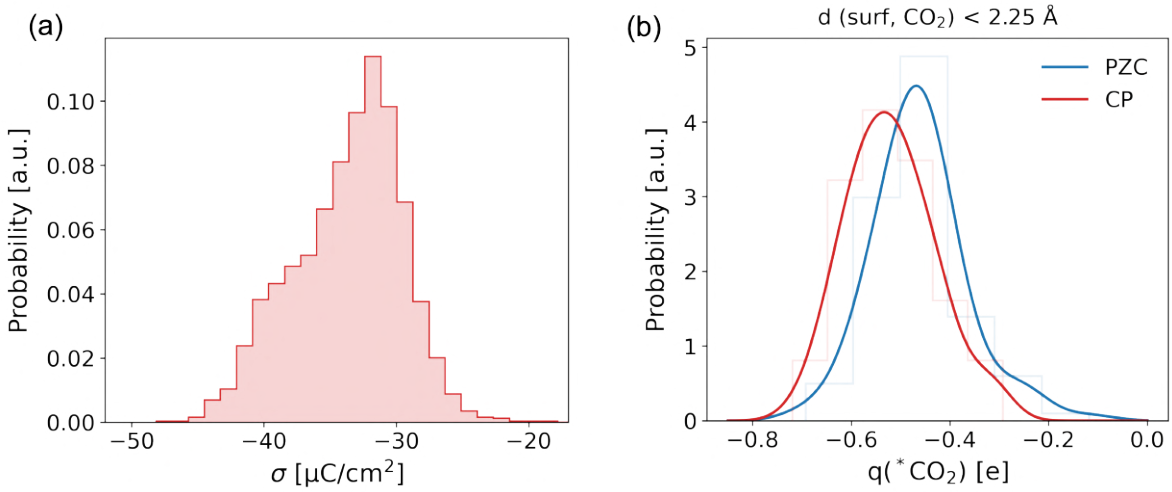

Figure 5.6: **Charges of surface and adsorbed CO_2.** (a) Distribution of surface charge density (σ) at CP condition. σ is calculated as the excess number of electrons ($N_{e,PZC} - N_e$) normalized by the surface area. The data is collected from the DFT dataset employed to train the MLP at CP. (b) Distribution of net Bader charge of $^*CO_2$ at PZC and CP conditions. The data were obtained from 105 representative structures in the training dataset for each scenario.

energy is 0.13 eV and 0.08 eV at PZC and CP conditions, respectively. The enhanced stability of $^*CO_2$ can be attributed to the increased surface charge density at CP (Fig. 5.6a), which facilitates charge transfer from the surface to the adsorbed CO_2, as confirmed by the Bader charge analysis (Fig. 5.6b) [242]. The effect of the electric bias is to decrease the barrier for adsorption from 0.20 eV at PZC to 0.16 eV at CP and shifts slightly the transition state (TS) position from d = 2.45 Å to d = 2.6 Å. This TS stabilization is likely due to its increased solvation at CP compared to PZC condition, since in the CP state, the free hydrogen atoms in the first H-density peak (~2.7 Å) could effectively form hydrogen bond with CO_2. In contrast, under PZC condition, the water structure is more rigid, thereby hindering interfacial water reorientation and CO_2 solvation. This is reflected in the histogram shown in Fig. 5.5c, where we observe a decrease in the percentage of configurations with zero hydrogen bonds and an increase in those where CO_2 is solvated by three or more water molecules under CP conditions.

Furthermore, to investigate the influence of the electrochemical potential on the dynamics of $^*CO_2$, we performed 1 ns unbiased molecular dynamics simulations under both PZC and CP conditions. The $^*CO_2$ adopts a bidentate structure and adsorbs near the top-site of the Ag surface. One C−O bond aligns parallel to the surface and coordinates with two sites, while the other C−O bond tilts toward the electrolyte (Fig. 5.7a,b). Furthermore, the $^*CO_2$ molecule exhibits large diffusion at CP exploring a wide range of adsorption sites, whereas it remains largely confined to a few sites under PZC conditions (Fig. 5.7c,d). Our results differ significantly from those reported in previous studies based on implicit solvent models [243, 225]. The most remarkable difference is that, in the implicit solvent framework, CO_2 does not chemisorb on the surface under PZC conditions, as indicated by a positive adsorption free energy ($\Delta G_{ads} > 0$). Moreover, the influence of the ap-

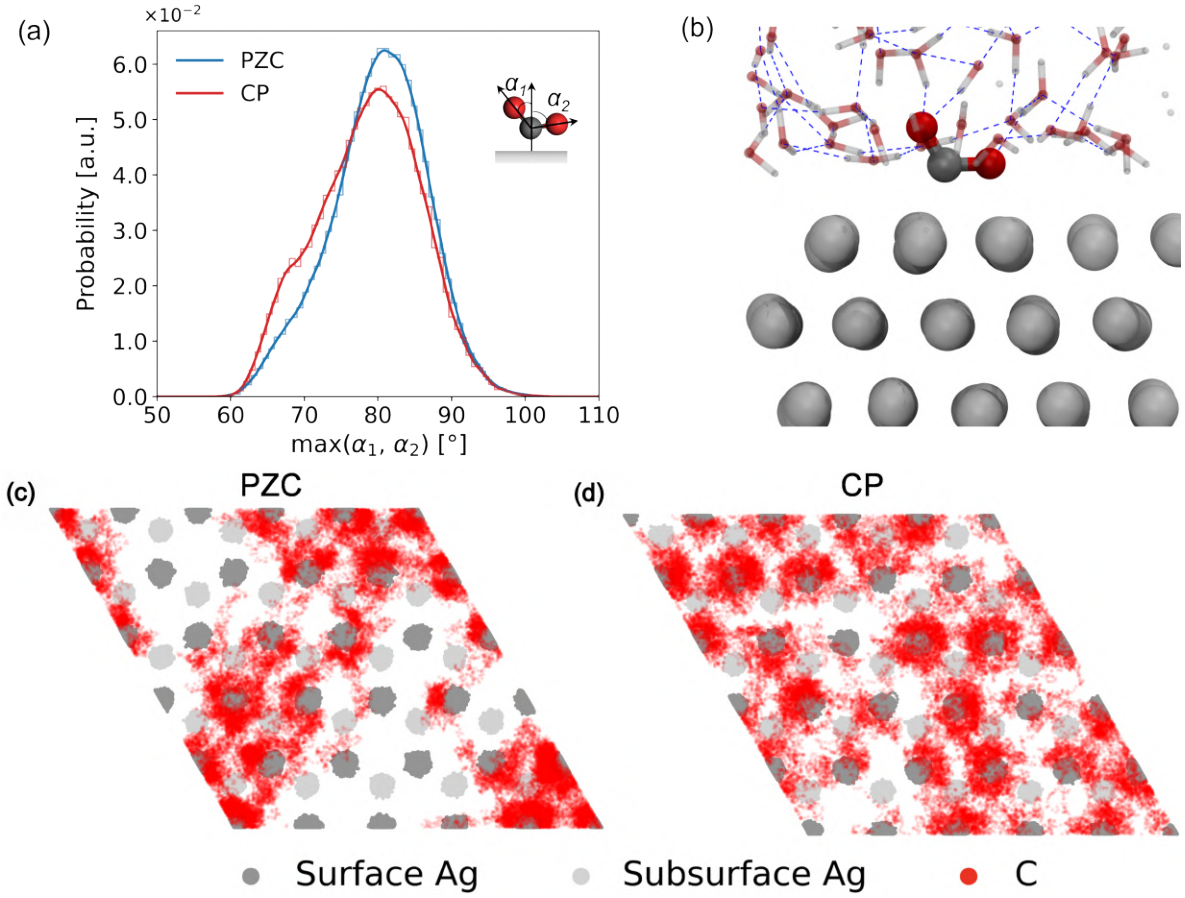

Figure 5.7: **Adsorbed CO_2 dynamics**. (a) Probability distribution of the maximum angle between carbon-oxygen bonds of $*CO_2$ and the surface normal (α_1 and α_2, denoted in the inset). Data were collected from 1 ns unbiased trajectories of $*CO_2$ under PZC and CP conditions. (b) A representative snapshot depicting the adsorption structure of CO_2 on the Ag surface, extracted from the trajectory at CP condition. The colors in the snapshot represent Ag in light gray, C in dark gray, O in red, and H in white. Hydrogen bonds are indicated by blue dashed lines. Water molecules are shown in transparent stick type for visual clarity. Scatter plots of Ag and C atom positions of $*CO_2$ under (c) PZC and (d) CP conditions, as seen from the top view. The Ag atoms of the two top layers are plotted. Data were collected from 1 ns unbiased trajectories of $*CO_2$, with a 25 fs frame interval reported.

plied potential on the free energy profile is much more pronounced in the implicit solvent case, with marked shifts observed as the potential is varied from PZC to negative potentials. We further investigated this effect by training MLP models at different applied potentials using an implicit solvent environment, as detailed in § 5.5 (Fig. 5.15), recovering the literature results.
Our findings suggest that the inclusion of explicit solvent in our simulations plays a key role in moderating the effect of the applied potential. The explicit solvent environment appears to partially screen the electric field at the interface, thereby reducing the potential-induced shifts in the free energy during CO_2 adsorption. This comparison underscores the crucial role of solvent representation in electrochemical modeling and highlights the necessity of incorporating explicit solvent to accurately capture interfacial molecular interactions and reaction mechanisms.

5.2.3 *CO_2 hydrogenation step

The adsorption of CO_2 involves an electron transfer from the metal surface to the molecule, leading to a slightly negative intermediate. However, this intermediate is not sufficiently activated for C−O bond cleavage, necessitating its hydrogenation as a critical step toward CO formation.
To investigate the hydrogenation of *CO_2, we designed two collective variables (CVs): the first CV is the coordination number between the oxygen atoms of CO_2 and the hydrogen atoms of all water molecules [$\mathrm{C.N.}(O_{cd}, H_w)$], which is able to describe the proton transfer from H_2O to *CO_2. The second CV is the ion-pair distance (d_{ions}) between *COOH and OH^-, that accounts for the formation of the final product.
The resulting free energy profile projected along the ion-pair distance is shown in Fig. 5.8a, while representative snapshots are shown in Fig. 5.8b. The reaction proceeds via three different stages, as reflected in the FES profile. The first one involves the proton transfer from one water molecule to *CO_2, leading to the formation of *COOH and OH^- ion pair (first proton transfer, 1st PT). However, this ion pair is initially tightly bound, exhibiting a strong tendency to recombine and return to the reactant state. To stabilize the protonated product, the *COOH-OH^- pair must dissociate. This occurs through the diffusion of OH^- away from *COOH (Stage 2, 2nd PT), facilitated by the Grotthuss mechanism, consistent with previous reports [244, 242]. Once the distance between *COOH and OH^- exceeds 3.2 Å, the free energy profile reaches a plateau (Stage 3).
An interesting feature emerges in the stage 3 region of the free energy profile, where the free energy remains nearly constant with increasing ion-pair distance under PZC condition, but gradually decreases beyond ~ 5 Å under CP condition. To explore the origin of this behavior, we carried out 10 independent unbiased molecular dynamics simulations under PZC and CP conditions, respectively. Each trajectory was initialized with a single OH^- ion positioned above the surface and run for 1 ns. The statistical distribution of the z-distance between OH^- and the surface is shown in Fig. 5.8c. At PZC, the OH^- ion remains tightly bound to the electrode, diffusing primarily within the interfacial region. In contrast, under CP condition, it rapidly desorbs from the surface and diffuses into the bulk region. This analysis highlights the strong influence of electrochemical conditions on OH^- dynamics. Under negative potentials, lower than PZC, the OH^- ion is electrostatically repelled from the electrode, whereas at PZC, its diffusion is hindered. Additionally, the orientation of OH^- near the interface exhibits distinct behavior under these two potential conditions (inset of Fig. 5.8b). Under CP condition, OH^- tends to adopt a vertical orientation, with the O−H bond

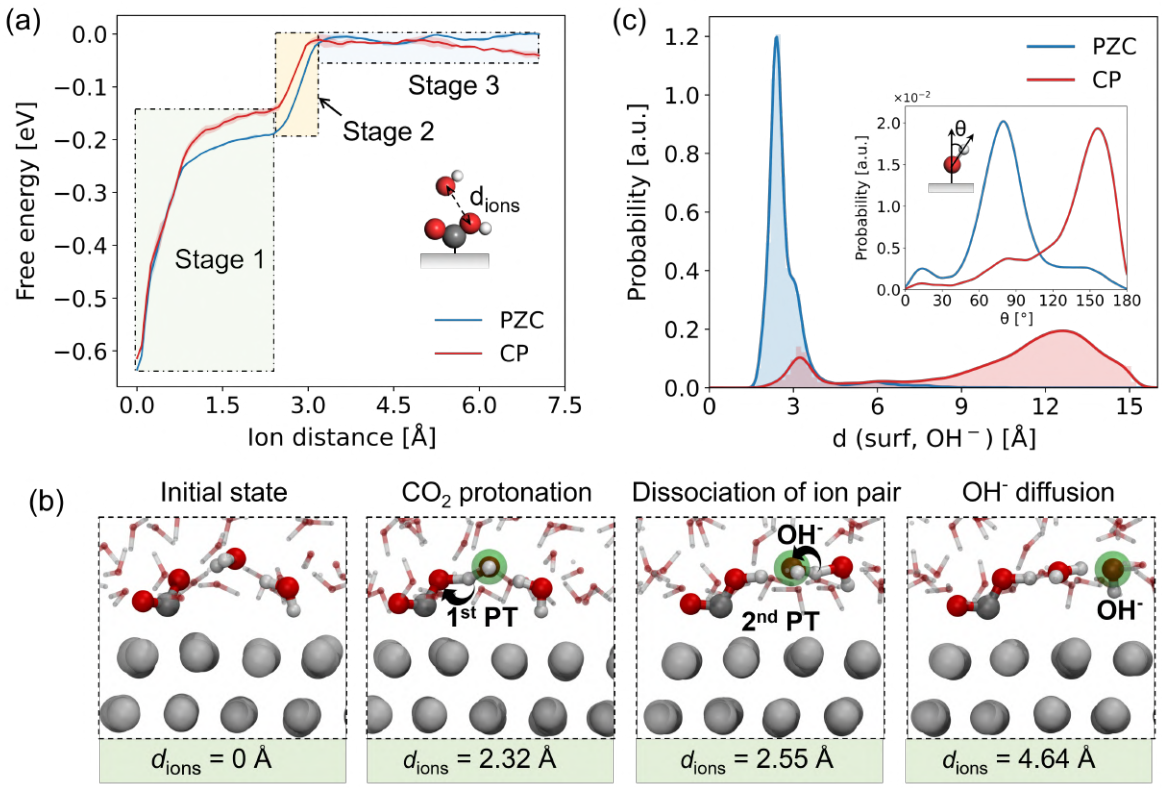

Figure 5.8: **Influence of potential on $^*CO_2$ hydrogenation process.** (a) One-dimensional projection of the FES along the distance of ions (*COOH and OH^-), with three distinct stages identified: CO_2 protonation (Stage 1), dissociation of the ion-pair of *COOH and OH^- (Stage 2), and OH^- diffusion (Stage 3). The endpoint of stage 2 is used as the reference point to align the free energy curves to allow a direct comparison of the free energy of OH^- diffusion under two potential conditions. (b) Representative snapshots illustrating three stages of $^*CO_2$ hydrogenation process, captured from the biased trajectory at CP condition. The colors in the snapshots represent Ag in light gray, C in dark gray, O in red, and H in white. Only the water molecules involved in the reaction are shown in ball-stick style, while others are displayed in transparent stick style for clear illustration. The green transparent ball illustrates the location of hydroxyl or OH^-. (c) Probability distribution of the vertical distance between OH^- and the surface from 10 independent unbiased trajectories of OH^- diffusion. The inset shows the probability distribution of OH^- orientation when the ion stays within the interfacial region (with a vertical distance less than 5 Å).

pointing toward the surface. In contrast, at PZC, the ion predominantly adopts a horizontal configuration, lying flat on the electrode surface while remaining confined to the interfacial region. The freedom of OH^- to diffuse away from the electrode results from the weakening of the hydrogen-bond network in the interfacial region at CP, where the interfacial water molecules reorient such that one hydrogen atom points toward the electrode, facilitating its diffusion via Grotthus mechanism. Together, these effects promote OH^- diffusion into the bulk and contribute to the lower free energy barrier for ion-pair dissociation observed in stage 2. This change in diffusion also explains

the energy trend in stage 3: restricted OH^- motion at PZC leads to a nearly constant free energy, while bulk diffusion at CP results in a gradual decrease of the free energy.

5.2.4 *COOH splitting mechanism

Before proceeding to describe the last electrochemical step that leads to the formation of *CO, we replaced the OH^- formed in the previous step by one water molecule to simulate the neutralization effect of the buffer solution under experimental conditions [242].
Once electrolyte neutrality is re-established, the *COOH intermediate remains stably adsorbed on the surface under both PZC and CP conditions. Furthermore, the length of the protonated C−O bond in *COOH is slightly longer under CP compared to PZC by approximately 0.12 Å. This bond elongation suggests greater activation of the C−O bond for cleavage under the effect of the applied potential. The free energy profile for the *COOH dissociation process as a function of the C−O distance ($d_{C,O_{OH}}$) is shown in Fig. 5.9a. Three distinct stages can also be observed here, and these are: C−O bond cleavage (Stage 1), first proton transfer (1^{st} PT, Stage 2), and OH^- diffusion (Stage 3) (Fig. 5.9b).
The transition state appears at a $d_{C,O_{OH}}$ of 1.94 Å under both potential conditions, with similar activation energies (0.53 eV at PZC and 0.49 at CP, respectively). As observed in the $^*CO_2$ hydrogenation step, the free energy profile gradually decreases under CP condition as the $d_{C,O_{OH}}$ distance increases beyond ~ 5 Å, while it remains nearly flat at PZC. This behavior can be explained similarly to the $^*CO_2$-to-*COOH hydrogenation case, considering the distinct OH^- dynamics at PZC and CP. Indeed, also here, the formed OH^- remains surface-bound at PZC, whereas under CP condition, it tends to diffuse into the bulk electrolyte, contributing to the observed decrease in free energy.
From a mechanistic point of view, our simulations suggest that the cleavage of the C−OH bond in *COOH is not initiated by protonation of the hydroxyl group. Instead, the bond breaks first, resulting in the formation of an OH^- species, which subsequently migrates away from the surface via the Grotthuss mechanism. This sequential bond cleavage followed by proton transfer contrasts with the mechanism proposed in previous studies based on the computational hydrogen electrode (CHE) model [245, 246], where this step is typically treated as a concerted proton-coupled electron transfer process.

5.2.5 *CO desorption process

As the final step, we have investigated the desorption of *CO. The z-distance between the carbon atom of CO and the surface is selected as the CV to obtain the FES of the desorption process, and the resulting free energy surface is shown in Fig. 5.9c. The equilibrium position of the *CO lies at approximately 1.7 Å from the surface, sitting in the first layer of interfacial water. As shown in Fig. 5.10a-b, *CO primarily binds to the hollow site of the second Ag layer, but migrates through multiple adsorption sites under both potential conditions. Similar to the behavior of $^*CO_2$, *CO exhibits larger mobility under CP condition, sampling nearly all surface sites within 1 ns. The FES reveals that the barrier for CO desorption is 0.07 eV lower under CP condition compared to PZC. Interestingly, the TS is located approximately 5 Å above the surface, which corresponds to the thickness

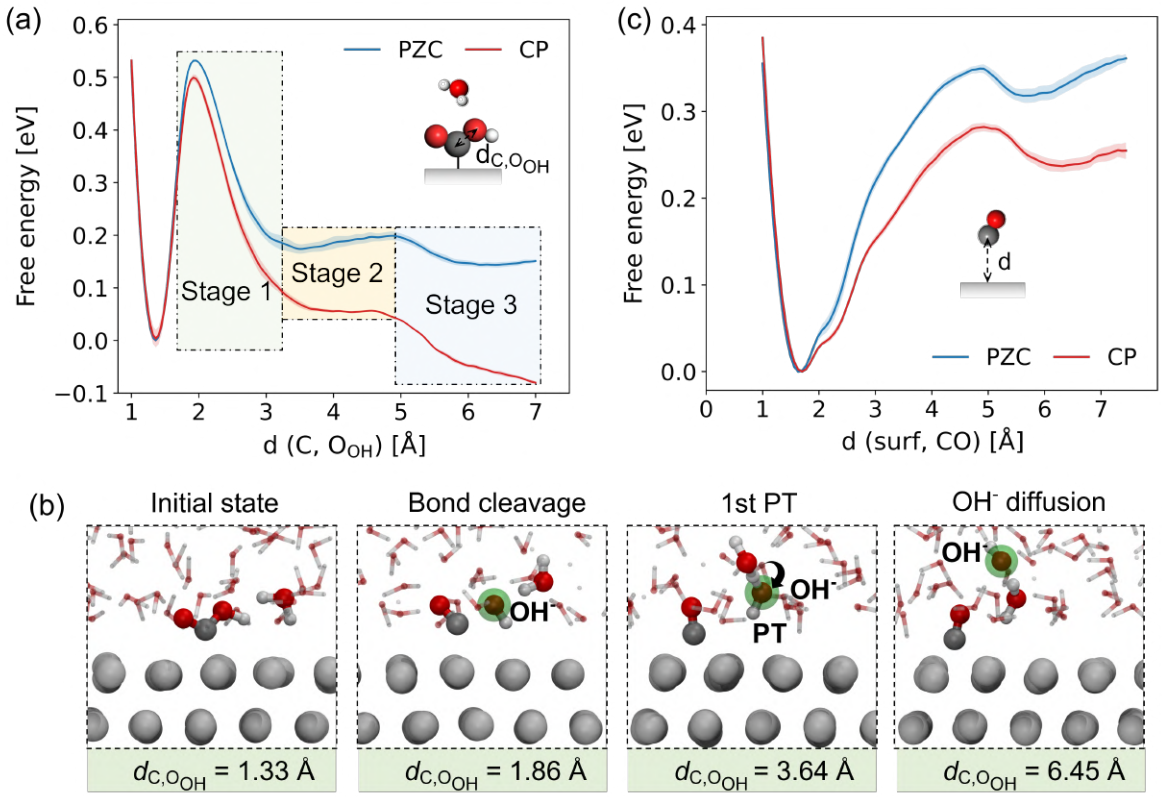

Figure 5.9: ***COOH splitting and CO desorption processes.** (a) One-dimensional FES of *COOH splitting along the distance between C atom and O atom of hydroxyl/hydroxide group, with three stages identified: bond cleavage (Stage 1), first proton transfer of OH^- (Stage 2), and OH^- diffusion (Stage 3). (b) Representative snapshots illustrating three stages of the *COOH splitting process, captured from the biased trajectory at CP condition. The colors in the snapshots represent Ag in light gray, C in dark gray, O in red, and H in white. The green transparent ball illustrates the location of OH^-. Only the water molecules involved in the reaction are shown in ball-stick style, while others are depicted in transparent stick style for clear illustration. (c) One-dimensional FES of CO desorption process along the vertical distance of CO from the surface. The reference minima of the free energy profiles are aligned to the initial states in both FES plots for better comparison of the free energy difference at PZC and CP conditions.

of the interfacial solvent layer. Bader charge analysis (Fig. 5.10c) indicates that the net charge of *CO is nearly identical in both potential conditions, with the adsorbed molecule being slightly negatively charged. At CP the electrostatic repulsion between the negatively charged surface and *CO weakens the Ag—CO bonding interaction, lowering the barrier for *CO desorption. Furthermore, the desorption process requires the breaking of the interfacial hydrogen bond network and releases a surface adsorption site, which is subsequently reoccupied by a nearby water molecule [247]. At CP, the weakened hydrogen-bond network facilitates both the migration of CO through the interfacial region and the rapid reorganization of water molecules to refill the adsorption site. These two effects collectively contribute to lowering the desorption barrier under CP condition. Therefore, although CO desorption is typically considered as a non-electrochemical process, our results indicate that it is indirectly influenced by the applied potential.

5.3 Discussion

The schematic mechanism for the overall CO_2 electroreduction to CO under working electrochemical conditions is shown in Fig. 5.11. Our simulations identify the hydrogenation of $*CO_2$ as the rate-determining step, with a free energy barrier of 0.59 eV at –1.5 V vs. SHE. Crucially, this barrier is not simply dictated by the direct electrostatic stabilization induced by the applied potential, but is strongly modulated by the electric-bias-driven reorganization of the interfacial water network. Prior studies often underestimated this barrier by overlooking the dissociation dynamics of the *COOH-OH^- ion pair. Our enhanced sampling-based approach captures this dissociation as a distinct, water-mediated step controlled by proton transfer and solvent structure, refining the mechanistic picture of CO_2 reduction.

Indeed, a central insight from our work is that the electric bias reshapes interfacial water structure and dynamics, which in turn governs key reaction steps. Under negative potential, water molecules adopt a “H-down” orientation and form a weakened hydrogen-bond network. This disruption reduces interfacial rigidity, enhancing solvation flexibility and facilitating proton transfer and OH^- diffusion via the Grotthuss mechanism. These solvent-mediated effects are not secondary, as OH^- diffusion plays a critical role across multiple steps and is highly sensitive to interfacial water structure. At PZC, OH^- remains confined near the surface, whereas under negative potential, it diffuses into the bulk, stabilizing intermediates and contributing to the overall free energy landscape. Even CO desorption, traditionally seen as a non-electrochemical event, is indirectly modulated by the interfacial solvent. The weaker hydrogen-bond network at negative potential lowers the energetic barrier for CO to escape into the electrolyte. Thus, even non-redox steps become field-sensitive through water-mediated effects.

This broader understanding of the reaction mechanism also challenges once more the traditional idea that catalytic activity depends only from specific sites on the surface [47, **?**, 28, 9]. The active catalytic environment should instead be seen as the coupled system of the electrode and the interfacial water layer. Moreover, these are not purely local effects confined to the adsorption site, but they reflect a non-local, collective response of the interfacial layer to the electric potential. As such, the catalytic performance emerges from global electric bias-induced restructuring of the electrode–electrolyte interface, rather than arising purely from local adsorption interactions at

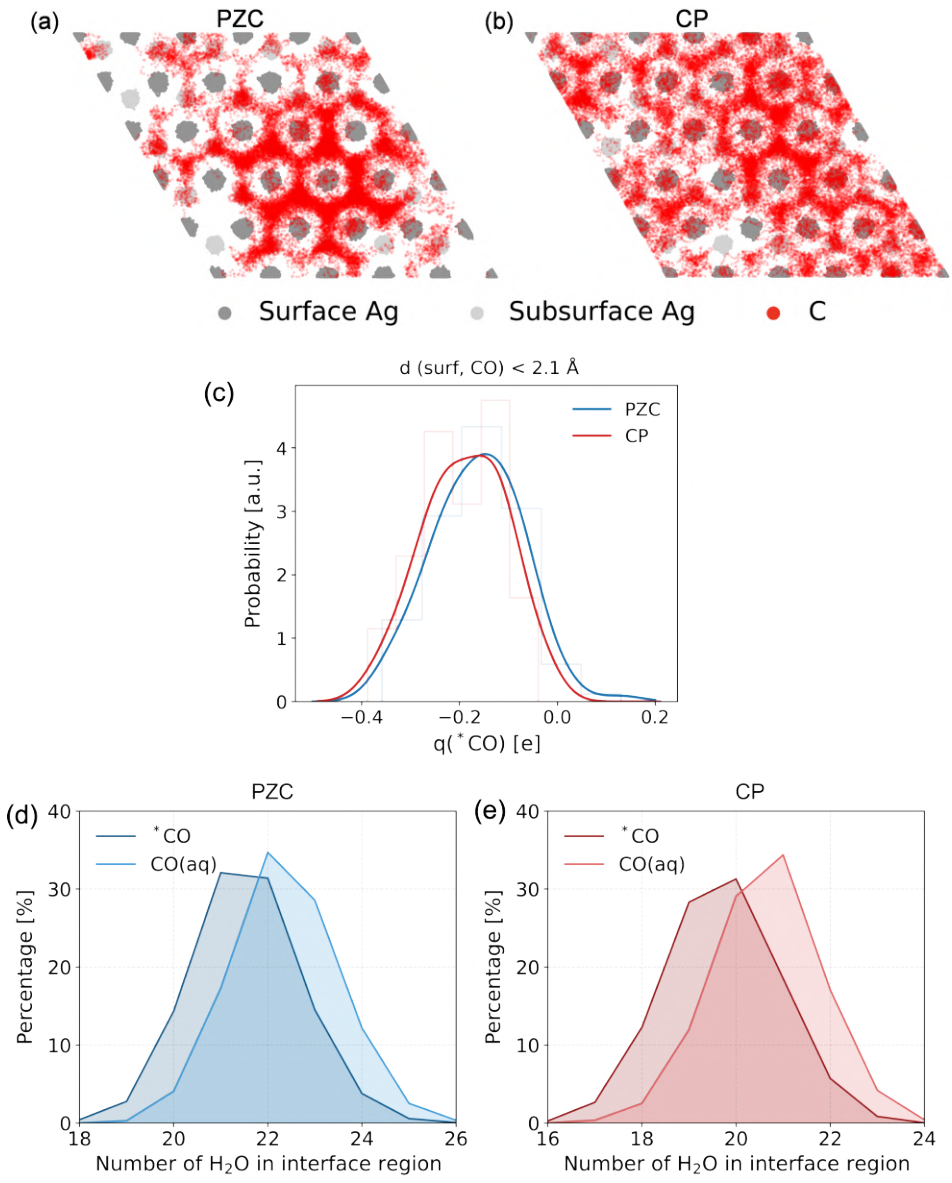

Figure 5.10: **Adsorbed CO dynamics**. Scatter plots of Ag and C atom positions of *CO under (a) PZC and (b) CP conditions, as seen from the top view. The Ag atoms of the two top layers are plotted. Data were collected from 1 ns unbiased trajectories of *CO, with a 25 fs frame interval reported. (c) Distribution of net Bader charge of *CO at PZC and CP conditions. The data were obtained from 105 representative structures in the training set at each scenario. Number of interfacial water molecules before (d<3 Å) and after (d >6 Å) CO desorption under (d) PZC and (e) CP conditions. d is the vertical distance of CO from the Ag surface.

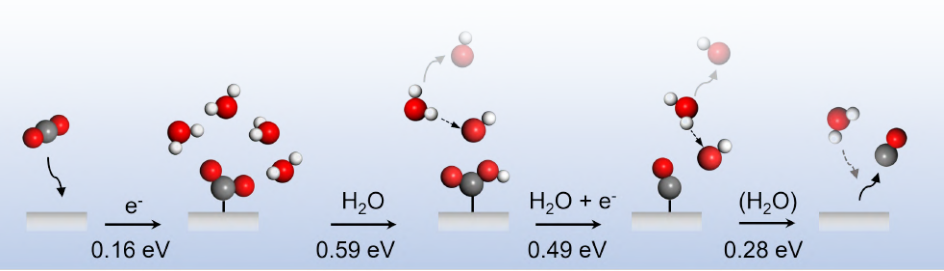

Figure 5.11: **Mechanism of CO_2 reduction to CO.** Schematic illustration of the proposed reaction pathway for CO_2 reduction in a neutral electrolyte under electrochemical working potentials. The free energy barrier of each elementary reaction step at an applied potential of -1.5 V vs. SHE is shown below the corresponding arrow. The free energy barriers ($\Delta G^{\ddagger}$ and free energy differences ΔG) of each elementary step are listed in Table.

the surface.

Our results underscore the need to treat solvent explicitly and dynamically, and they establish machine learning-accelerated molecular dynamics coupled with enhanced sampling techniques as a powerful framework for capturing the complexity of electrochemical interfaces. Looking ahead, tuning this interfacial environment offers a promising route to enhance CO_2 reduction kinetics. In practical systems, alkali metal cations (Na^+, K^+, Cs^+, etc.) play a crucial role at the electrode-electrolyte interface by promoting surface charge accumulation, modulating the interfacial electric field, and stabilizing key reactive intermediates [241, 248, 249]. While explicit cation effects were not included in the present simulations, their strong coupling to electrochemical potential-driven water dynamics makes a systematic investigation of their role under working conditions a promising direction for future studies, and our computational framework is readily extendable to incorporate these effects.

5.4 Methods

DFT single point calculations

The Ag(111)/H_2O interface was modeled using 5-layer p(3×3) and 4-layer p(4×4) Ag(111) slabs, with 30 and 58 water molecules placed above the Ag surface, respectively, corresponding to a density of 1 g/cm^3. All DFT energy and force calculations were performed with Vienna Ab initio Simulation Package (VASP) [250, 251], where the projector augmented wave (PAW) pseudopotentials were used [252, 253]. The electron exchange-correlation potential was described using r^2SCAN functional [254], with rVV10 correction applied to account for the nonlocal van der Waals interactions [255]. The kinetic energy cutoff for the plane wave basis expansion was set to 500 eV. The first order Methfessel-Paxton scheme with a smearing width of 0.2 eV was used. The convergence threshold for self-consistency was set to 1×10^{-6} eV. The Brillouin zone was sampled using a k-point grid of $3 \times 3 \times 1$ and $2 \times 2 \times 1$ for p(3×3) and p(4×4) slab models, respectively. A 15 Å vacuum layer was perpendicularly added to the model to eliminate the spurious interactions between periodic

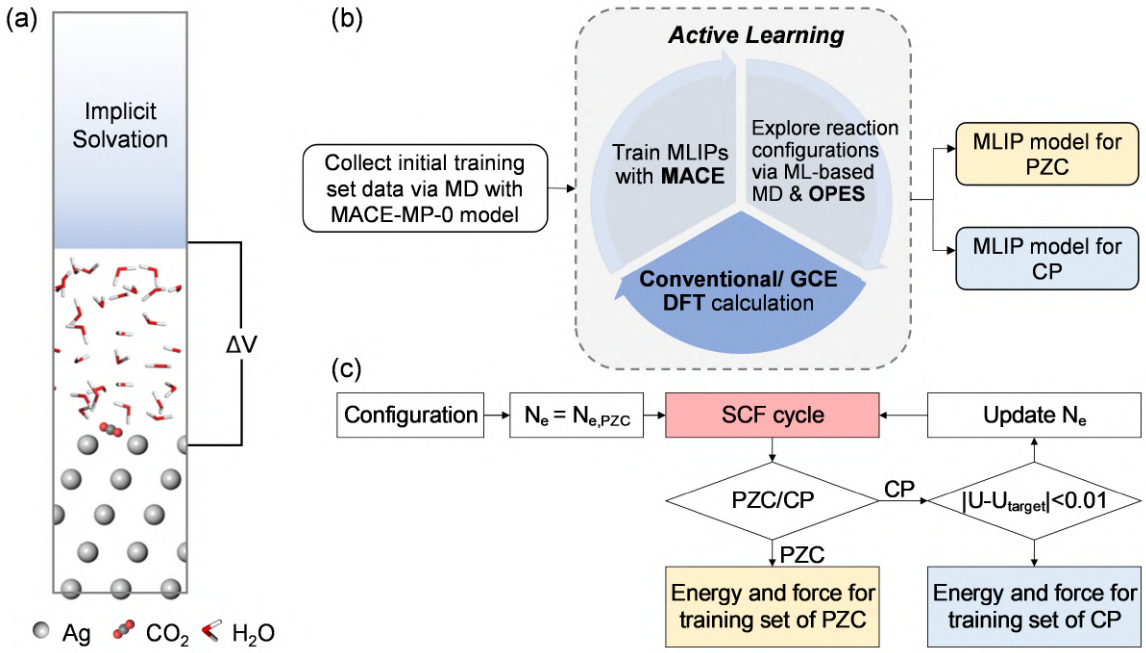

Figure 5.12: **Model and computational flowchart**. (a) Schematic illustration of the Ag(111)/H_2O hybrid explicit-implicit solvation model; (b) Active learning procedure for training MLPs under PZC and CP conditions; (c) Flowchart of conducting conventional/GCE DFT calculations.

images.

The self-consistent implicit solvation model implemented in VASPsol [256, 227] was employed at the region of vacuum layer, with the aim to calculate the electric potential (U) by referencing the Fermi level (E_{Fermi}) of the system against the standard hydrogen electrode (SHE):

$$U(\text{v.s. SHE}) = -\frac{E_{Fermi}}{e} - U_{SHE} \tag{5.1}$$

where U_{SHE} is the absolute electrode potential of SHE, with a fitted value of 4.21 V in this work. In VASPsol, the relative permittivity of the solvent was set to 78.4 to represent water, and the Debye length was set to 3.04 Å , corresponding to a bulk electrolyte concentration of 1M. A SOL-HYBRID model [60] was patched to the original VASPsol to avoid the possible nonphysical interaction of implicit solvation with explicit water molecules. This hybrid explicit/implicit solvent model has been validated for computing electrode potential and performing GCE-DFT calculations [229, 60, 257, 258].

Two types of DFT calculations were performed to calculate the energies and forces of configurations under PZC and CP conditions. At PZC, the number of valence electrons ($N_{e,PZC}$) was used, corresponding to the conventional constant-charge DFT calculations. In contrast, under CP condition, the number of electrons (N_e) of each configuration was externally optimized to reach the preset constant electric potential with a convergence threshold of 0.01 V, which is denoted as GCE-DFT. The optimization of N_e was implemented with a fully-converged constant potential (FCP) algorithm in ref. [106]. The potential-dependent electronic grand canonical free energy of

the system under CP condition was calculated as

$$\Omega_{el}(U) = E(U) - (N_e - N_{e,PZC}) \cdot E_{Fermi} \tag{5.2}$$

where $E(U)$ is the electronic energy under the applied potential.

Machine learning interatomic potentials

Two sets of MLP were constructed for describing CO_2 reduction reaction at the Ag(111)/H_2O interface under PZC and CP (-1.5 V vs. SHE) conditions, respectively. Fitting separate MLPs for a grand canonical potential energy surface at a desired electric potential is a practical strategy, since electrochemical reactions proceed at the electronic ground state in an open system [106]. In this work, each MLP was trained using MACE version 0.3.4 [148], one of the state-of-the-art graph neural networking (GNN) architectures, which combines equivalent message passing with a high body-order description of the local atomic environment via spherical harmonic polynomials, leading to excellent performance in accurate modeling of complex atomic interactions while maintaining efficiency [259]. Each model was trained with 128 equivariant messages and 2 interaction layers. A cutoff of 6 Å was employed. The dataset was split into training/validation subsets with a ratio of 95:5 with a batch size of 5 for training and 10 for validation. A weighted loss function was employed with force and energy weights set to 1000 and 10, respectively. The training process was carried out with 350 epochs, with last 20% of epochs switching to the Stochastic Weight Averaging strategies to lower the energy errors. Four different MLP models were trained to estimate model uncertainty using a query-by-committee approach.

Active learning procedure

The initial configurations of the training set were collected from MD simulations with the MACE-MP-0 [149, 150], a pre-trained MLP model that covers 89 elements on 1.6 M bulk crystals in the MPTrj dataset. The model has been proved to reasonably describe the behavior of solid-liquid interfaces [259]. We then applied an active learning protocol to generate new configurations through MD simulations with MLPs that have been built during training. Specifically, we evaluated the maximum standard deviation of the atomic forces predicted by the committee of four different MLP models in each electric potential scenario (see § 3.5.5). Training configurations were selected according to their associated σ values, following a strategy similar to that adopted in ref. [47, 28].
Even with the availability of MLPs, numerous reactive process, such as complex chemical reactions, typically occur on timescales far exceeding those accessible by performing standard MD simulations. The limitation makes it prohibitive to thoroughly sample a broad spectrum of reactive species across all possible reaction pathways. To this effect, we combined the protocol with enhanced sampling methods to ensure a comprehensive exploration of the reaction landscape of CO_2 reduction. This procedure was iterated until reliable MLPs capable of fully describing the CO_2 reduction to CO were obtained. The illustration of our workflow is depicted in Fig. 5.12. It is important to note that the training of MLPs for PZC and CP conditions are independent processes, except for the collection of initial configurations during the first cycle. Ultimately, we constructed

two training sets for describing CO_2 reduction process at the Ag(111)/H_2O interface under PZC and CP conditions, comprised of 6449 and 5777 configurations, respectively. The composition of two training sets is reported in Fig. 5.13: 600 configurations along the reaction pathways and transition states were selected as test sets for evaluating the accuracy of forces predicted by the MLPs under PZC and CP conditions, respectively (Fig. 5.14). The MLPs for modeling the CO_2 adsorption process at the Ag(111) surface with a fully implicit solvent model at PZC and constant potentials of -1 V, -1.5 V, and -2 V vs. SHE were obtained using the same protocol but trained separately. The final training set for these four MLPs consists of 911 configurations.

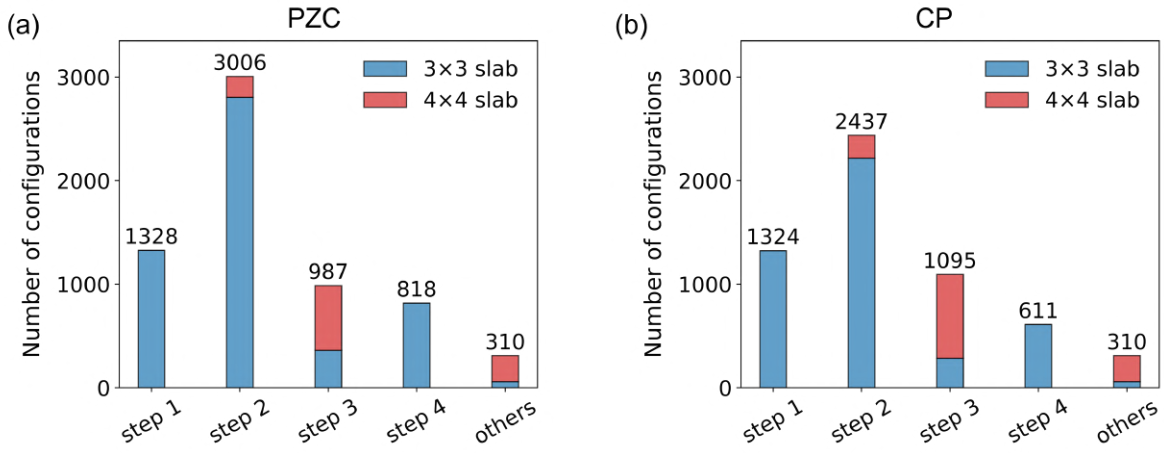

Figure 5.13: **Training set composition**. Histogram statistics of the training set for building the MLIPs under (a) PZC and (b) CP conditions.

Molecular Dynamics Simulations

Classical molecular dynamics simulations were preformed with the Large scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) program [216], patched with MACE 0.3.4 and PLUMED 2.9 [127]. NVT simulations were performed with an integration time step of 0.5 fs. The temperature was controlled using stochastic velocity scaling thermostat [83] with a coupling constant of 50 fs. During the active learning stage, simulation were performed with p(3×3) and p(4×4) Ag(111)/H_2O models. One interfacial water molecule would be substituted by CO_2 molecule when considering the reaction of CO_2 to CO conversion. When the final MPILs were obtained, a 5-layer p(6×6) Ag(111) slab containing 180 water molecules (the thickness of explicit water is around 20 Å and the density of water is 1 g/cm^3) was utilized for simulating properties of the Ag(111)/H_2O interface. The water density profiles in Fig. 5.2a confirm that the thickness of explicit water used in our simulation is sufficient to capture the transition from interfacial to bulk water properties. To evaluate the CO_2 reduction elementary reaction steps with OPES biased simulation, a 5-layer p(5×5) Ag(111) slab with 119 water molecules and one CO_2 molecule was used, representing the reaction in neutral electrolyte (pH=7). In all simulations, the bottom two layers of the Ag slab were fixed to simulate bulk phase structure. The temperature was set to 330 K to better describe

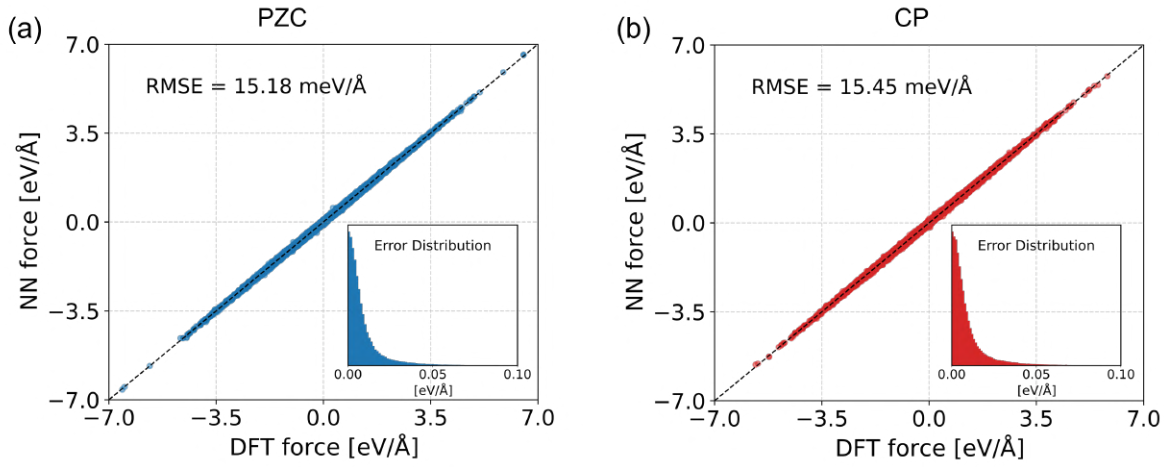

Figure 5.14: **Validation of the MLPs**. The atomic forces under (a) PZC and (b) CP conditions are computed with conventional/GCE DFT and corresponding MLIP model, respectively. Insets illustrate the probability distributions of the absolute difference of atomic force between the DFT and MLIP models. Each test set consists of 600 configurations collected from intermediates and transition states of CO_2R elementary reaction steps sampled during ML-based OPES simulations.

the geometrical properties of water at room temperature in all the simulations. The use of MLPs provided an approximately 10^5 fold speed-up over grand canonical DFT calculations, allowing the extensive sampling required to capture slow interfacial dynamics and obtain statistically robust thermodynamic and kinetic insights.

Enhanced Sampling simulations

The On-the-fly Probability Enhanced Sampling (OPES) [115], implemented in the open source PLUMED plugin, was used for exploring the reactive space and obtaining the free energy surfaces. Two types of collective variables (CVs) were employed during the OPES-biased simulations. One CV corresponds to the distance between two reactive intermediates or the vertical distance of the intermediate from the surface. The other CV represents the maximum coordination number, which captures the chemical bond formation and cleavage of critical components:

$$C.N._{\mathrm{A,B}} = \sum_{j \in \mathrm{B}}^{N_j} \frac{1 - \left(\frac{d_{ij}}{r_0}\right)^6}{1 - \left(\frac{d_{ij}}{r_0}\right)^{12}} \tag{5.3}$$

$$S_{\mathrm{AB}}^{\max} = \beta \log \sum_{i \in \mathrm{A}} \exp\left(\frac{C.N._{\mathrm{A,B}}}{\beta}\right) \tag{5.4}$$

Here, d_{ij} denotes the distance between atoms i of species A and atoms j of species B, and r_0 is the reference distance between two species, while β governs the smoothness of the soft maximum function, which is set to 0.05.

5.5 Supplementary note

The CO_2 adsorption barrier in Fig. 5.5b differs from those reported in previous studies where an implicit solvent model was used[243, 225]. To gain deeper insight into the role of explicit solvent in CO_2 adsorption, we trained four MLP models to describe the CO_2 adsorption process in a fully implicit solvent at PZC and constant potentials of -1V, -1.5V, and -2V vs. SHE, respectively. The resulting free energy profiles are presented in Fig. **??**. From the figure it is clear that no stable CO_2 adsorption state forms at PZC, and the free energy of $*CO_2$ decreases markedly as the applied potential shifts from PZC to -2V, with stable $*CO_2$ states only forming at -1.5V and -2V. Moreover, the barrier for CO_2 adsorption appears to be more sensitive to potential variation in the implicit solvent model. These findings suggest that the explicit solvent environment could partially buffer the influence of the applied potential during the CO_2 adsorption process, thereby moderating its effect on the free energy landscape.

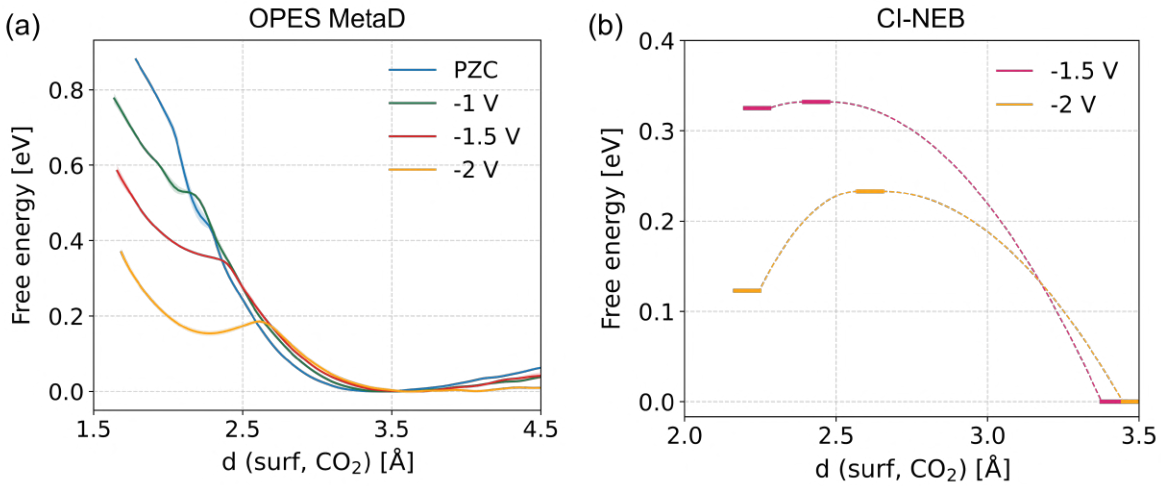

Figure 5.15: **Potential dependence of CO_2 adsorption process using an implicit solvent model**. (a) Free energy profiles of CO2 adsorption under PZC, -1V, -1.5V, and -2V vs. SHE. The results are obtained by conducting OPES enhanced sampling simulations at 330 K, with statistics errors shown as shaded regions. (b) Free energy barrier and difference under -1.5V and -2V. The transition states are identified with the climbing image nudged elastic band (CI-NEB) method. The barriers are not available for the PZC and -1V conditions as no stable chemisorbed (bent) CO2 can be obtained with structure optimization.

Chapter 6

Lithium-sodium alloys for ammonia synthesis

In the previous chapters, we explored how dynamic behavior is deeply influenced by operando variables. In this chapter, we present a compelling case where the catalytic system itself is inherently liquid, making atomic mobility and dynamic fluctuations intrinsic to its reactivity. Here, catalysis does not occur at fixed active sites, but rather emerges from the continuous restructuring of the system at the atomic scale. The following results are included in a manuscript under preparation, and are subject to a filed patent application [30].

6.1 Introduction

Efforts to identify novel catalytic systems have focused on materials such as alkali and alkaline earth-metal hydrides, liquid metals, and their alloys. In particular, liquid alloys such as Cu-Ga [70], Ga-Pt [260], and Li-Sn [77, 261] have already demonstrated promising catalytic properties. All these systems challenge the traditional notion of static active sites in catalysis. Instead, they feature dynamic active sites, driven by the intrinsic mobility of their constituent atoms, and exhibit significant system fluctuations. Indeed, emerging studies propose that system fluctuations play a pivotal role in driving catalytic reactions forward by enabling highly dynamic and adaptive active sites [28, 48, 47, 38]. Such a behavior becomes especially relevant near phase transitions, which frustrate the catalyst's structure and create critical fluctuations that enhance catalytic activity. These properties create a catalytic environment that transcends the limitations of static active sites, offering a new paradigm for catalyst design.

Alkali metals are known to react with N_2 to form nitrides, which are key intermediates in ammonia synthesis. For instance, sodium was reported to promote the nitridation of silicon [262, 263] and the dissolution of N_2 in the Ga−Na melt [264]. Furthermore, liquid Na was explored as a potential catalyst for NH_3 synthesis [76]. In the experimental setup proposed by Kawamura and Taniguchi, a gas mixture with composition 4%H_2-96%N_2 was directly blown into the liquid Na [265]. The ammonia production rate was reported to be $\sim$ 1e-4 μmol s^{-1}. However, its catalytic performance is hindered by the short dwell time of bubbles in liquid Na and their small reactive surface area. In addition, the formation of NaH may acts as a poison to the catalyst, forcing to operate at temperature $> 500^{\circ}$C to reduce its formation [76, 266]. Lithium (Li) has also attracted attention because of

its exceptional reducing ability: it can reduce N_2 to form Li_3N [267, 268] and dissociate H_2 to form LiH [269, 270]. These properties suggest that liquid Li could effectively promote NH_3 synthesis. To enhance catalytic performance and mitigate poisoning from hydride or nitride formation, we want to explore liquid mixtures that can suppress the nucleation of these phases.
As demonstrated by Rahim et al. [260], the atomic dispersion of active metals in liquid metal matrices prevents aggregation, while the inherent fluxionality of the liquid environment allows to solvate reactants and intermediates. This adaptability enhances reaction efficiency and selectivity, making liquid metal systems a promising avenue for advancing catalytic technologies.
To overcame the limitations of pure liquid alkali metals, here we propose the use of Li−Na mixtures. Extensive experimental data on compositions and miscibility are available for this system [271, 272, 273]. In particular, we explored the system with composition $Li_{0.64}Na_{0.36}$, corresponding to the consolute composition [274], aiming at investigating its catalytic properties for ammonia synthesis.

6.2 Results

6.2.1 $Li_{0.64}Na_{0.36}$ dynamics

The Li−Na system exhibits extensive immiscibility among alkali-metal alloys, giving rise to a broad liquid-liquid miscibility gap that persists over most of the composition range. Early calorimetric and chemical analyses established that the consolute point occurs near $x_{Li} = 0.64$ and T = 577± 2 K [272, 273]. Below this critical temperature, the alloy separates into two coexisting liquids, one Li-rich and the other Na-rich. This wide miscibility gap, driven by a large positive enthalpy of mixing [275], leads to nanoscale phase segregation that persists even in the molten state.
To analyze the segregation behavior, we performed MLP-based MD simulations of the $Li_{0.64}Na_{0.36}$ alloy for ∼10ns at a series of temperatures spanning the miscibility gap. We then constructed surface meshes based on the alpha-shape algorithm [194, 195], which allow a direct visualization of how atoms of the same species aggregate into Li- and Na-rich domains (Fig. 6.1). Representative MD snapshots at 500, 600, and 700 K (Fig. 6.1) reveal clear phase separation into Li- and Na-rich regions. With increasing temperature, the domains become progressively smaller and more diffuse, reflecting enhanced mixing. The system can thus be described as microheterogeneous, with Li-rich and Na-rich domains coexisting on the nanoscale, a feature that, as we shall discuss later, has functional implications in catalysis.
For a quantitative assessment, we computed the local Na molar fraction by dividing the simulation box into a three-dimensional grid and evaluating the the Na concentration in each voxel. The resulting distributions (Fig. 6.2a) are bimodal at low temperature, corresponding to the coexistence of Na-rich and Li-rich regions, but become unimodal with increasing temperature, indicating enhanced mixing. From this trend, the mixing-demixing transition temperature can be roughly estimated between 560 and 600K in qualitative agreement with the experimental phase diagram.
Further insight into the liquid structure is obtained from the pair distribution functions, $g(r)$, for Li−Li, Na−Na, and Li−Na pairs (Fig. 6.2b-d). The Li−Na correlation exhibits markedly lower peaks than the homonuclear correlations, indicating a pronounced homocoordination tendency.

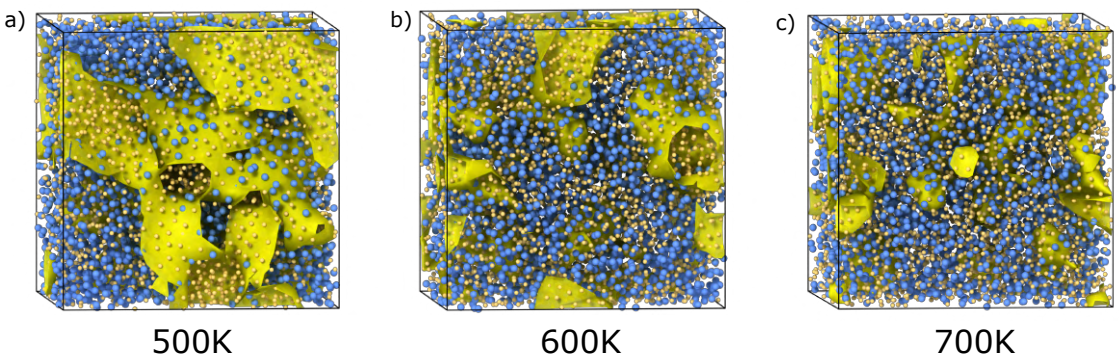

Figure 6.1: **Bulk $Li_{0.64}Na_{0.36}$ at different temperatures**. Snapshots of bulk $Li_{0.64}Na_{0.36}$ at a) 500K, b) 600K, and c) 700K. The yellow solid surfaces contains the domains rich in Li atoms and are computed using the alpha-shape method [194, 195] implemented in OVITO [186]. Details are reported in the Methods section 6.4.

With increasing temperature, the peaks of the Li−Li and Na−Na $g(r)$ functions decrease and the oscillations beyond the first shell become progressively damped, reflecting the gradual loss of short- and medium-range order as the melt approaches a more homogeneous state.
This observation is reinforced by the radial concentration correlation function (RCF(r)) [276] defined as:

$$\mathrm{RCF}(r) = 4\pi r^2 \rho x_{\mathrm{Li}} x_{\mathrm{Na}} \left[g_{\mathrm{Li-Li}}(r) + g_{\mathrm{Na-Na}}(r) - 2g_{\mathrm{Na-Li}}(r) \right] \tag{6.1}$$

Across all studied temperatures, RCF(r) remains positive (Fig. 6.2e), which is indicative of a marked tendency for homocoordination. Moreover, its amplitude decreases progressively with temperature and approaches zero at high T, reflecting the gradual loss of segregation as the melt becomes more homogeneous.
Finally, we analyzed the structure factors obtained from Fourier transforms of the partial pair correlation functions. Within the Bhatia-Thornton formalism [277], the total structure factor is decomposed into contributions from number-density and concentration fluctuations. Among these, the concentration–concentration term, $S_{CC}(q)$, is particularly informative, as its behavior reflects the balance between homo- and heterocoordination (Fig. 6.2f):

$$S_{CC}(q) = x_{\mathrm{Li}} x_{\mathrm{Na}} \left(1 + x_{\mathrm{Li}} x_{\mathrm{Na}} \left(S^{FZ}_{\mathrm{LiLi}}(q) + S^{FZ}_{\mathrm{NaNa}}(q) - 2S^{FZ}_{\mathrm{LiNa}}(q) \right) \right) \tag{6.2}$$

where $S^{FZ}_{ij}(q)$ is the Faber-Ziman partial structure factor obtained from the Fourier transformation of the radial distribution function for the ij pair. At high temperatures, $S_{CC}(q)$ remains low for all q values, indicating suppressed long-range concentration fluctuations. As the temperature is lowered toward the critical point, $S_{CC}(q)$ develops a pronounced increase at small q, signaling the growth of long-wavelength concentration fluctuations and the onset of compositional heterogeneity. By fitting the low-q behavior of $S_{CC}(q)$ to an Ornstein–Zernike–type Lorentzian form (Eq. 6.3), we extracted the correlation length $\xi(T)$ as a function of temperature.

$$S_{CC}(q) = \frac{a(T)}{\xi^{-2} + q^2} \tag{6.3}$$

The correlation length diverges according to the expected critical scaling

$$\xi(T) = \xi_0 \left(\frac{T - T_c}{T_c} \right)^{-\nu} = \xi_0 \varepsilon^{-\nu} \tag{6.4}$$

yielding a critical temperature $T_c = 577 \pm 1$ K and a critical exponent $\nu = 0.63$, consistent with the 3D Ising universality class. (Fig. 6.2f-inset). The computed critical temperature T_c is in agreement with both the experimental phase diagram and our voxel-based compositional analysis.

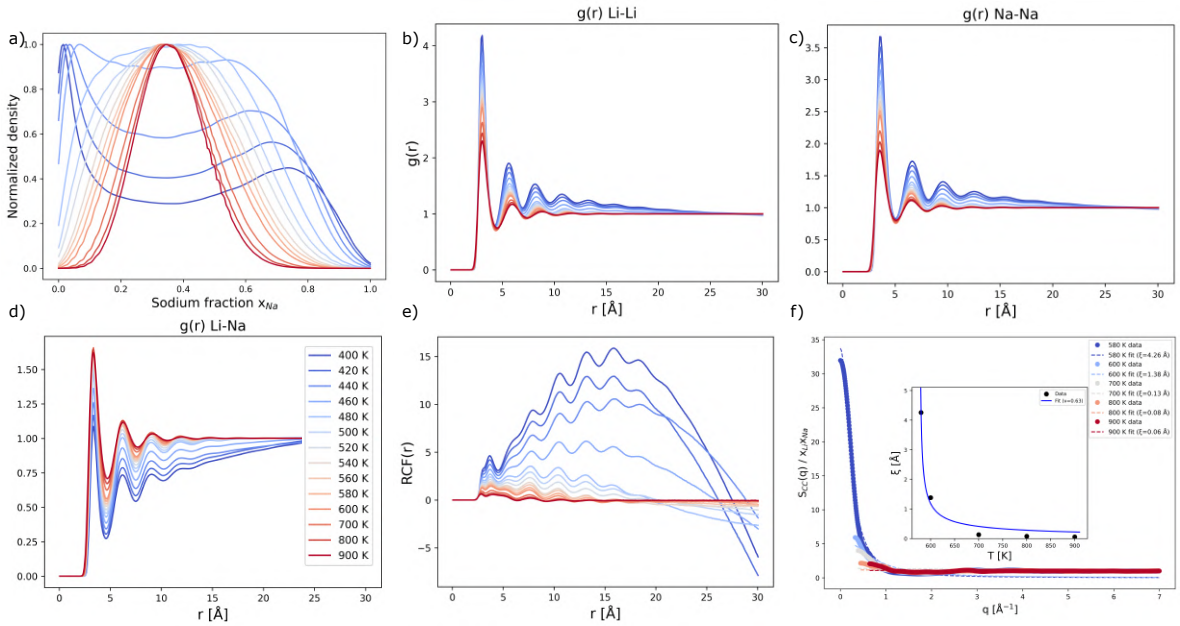

Figure 6.2: **Structure of bulk $Li_{0.64}Na_{0.36}$ at different temperatures**. a) d) Distribution of Na fraction in the simulation cell. Radial distribution function for b) Li−Li pair; c) Na−Na pair; d) Li−Na pair. e) Radial concentration correlation function RCF(r). f) Concentration-concentration structure factor $S_{CC}(q)$. The inset shows the temperature dependence of the correlation length $\xi(T)$, extracted by fitting the low-q region of $S_{CC}(q)$ to an Ornstein–Zernike form. Details of the simulations are reported in the Methods section 6.4.

6.2.2 H_2 and N_2 dynamics

To resemble the methodology proposed in ref. [76], we simulated a system in which N_2 and H_2 are let to react in the bulk alloy. To do so, we inserted the reacting N_2 and H_2 in the simulation cell. Interestingly, H_2 spontaneously dissociates into two hydride ions (2 H^-) at all temperatures in the range 500-700K. Since the pair distribution functions $g(r)$ show little temperature dependence across this range, we report here the results at 600K as representative (Fig. 6.3a). In the $Li_{0.64}Na_{0.36}$ alloy, the $g_{H-Li}(r)$ exhibits well-defined peaks at 1.85Å, 4.45Å, and 7.05Å indicating strong and structured coordination of H^- with lithium atoms. For the $g_{H-Na}(r)$, peaks appear at 4.45Åand 7.55Å, with the absence of a first-shell peak confirming a weaker interaction with sodium. These

results clearly indicate a preferential solvation of H^- by lithium in the alloy. For comparison, in pure liquid lithium, the $g_{H-Li}(r)$ shows peaks at 1.85Å, 4.35Å, and 6.95Å, consistent with the alloy case. In contrast, in pure sodium, the $g_{H-Na}(r)$ peaks are found at longer distances: 2.25Å, 5.35Å, and 8.45Å. Analysis of the local coordination environments reveals that in pure liquid lithium, H^- is surrounded by six Li atoms arranged in an approximately octahedral geometry. Instead, in pure sodium, H^- is preferentially coordinated by four Na atoms in a distorted tetrahedral configuration. In the $Li_{0.64}Na_{0.36}$ alloy, the local environment of H^- is composed a first coordination shell of Li atoms forming a distorted octahedron, surrounded by a second coordination shell of Na atoms (Fig. 6.3b). Regarding N_2, in pure sodium, the N−N bond length distribution exhibits a maximum

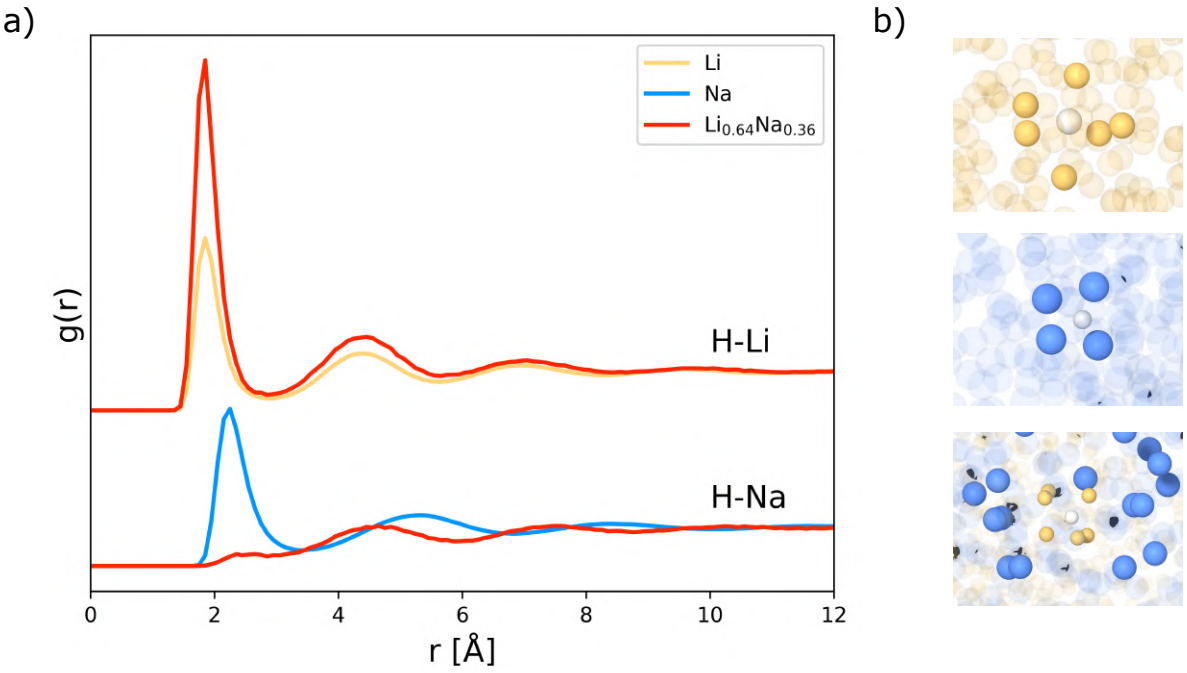

Figure 6.3: **Structure of H^-**. a) Radial distribution functions for H−Li and H−Na pairs at 600K in pure Li, pure Na and in the $Li_{0.64}Na_{0.36}$ alloy. b) Coordination environments of H^-. From top to bottom: H^- in Li, H^- in Na, and H^- in $Li_{0.64}Na_{0.36}$.

around 1.15Å at both 500K and 600K, indicating that the N_2 molecule largely retains its bond order at these temperatures (Fig. 6.4a). The distributions at 500K and 600K are nearly identical, suggesting minimal thermal influence on the bond length. At 700K, the distribution broadens significantly, reflecting greater bond length fluctuations and stronger thermal perturbation of the N_2 molecule within the liquid sodium matrix. The overall trend highlights the relative stability of N_2 in liquid sodium at moderate temperatures, and the onset of significant dynamic perturbations as the system approaches higher thermal regimes.

In contrast, in both liquid lithium and the $Li_{0.64}Na_{0.36}$ alloy, the N−N bond length distribution shows a peak centered around 1.4Å across all temperatures investigated (Fig. 6.4a). This value is larger than in pure sodium, indicating a substantial weakening and elongation of the N≡N bond due to the presence of lithium. Moreover, the distribution in the alloy is only weakly affected by

temperature, as the position and shape of the peak remain largely consistent between 500K and 700K. These results highlight the dominant role of lithium in promoting N_2 activation, with bond weakening primarily driven by the chemical environment rather than by thermal effects.

To further characterize the interaction between molecular nitrogen and the surrounding liquid metals, we computed the radial distribution function $g(r)$ for N-Na at 500K, 600K, and 700K (Fig. 6.4b). In all cases, the $g(r)$ shows no significant structuring: the distribution remains essentially flat, with no pronounced peaks at any distance. This lack of correlation indicates that the N_2 molecule and sodium atoms are largely unassociated in the liquid phase, with no specific solvation shell or preferential arrangement around N_2. The absence of local structuring around the nitrogen molecule is consistent with the observed preservation of its bond order, as reflected in the N−N bond length distributions.

The radial distribution function for N-Li provides insight into the local structuring of lithium atoms around the nitrogen molecule (Fig. 6.4c). In contrast to the N-Na case, the N-Li RDF exhibits distinct peaks, indicating a well-defined solvation structure. Well defined peaks are observed at approximately 1.98Å, 3.13Å, and 4.78Å, corresponding to successive coordination shells of lithium atoms around N_2. The first peak at 1.98Å reflects strong and direct interactions between lithium and nitrogen, consistent with the activation of the N≡N bond as seen in the bond length distributions.

To further characterize the local environment of molecular nitrogen in the $Li_{0.64}Na_{0.36}$ alloy, we employed the Localized Many-Body Tensor Representation (LMBTR) descriptor [278], as implemented in the DScribe library [279]. Using LMBTR to represent the atomic environment surrounding each N_2 molecule along the trajectory, we performed principal component analysis (PCA) to reduce the descriptor space and identify recurring structural patterns. The projection along the two main components revealed distinct clusters corresponding to different solvation geometries (Fig. 6.4d). In particular, we found that N_2 can be coordinated by 6 to 9 lithium atoms in its first solvation shell, depending on the instantaneous configuration and thermal fluctuations. These varying local environments suggest that N_2 experiences a dynamic coordination landscape in the liquid alloy, which may play a key role in its progressive activation and reactivity.

6.2.3 N_2 splitting

The free energy profiles as a function of the N−N distance for the reaction $N_2 \longrightarrow 2\ N^{3-}$ were analyzed in three liquid systems at 600 K: Li, Na and $Li_{0.64}Na_{0.36}$ (Fig. 6.5).

The energy barrier for N_2 dissociation in liquid Na is approximately 350 $kJmol^{-1}$, whereas in liquid Li, it is significantly lower, around 70 $kJmol^{-1}$. For the $Li_{0.64}Na_{0.36}$ alloy, the free energy profile closely resembles that of Li rather than Na. Moreover, the equilibrium distance between nitrogen atoms is shorter in Na compared to Li and the $Li_{0.64}Na_{0.36}$alloy, indicating the stronger reducing ability of Li.

The reaction free energy is positive in Na, suggesting that nitride formation is unfavorable. In contrast, it is negative in Li and $Li_{0.64}Na_{0.36}$, implying that nitrides are the most stable phase in these systems. Once nitrides form in Li and $Li_{0.64}Na_{0.36}$, a stable phase emerges, characterized by an N−N bond distance of approximately 3.5 Å, which can be associated with a Li_3N crystal seed.

The energy barrier for complete nitrogen dissociation is around 50 $kJmol^{-1}$ in pure Li and 30

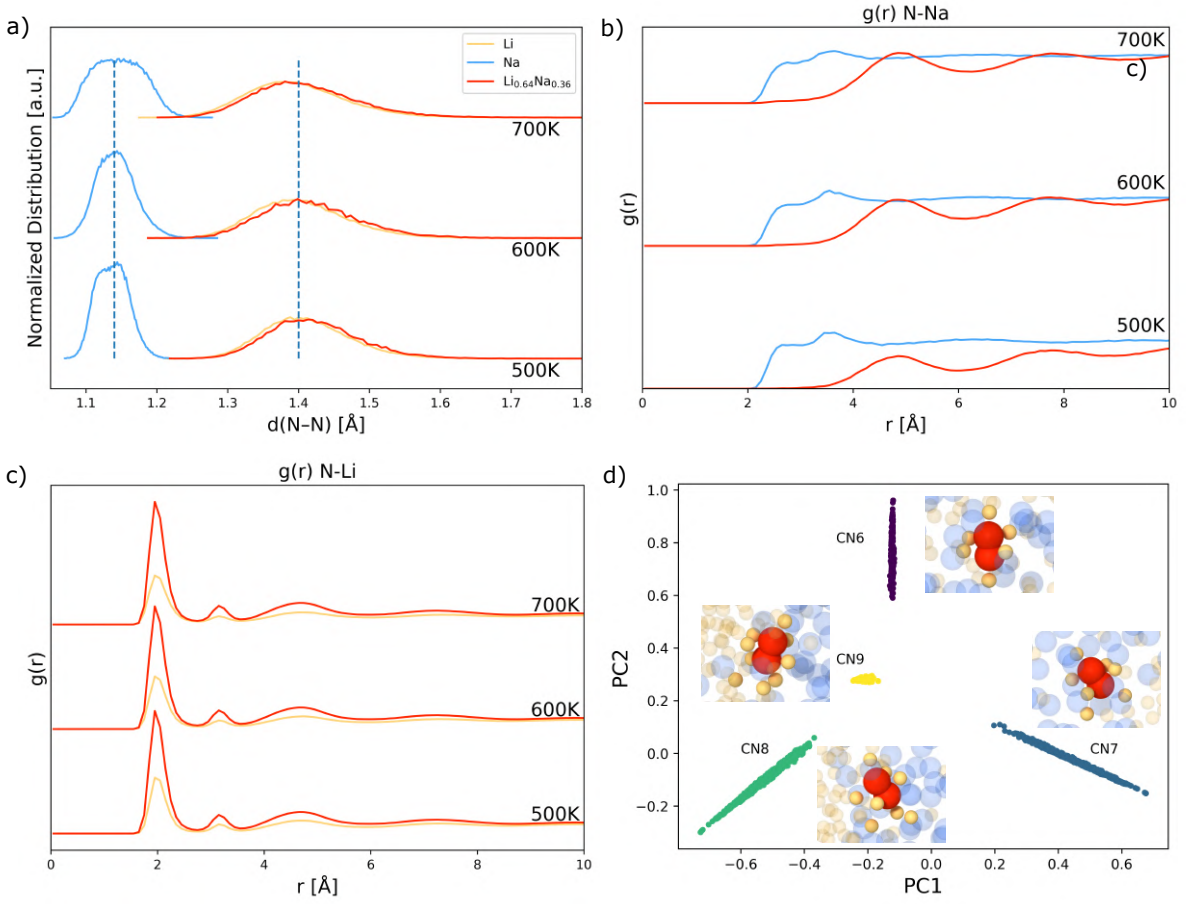

Figure 6.4: **Structure of N_2.** a) Distribution of the N−N distance at three different temperatures in pure Li, pure Na and in the $Li_{0.64}Na_{0.36}$ alloy. b) Radial distribution function for N−Na pair at three different temperatures. c) Radial distribution function for N−Li pair at three different temperatures. d) Clustering of different coordination environments of N_2 in $Li_{0.64}Na_{0.36}$. Representative configurations for each type of environment are shown in the inset figures. Details of the simulations are reported in the Methods section 6.4.

kJmol^{-1} in $Li_{0.64}Na_{0.36}$. The lower dissociation barrier in the alloy is attributed to the solvation of nitrides in Na. This suggests that Na can hinder the formation of stable Li_3N, which would otherwise act as a catalyst poison.

Upon formation, the N^{3-} anion exhibits an octahedral coordination environment in both pure Li and pure Na. In the $Li_{0.64}Na_{0.36}$ alloy, N^{3-} is solvated by a first coordination shell of six Li atoms arranged in an octahedral geometry, followed by a second solvation shell composed primarily of Na atoms (Fig. 6.5b). This solvation structure mirrors that observed for hydride ions in the alloy (Fig. 6.3b). The preference for Li coordination is also evident in the radial distribution functions $g(r)$ shown in Fig. 6.6b (red line), which display a pronounced first peak for N−Li interactions around 2Å, while the first peak for N−Na appears at a longer distance of approximately 5Å.

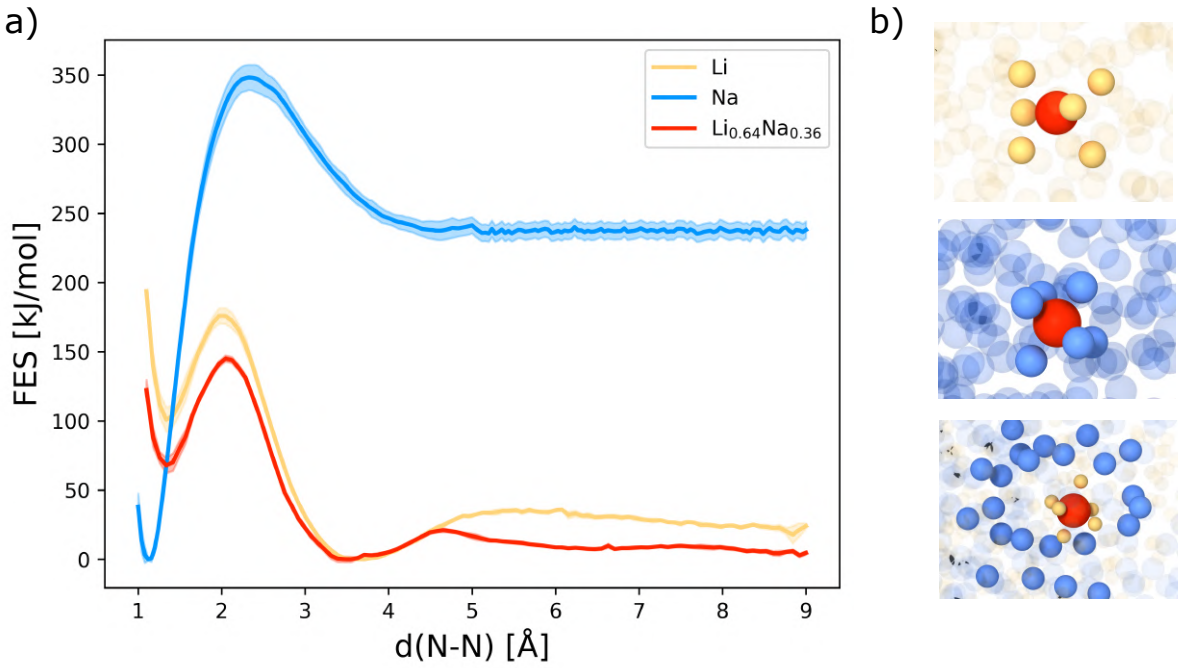

Figure 6.5: **Dissociation of N_2.** a) Free energy profiles for N_2 dissociation at 600K in pure Li, pure Na and in the $Li_{0.64}Na_{0.36}$ alloy along the distance N−N. Details of the simulations are reported in the Methods section. b) Coordination environments of N^{3-} after N_2 dissociation. From top to bottom: N^{3-} in Li, N^{3-} in Na, and N^{3-} in $Li_{0.64}Na_{0.36}$.

6.2.4 NH_3 synthesis

Following the dissociation of N_2 into N^{3-}, we performed simulations of the subsequent hydrogenation steps leading to NH_3. For each N atom, three H atoms were introduced into the system, and the N−H coordination number was employed as the CV. The corresponding FESs at 600 K are shown in Fig. 6.6a. As one can observe, the free energy barriers for the hydrogenation of N to NH_3 are significantly higher in pure Li compared to pure Na, which is in contrast with the behavior observed in the previous dissociation step (Table 6.1). For the $Li_{0.64}Na_{0.36}$ alloy, the barriers lie between the two limiting cases of pure Li and pure Na (Table 6.1). Moreover, as shown in the two-dimensional FES (Fig. 6.6b) the minima corresponding to the hydrogenated intermediates progressively broaden and shift toward higher values of the N−Na coordination number (CN(N,Na)). We refer to this increased affinity of nitrogen species for Na as *sodiophilicity* and is also reflected in the corresponding radial distribution functions $g(r)$ (Fig. 6.6c). As the degree of protonation increases, the intensity of the first peak in the $g(r)$ of N−Li gradually decreases, whereas the first N−Na peak becomes more pronounced and shifts toward shorter r values, indicating a stronger interaction between the protonated intermediates and sodium atoms. This tendency of the intermediates to be increasingly coordinated by Na also drives their location from the Li−Na interface (Fig. 6.6d). As the degree of protonation increases, the probability of finding NH_x species near the interface decreases, and the distribution progressively shifts to larger distances.

These findings point to a clear synergistic effect between the two metals: while Li facilitates N_2 dissociation, Na stabilizes the intermediates and lowers the barriers associated with the hydrogenation steps. We therefore argue that increasing the Na concentration can further reduce these barriers and enhance ammonia production. Optimizing the Li-Na ratio to balance efficient nitrogen dissociation and hydrogenation is subject of ongoing investigations.

Reaction	Li	Na	$Li_{0.64}Na_{0.36}$
$\Delta F^{\ddagger}$ (kJ/mol)			
$N^{3-} + H^- \longrightarrow NH^{2-}$	154.5	53.7	122.4
$NH^{2-} + H^- \longrightarrow NH_2^-$	165.7	77.5	154.8
$NH_2^- + H^- \longrightarrow NH_3$	200.8	145.4	188.0
ΔF (kJ/mol)			
$N^{3-} + H^- \longrightarrow NH^{2-}$	52.9	-175.0	5.8
$NH^{2-} + H^- \longrightarrow NH_2^-$	84.5	-74.5	43.5
$NH_2^- + H^- \longrightarrow NH_3$	112.5	6.6	89.1

Table 6.1: **Thermodynamics of NH_3 synthesis**. Activation energies ($\Delta F^{\ddagger}$) and free energy differences (ΔF) for the different steps in NH_3 synthesis for pure Li, pure Na and $Li_{0.64}Na_{0.36}$ alloy at 600K.

6.3 Discussions

The catalytic behavior of the $Li_{0.64}Na_{0.36}$ alloy for ammonia synthesis is controlled by its unique microheterogeneous structure, characterized by the coexistence of Li-rich and Na-rich domains. This liquid-liquid phase separation, driven by immiscibility and a large positive enthalpy of mixing [275], gives rise to a dynamic and adaptive interface that plays a key role in facilitating the overall reaction.

Our MLP-MD simulations reveal that the reactions are facilitated by the synergistic interplay between the components distributed across the two domains. Lithium atoms serve as the primary sites for N_2 activation and splitting, consistent with the strong N−Li interactions and the low energy barriers for dissociation. In contrast, sodium atoms progressively stabilize protonated intermediates (NH^{2-}, NH_2^{1-}, NH_3), as demonstrated by the increased sodiophilicity observed in the radial distribution functions and free energy surfaces. Functionally, the alloy mimics dual-site catalysis, where different catalytic tasks, N_2 activation and hydrogenation, are handled by different components. Moreover, the intrinsic mobility of the liquid system enables interfacial rearrangements and compositional fluctuations that dynamically adapt to the local chemical environment, creating favorable conditions for each step of the reaction. Thus, domain mobility and compositional fluctuations act in concert, enhancing the catalytic performance beyond what either metal could achieve alone.

Importantly, the Li−Na ratio serves as a tunable parameter that controls the size and distribution

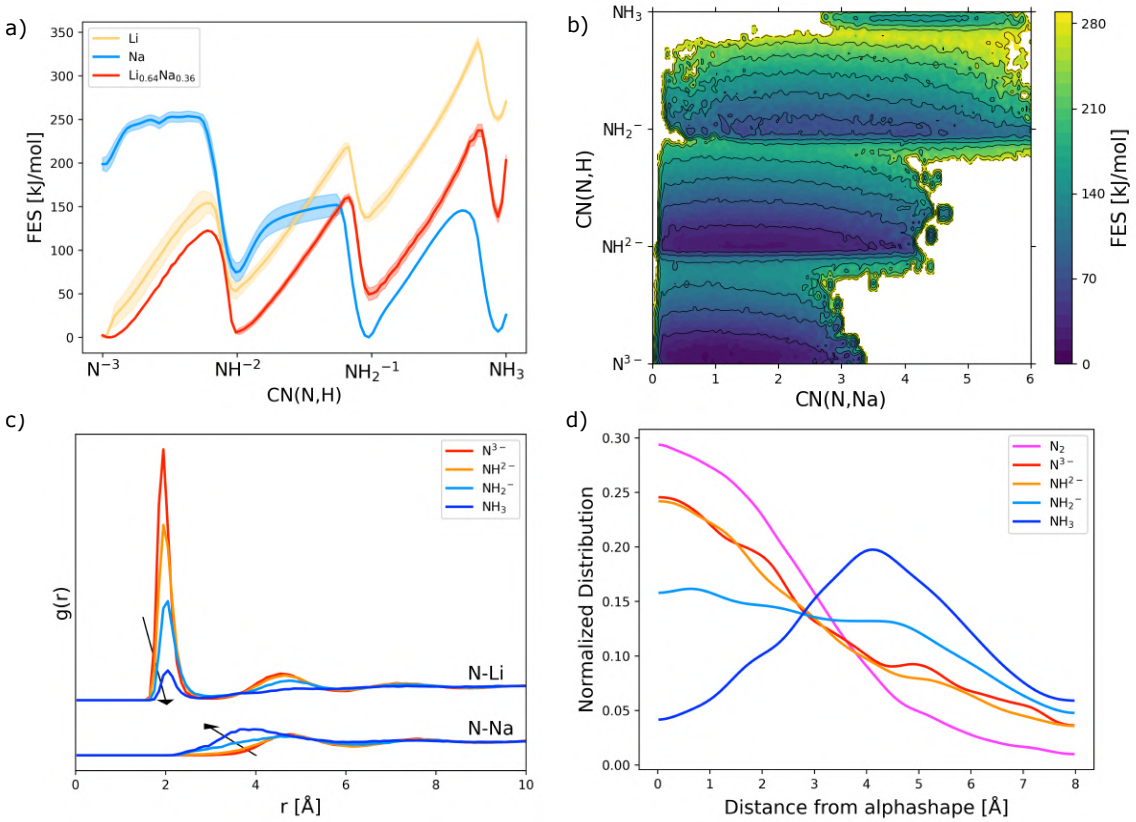

Figure 6.6: **Formation of NH_3**. a) Free energy profiles along the coordination number N−H for the formation of NH_3 in pure Li, pure Na and in the $Li_{0.64}Na_{0.36}$ alloy at 600K. b) 2D-Free energy surface of NH_3 formation in $Li_{0.64}Na_{0.36}$ as a function of the coordination number between N and H (CN(N,H)) and the coordination number between N and Na (CN(N,Na)). c) Radial distribution functions for N−Li and N−Na pairs in the $Li_{0.64}Na_{0.36}$ alloy. The black arrows indicate the effect of protonation on the first peak of the $g(r)$. d) Distributions of distances of N species from the Li−Na interface determined with the alpha-shape method [194, 195] as implemented in OVITO [186]. As the protonation state of the species increases, the probability of being close to interface decreases, while increases its probability to be found at larger distances. Details of the simulations are reported in the Methods section 6.4.

of the microdomains, thereby modulating the reaction free energy landscape. We hypothesize that increasing the lithium content could promote more efficient N_2 activation. On the other hand, increasing sodium content could favor the hydrogenation pathways.

By tuning the Li−Na ratio, one can effectively program the distribution and relative sizes of the microdomains, thereby influencing the reaction free energy landscape and the overall production rate. However, composition tuning must also be carefully optimized to avoid the formation of poisoning phases, such as LiH_3 or NaH, which can hinder long-term performance. Ongoing simulations are currently being conducted to identify the optimal compositional window and tempera-

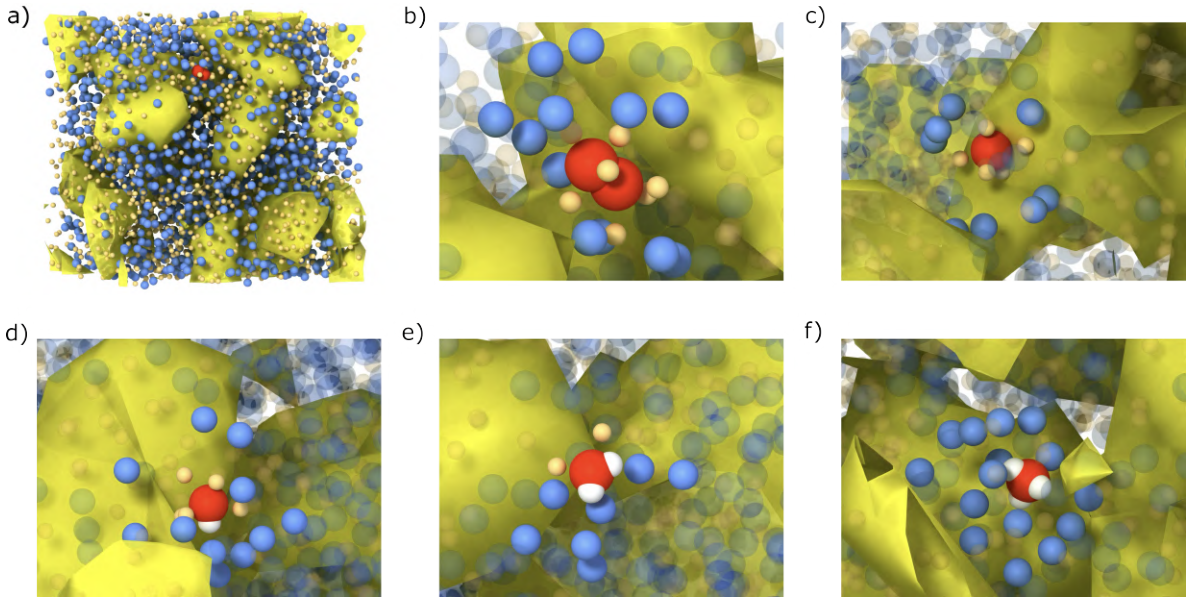

Figure 6.7: **Different steps of ammonia synthesis at 600K**. a) Snapshot of the Li-Na system with one N_2 molecule dissolved. b) View of the coordination of dissolved N_2. c) Dissolved N after the N—N splitting. d) Dissolved NH. e) Dissolved NH_2. f) Dissolved NH_3. Details of the simulations are reported in the Methods section 6.4.

ture that maximize catalytic activity while minimizing the formation of poisoning phases, thereby ensuring sustained performance under realistic conditions.
In conclusion, the $Li_{0.64}Na_{0.36}$ system provides a compelling example of how phase-separated, microheterogeneous environments can be leveraged for catalytic design. The interface itself becomes an active participant, not a passive spectator. This behavior challenges once more the conventional view of catalysis as a static phenomenon and points toward a new paradigm of interfacial catalysis. Future work will explore how these principles apply to other liquid alloys, and how the properties of the interface influence reactivity and selectivity. Such efforts could lay the groundwork for designing liquid catalytic materials that exploit heterogeneity not as a limitation, but as a central design strategy.

6.4 Methods

AIMD simulations

The initial data set was collected via ab initio molecular dynamics (AIMD) of the mixture alloy and pure liquid Li and Na. These simulations were performed using CP2K 9.1 [206, 207], patched with PLUMED 2.9 [127]. Initially, equilibration of the simulation cell volume was conducted by sampling the isothermal-isobaric ensemble at 1 atm in the temperature range 500-800K, employing a barostat with a coupling constant of 100 fs and a timestep of 0.5 fs. The temperature was maintained using a Canonical Sampling through Velocity Rescaling (CSVR) thermostat [83] with a coupling

constant of 10 fs. The total equilibration duration was set to 10 ps. After equilibration, AIMD simulations were performed in the NVT ensemble in the temperature range 500-800K using the same CSVR and timestep settings. The electronic structure was computed via the Kohn–Sham density functional theory (DFT) approach within the Quickstep implementation of CP2K, employing the Gaussian Plane Wave (GPW) method. The Perdew–Burke–Ernzerhof (PBE) [208] generalized gradient approximation (GGA) exchange-correlation functional was used. Gaussian-type DZVP-MOLOPT-PBE basis sets and GTH pseudopotentials [209, 210] were selected from the UZH library. The plane-wave cutoff energy for the density grid was set to 400 Ry, employing five multigrid levels. Enhanced sampling simulations for collecting reactive configurations for N_2 splitting and NH_3 formation in the three systems were performed using OPES [115] bias potential, as detailed in Methods section 6.4. From the AIMD trajectory, configurations were selected for DFT relabelling of energy and forces.

DFT calculations

Single-point DFT calculations were performed using Quantum ESPRESSO 7.1 [211, 212, 213]. The calculations employed the Perdew–Burke–Ernzerhof (PBE) generalized gradient approximation (GGA) exchange-correlation functional [208, 214]. Ultrasoft pseudopotentials (USPP) and Optimized Norm-Conserving Vanderbilt (ONCV) pseudopotentials from the SSSP PBE Precision v1.3.0 library were used for lithium and sodium, respectively. Plane-wave basis sets were used with kinetic energy cutoffs set to 75 Ry for wavefunctions and 300 Ry for charge densities. Gaussian smearing was applied with a smearing width of 0.01 Ry. Calculations were performed at the Gamma point only. The self-consistent field convergence was assisted by a mixing parameter of 0.1, with initial potentials set to atomic and wavefunctions initialized with atomic orbitals plus a random perturbation.

Machine learning potentials training

Three separate MLP models were trained: one for pure liquid Li, one for pure Na, and one for the $Li_{0.64}Na_{0.36}$ alloy (Table 6.2). Each model was trained using the Deep Potential Molecular Dynamics Smooth Edition scheme, as implemented in the DeePMD-kit software (version 2.2.7) [144]. An embedding and a fitting network with [30, 60, 120] and [240, 240, 240, 240] nodes per layer were used, respectively. A cut-off of 7Å was used to describe to local environment of each atom. The ADAM optimizer was used to train the network with an exponentially decaying learning rate from 1×10^{-3} to 5×10^{-8}. The pre-factors used in the in loss function for the energy and force terms are varied during the optimization process from 0.05 to 5 and from 1000 to 1, respectively. The batch size was set to 8 and the final models used for the production runs were trained for 2×10^6 steps.

	Training Set			Test Set		
System	Num. configs	Energy RMSE/Natoms	Force RMSE	Num. configs	Energy RMSE/Natoms	Force RMSE
$Li_{0.64}Na_{0.36}$ alloy						
$Li_{320}Na_{180}$	5000	1.4e-3	2.3e-2	500	1.3e-3	2.1e-2
$Li_{320}Na_{180}N_2$	6521	1.6e-2	2.5e-2	300	1.5e-3	2.6e-2
$Li_{320}Na_{180}N_2H_3$	5121	1.7e-3	2.7e-2	300	1.4e-3	2.5e-2
$Li_{320}Na_{180}H_2$	3000	2.0e-3	2.7e-2	300	2.1e-3	2.8e-2
Pure Li						
Li_{300}	500	1.0e-3	2.1e-2	100	1.1e-3	2.2e-2
$Li_{300}N_2$	1967	1.5e-3	2.2e-2	200	1.4e-2	2.3e-2
$Li_{300}N_2H_3$	2731	1.7e-3	2.5e-2	200	1.6e-3	2.5e-2
$Li_{300}H_2$	1832	1.2e-3	2.2e-2	200	1.3e-3	2.1e-2
Pure Na						
Na_{300}	1022	1.7e-3	3.0e-2	100	1.8e-3	2.9e-2
$Na_{300}N_2$	4905	2.0e-3	2.9e-2	200	2.1e-3	3.0e-2
$Na_{300}N_2H_3$	4501	2.3e-3	3.2e-2	200	2.5e-3	3.4e-2
$Na_{300}H_2$	3105	1.7e-3	2.8e-2	200	1.8e-3	2.9e-2

Table 6.2: **Machine learning potentials details**. Energy RMSE/Natoms and Forse RMSE are given in [eV/atom] and [eV/Å], respectively.

Molecular Dynamics simulations

MLP-MD simulations were performed with Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) software (version 24Mar2022) [216], patched with DeePMD-kit and PLUMED (version 2.9) [127]. NVT simulations were performed with an integration time step of 0.5 fs. The temperature was controlled using stochastic velocity scaling thermostat [83] with a coupling constant of 50 fs. MLP-MD simulations of bulk $Li_{0.64}Na_{0.36}$ were performed in the temperature range of 400-900K on a system containing 8000 atoms. The corresponding results are reported in Fig. 6.2. Simulations to study the H_2 and N_2 dynamics (Fig. 6.4), N_2 splitting (Fig. 6.5) and NH_3 formation (Fig. 6.6) were performed on systems containing 1002 atoms for pure Li and pure Na, and 4002 atoms for the $Li_{0.64}Na_{0.36}$ alloy. The alpha-shape analysis presented in Fig. 6.1, Fig. 6.6d, and Fig. 6.7 were carried out using the alpha-shape method [194, 195] as implemented in OVITO [186], with a probe radius of 5Å.

Enhanced sampling simulations

The PLUMED library [127] is used to calculate the collective variables and the OPES [115] bias potentials for both ab initio and deep potential molecular dynamics simulations. The N_2 splitting simulations were performed using the distance between the two N atoms as CV. For the synthesis of NH_3 in pure Li and pure Na the coordination number between the N atom and the three H atoms (CN(N,H)) was used as CV. In the case of $Li_{0.64}Na_{0.36}$ alloy, the coordination number between N atom and Li atoms (CN(N,Li)) was used in addition to CN(N,H). For the latter, reweighing along CN(N,H) and the coordination number between N atom and Na atoms (CN(N,Na)) were performed to give a clearer picture of the sodiophilicity trend of NH_x species.

System	Temperature	CV	Barrier	PACE
		N_2 splitting (Fig. 6.5)		
$Li_{0.64}Na_{0.36}$	600	$d(N, N)$	150	200
Li	600	$d(N, N)$	180	200
Na	600	$d(N, N)$	320	200
		NH_3 formation (Fig. 6.6)		
$Li_{0.64}Na_{0.36}$	600	CN(N,H) + CN(N,Li)	230 + 80	200 + 500
Li	600	CN(N,H)	230	200
Na	600	CN(N,H)	200	200

Table 6.3: **Enhanced sampling simulations details**.

Chapter 7

Ongoing works

In this part of the thesis, we briefly present preliminary results from ongoing projects that extend the insights gained from the previous chapters. In particular, we begin by investigating the dynamics of nickel nanoparticles supported by BaH_2, followed by the study of BaH_2 potential in olefin hydrogenation and carbon-carbon coupling reactions. We then examine the role of alkali cations in modulating interfacial reactivity during CO_2 electroreduction. Finally, we present a screening protocol for mixed hydride-nitride materials and highlight Sr_2LiH_2N as a promising candidate for ammonia synthesis. Together, these studies once again emphasize the central role of dynamics in understanding and designing efficient catalytic systems under realistic, *operando* conditions.

7.1 Enhancing BaH_2 activity with nickel nanoparticles

As anticipated in chapter 4, the combination of alkali and alkaline earth metal hydrides with transition metals can lead to an enhanced catalytic activity. This synergistic effect is evidenced by the significant boost in ammonia synthesis observed in Ru/CaH_2, $Ru/Ca_2N{:}e^-$ [204] and Ni-BaH_2 [205, 173] catalysts. In the latter case, nickel acts as a kinetic promoter in the chemical looping process, significantly accelerating both the N_2 fixation and hydrogenation steps while also lowering the corresponding activation energies. This is particularly noteworthy given that Ni is generally inactive in conventional thermocatalytic ammonia synthesis [173].

Preliminary MD simulations performed using the MACE-MP0 foundational model, revealed that when a Ni cluster is deposited on the BaH_2 surface, it undergoes significant structural and electronic rearrangements. Specifically, the cluster becomes partially covered by surface Ba and H atoms through a spillover mechanism [280] (Fig. 7.1). These atoms interact strongly with the Ni cluster with significant charge redistribution at the Ni-BaH_2 interface. In particular, from Bader charge calculations, we observed that Ba^{2+} cations tend to reduce to Ba^0, while H^- anions tend to oxidize to H^0. As a result, Ni atoms in the cluster acquire a negative charge that can act as a source of electrons in the N_2 adsorption process. Starting from this initial insights, we aim to expand the training set used in chapter 4 and simulate the N_2 activation on Ni-supported nanoparticles on BaH_2 (Fig. 7.1).

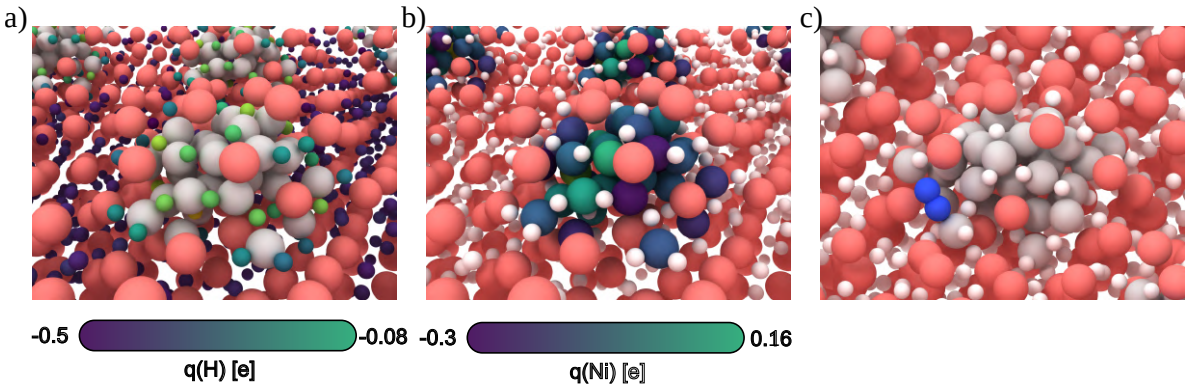

Figure 7.1: **Charges distribution and N_2 adsorption.** a) Distribution of H Bader charges. b) Distribution of Ni Bader charges. c) Representative snapshot for N_2 adsorption on Ni nanoparticle.

7.2 Olefine reduction and functionalization with BaH_2

In addition to its role in ammonia synthesis, BaH_2, along with CaH_2, has also been investigated for its catalytic activity in the hydrogenation of olefins and aromatic compounds [281, 282, 283, 284, 285]. Preliminary MD simulations reveal that olefins, such as 1,3-butadiene and acetylene, can interact with the surface of BaH_2 via π-cation interaction. This interaction facilitates the transfer of a surface hydride to the molecule, leading to the formation of a carbanion and a positively charged vacancy on the surface. Subsequent exposure to molecular hydrogen can result in hydrogenation of the carbanion, reducing its unsaturation and simultaneously regenerating the catalyst surface (Fig. 7.2). Furthermore, we propose that the carbanion intermediate formed

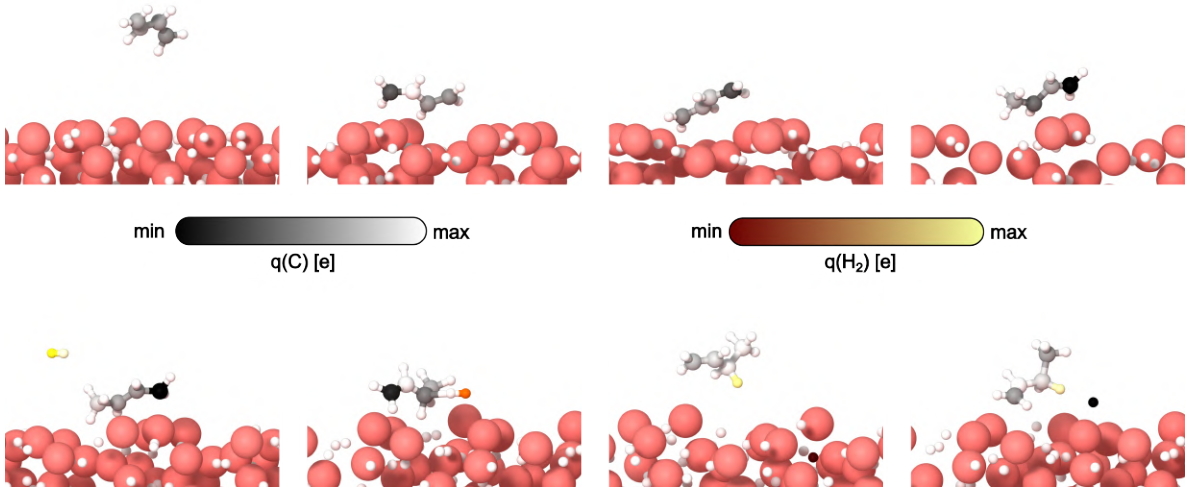

Figure 7.2: **Reduction of 1,3-dutadiene.** (Top) Reaction with one superficial hydride to form carbanion. (Bottom) Reaction with H_2 to form the monoinsaturated olefine. The C are colored according to their Bader charge. In the case of the reaction with H_2, H atom in H_2 are colored according to their Bader charges.

during hydride transfer could serve as a versatile synthon for other types of reactions, including C–C coupling processes, thereby expanding the catalytic utility of alkaline earth metal hydrides beyond hydrogenation reactions. As a proof of concept, we explored the nucleophilic addition of the butadienyl species to electrophilic carbons in CO_2 and oxirane (Fig. 7.3). In both cases, C–C bond formation is observed. These preliminary findings open the door to a broader investigation of alkali and alkaline earth metal hydrides as multifunctional catalysts for the synthesis of value-added chemicals. To accelerate this effort, we plan to use franken to finetune foundational models, which could enable screening of reaction pathways across a variety of substrates.

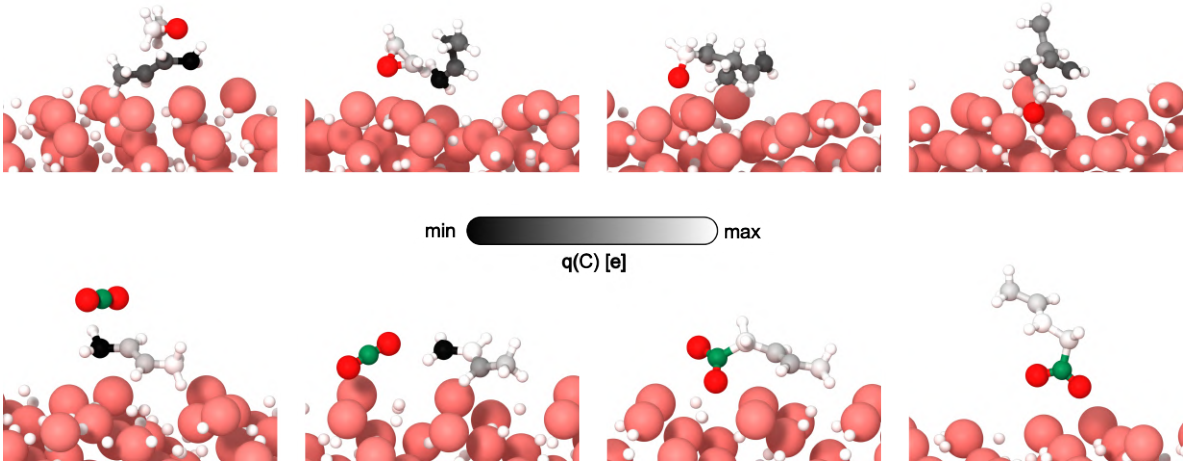

Figure 7.3: **Functionalization of 1,3-dutadiene**. (Top) Reaction with oxirane. (Bottom) Reaction with CO_2. The C are colored according to their Bader charge. In the case of CO_2, the C atom in CO_2 is omitted from the color coding to highlight changes in charge for hydrocarbon C.

7.3 Cations effect on electroreduction of CO_2

It is recognized that alkali metal cations play an active role in electrocatalytic reactions, not merely as spectators in the electrolyte but as entities that directly influence interfacial reactivity [286, 231]. In this context, cations should be viewed not simply as components of the bulk electrolyte, but as a functional element of the catalytic interface [287, 241]. For this reason, we aim to expand the results presented in chapter 5 by investigating the effect of K^+ on CO_2 electroreduction at the Ag(111) surface under constant-potential, operando conditions (-1.5V vs SHE). Preliminary results obtained from MLP-simulations at the r^{2}SCAN level of theory, confirm the strong influence of cations on CO_2 activation (Fig. 7.4). In particular, as the concentration of K^+ increases from 0 to 1/8 of a monolayer, the stabilization of adsorbed $*CO_2$ increases. Furthermore, if the concentration of of K^+ increases to 3/16 of a monolayer, we observed the outersphere activation of CO_2 (Fig. 7.4d). At this condition, the CO_2 molecule is not directly chemisorbed onto the metal surface, but rather stabilized and activated in the interfacial region by electrostatic interactions with nearby K^+ ions. Furthermore, the reaction selectivity shifts from CO production to formate formation [288].

On going simulations will allow us to quantify how ion-specific hydration structures, interfacial ion distributions, and dynamic solvent reorganizations modulate the free energy landscape for CO_2 activation. In particular, we will apply our novel enhanced sampling scheme based on the committor function to identify and characterize the transition-state ensembles associated with CO_2 activation, providing an atomistic understanding of how cations influence this process and rationalizing the shift in reaction selectivity with varying cation concentration.

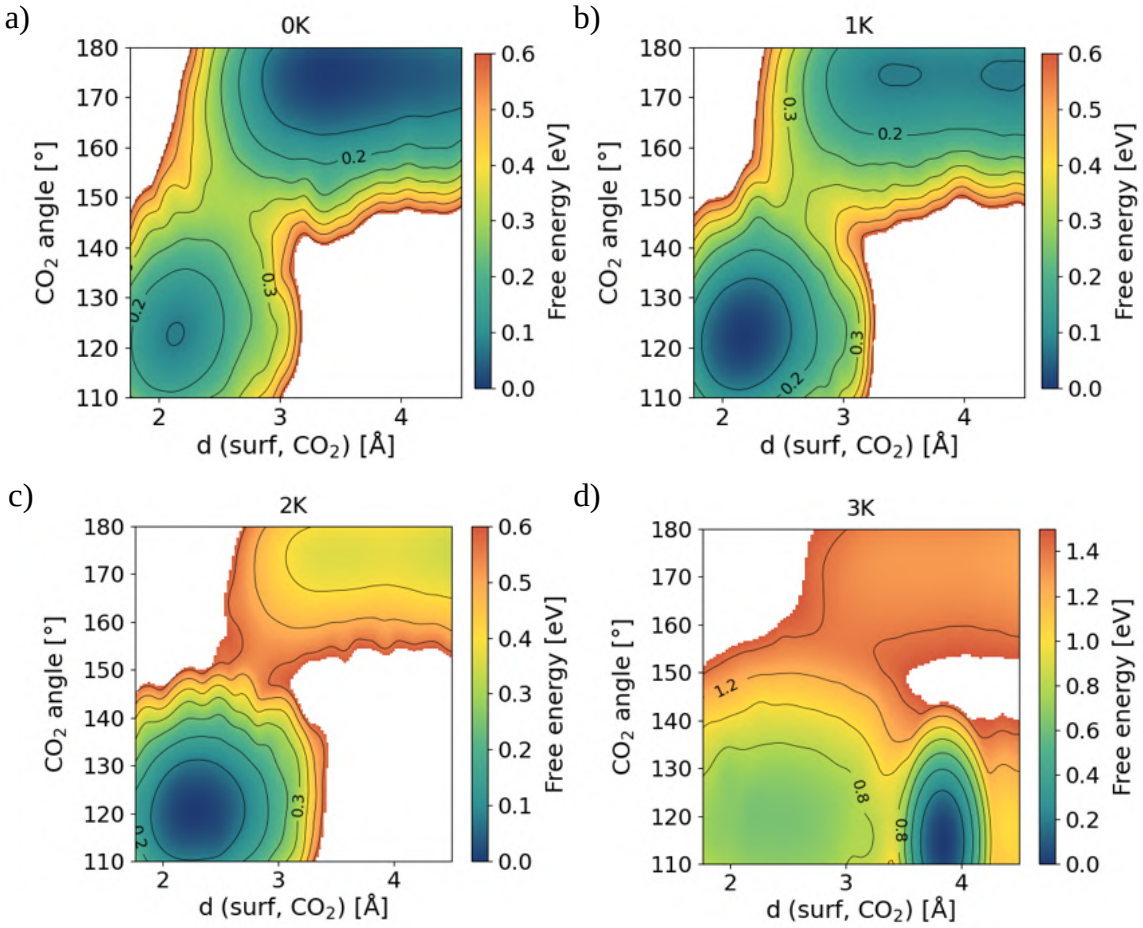

Figure 7.4: **K^+ concentration effect.** 2-D Free energy surfaces for the activation of CO_2 at different K^+ concentrations. In a) no K^+ is introduced in the simulation cell. In b) only one K^+ is introduced in the simulation cell, corresponding to 1/16 of a monolayer. In c), two K^+ are considered (1/8 of a monolayer), and in d) three K^+ are considered (3/16 of a monolayer).

7.4 Catalysts screening and Sr_2LiH_2N for ammonia synthesis

Building upon the insights gained from our previous studies on catalytic systems, we aim to establish a screening protocol to identify new candidate materials for ammonia synthesis or decomposition. Hydride and nitride compounds have attracted significant attention in the development

of efficient catalysts for ammonia synthesis, particularly under mild conditions [289]. Their relevance arises from several factors. Nitrides provide a source of lattice nitrogen that can directly participate in the reaction via a Mars-van Krevelen mechanism, enabling dynamic cycling between nitrogenated and de-nitrogenated states [290, 291]. For instance, nitrides such as LaN, Fe_4N, and Co_3Mo_3N [292, 293] have demonstrated catalytic activity, with lattice nitrogen being hydrogenated to form NH_3, while the generated vacancies serve as active sites for N_2 activation. Hydride compounds, on the other hand, can form anionic vacancies via reductive elimination and facilitate N_2 activation through reductive protonation reactions. Examples such as Li_4FeH_6 [294], Li_4RuH_6 [199], LiH, CaH_2, and BaH_2 have shown the ability to react with N_2 to form NH_x intermediates and regenerate in chemical looping cycles. Importantly, compounds that combine both hydride and nitride anions, such as Ba_5CrN_4H [295], and Ca_3CrN_3H [296] have been experimentally synthesized and tested, showing measurable catalytic activity in ammonia formation. These mixed-anion materials leverage the dual reactivity of both lattice nitrogen and hydride ions, creating a unique environment that promotes both N≡N activation and N−H bond formation.Furthermore, from our study on BaH_2, the high mobility of hydrides emerged as a key feature for its catalytic activity.
Therefore we explored untested materials with high hydride diffusion as potential catalyst for ammonia synthesis. Thanks to the availability of large-scale materials databases and foundational models, we can efficiently explore a vast chemical space to identify promising compounds. To this end, we propose a multi-stage filtering protocol (Fig. 7.5):

1 Selection of ternary, quaternary, and quinary nitride-hydride compounds.

2 Prioritization of materials with available experimental data.

3 Elimination of compounds containing catalytically irrelevant or unstable species.

4 Assessment of bulk-phase dynamical stability using molecular dynamics simulations with foundational interatomic potentials.

This protocol is designed to combine chemical intuition with data-driven selection criteria and atomistic simulations, ultimately accelerating the discovery of viable catalysts. We applied this screening protocol to Materials Project dataset [297, 298] and identified five promising candidates: Ca_2BHN_2, Ca_2HN, Li_4HN, Sr_2HN, Sr_2LiH_2N. The nitridoborate compound Ca_2BHN_2 was excluded from the screening because it is a molecular solid. Among the four remaining compounds, Sr_2LiH_2N exhibited the highest hydride diffusion coefficient (Fig. 7.5), making it a particularly attractive material for further investigation in the context of ammonia synthesis.
As a selected candidate, we trained a preliminary MLP model at the PBE level of theory and started investigating its catalytic properties. Similar to the approach followed for BaH_2, we began by analyzing the diffusion behavior in the bulk phase. Sr_2LiH_2N crystallizes in an orthorhombic structure (space group *Pnma*, Nr. 62), characterized by a layered arrangement in which lithium planes are "'sandwiched" between planes composed of strontium cations. The computed diffusion coefficients for H^- are particularly high, ranging from 1.84×10^{-3} Å^2/ps at 300K to 1.45×10^{-1} Å^2/ps at 700K (Fig. 7.6a). Hydrides occupy positions in both Li and Sr planes, while nitrides are located in the Sr planes. This structural difference leads to distinct dynamic behaviors in hydrides: H^- in the Li planes exhibits in-plane mobility, while H^- in the Sr planes tends to diffuse vertically (Fig.

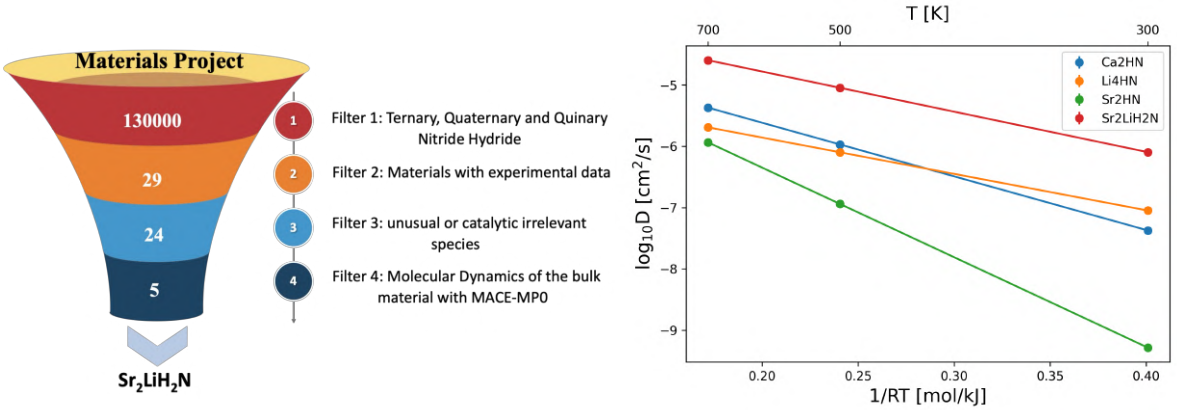

Figure 7.5: **Screening protocol and diffusion coefficients calculation.** Schematic representation of our screening protocol consisting of four filters. Starting from 130000 initial structures of nitride-hydride compounds, the screening protocol ends with selecting just five candidates. The dynamic properties of these candidates are then evaluated using the MACE-MP0 foundational interatomic potential. On the right, the diffusion coefficient for four compounds at temperatures 300K, 500K and 700K.

7.6b). The coexistence of these distinct modes enables hydride diffusion along specific channels within the crystal (Fig. 7.6c). To investigate the catalytic activity, we selected the (001) surface as a

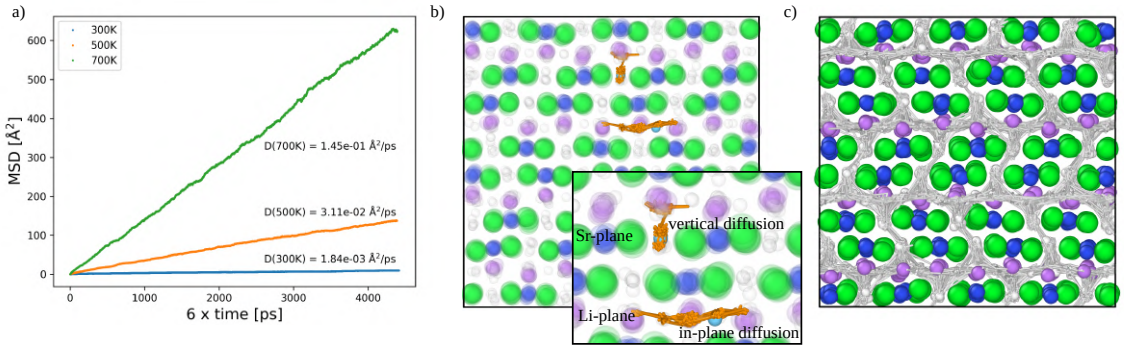

Figure 7.6: **Diffusion in bulk Sr_2LiH_2N.** a) Mean square displacements and diffusion coefficients at 300K, 500K and 700K. b) Different diffusion modes for H^- in Li-planes and Sr-planes. c) Positions of H^- during a trajectory of 500ps at 500K. Sr, Li, N, and H atoms are represented with green, violet, blue, and white spheres, respectively.

starting model interface (Fig. 7.7a). This surface exposes nitrides, which could actively participate in ammonia synthesis via a Mars-van Krevelen-type mechanism. We propose a chemical looping process, in which the material is alternately exposed to H_2 and N_2 atmospheres. Our simulations reveal that, under H_2, the exposed surface nitrides are reduced to form ammonia (Fig. 7.7b-e). For each proton added to nitrogen, a hydride ion is inserted into the lattice; therefore, the formation of one NH_3 molecule corresponds to the incorporation of three lattice hydrides. Remarkably,

from our simulations the crystal can accommodate a large number of additional hydrides, possibly leading to the formation of an intermediate compound through a phase transformation.
In the subsequent N_2 exposure phase, we hypothesize that dehydrogenation of the hydride-rich intermediate may occur, resulting in the formation of anionic vacancies or F-centers. These vacancies can adsorb and activate N_2, leading to the reformation of surface nitrides and completing the catalytic cycle [196, 28]. Ongoing simulations aim to elucidate the detailed mechanism underlying this process. Notably, preliminary experimental studies have demonstrated that Sr_2LiH_2N functions as a promising catalyst for ammonia synthesis. Combined computational and experimental efforts are currently underway to rationalize its catalytic mechanism. This positively reinforce the validity of our screening protocol. Looking ahead, integrating dynamic, *operando*-level descriptors will be key to improve the screening process.

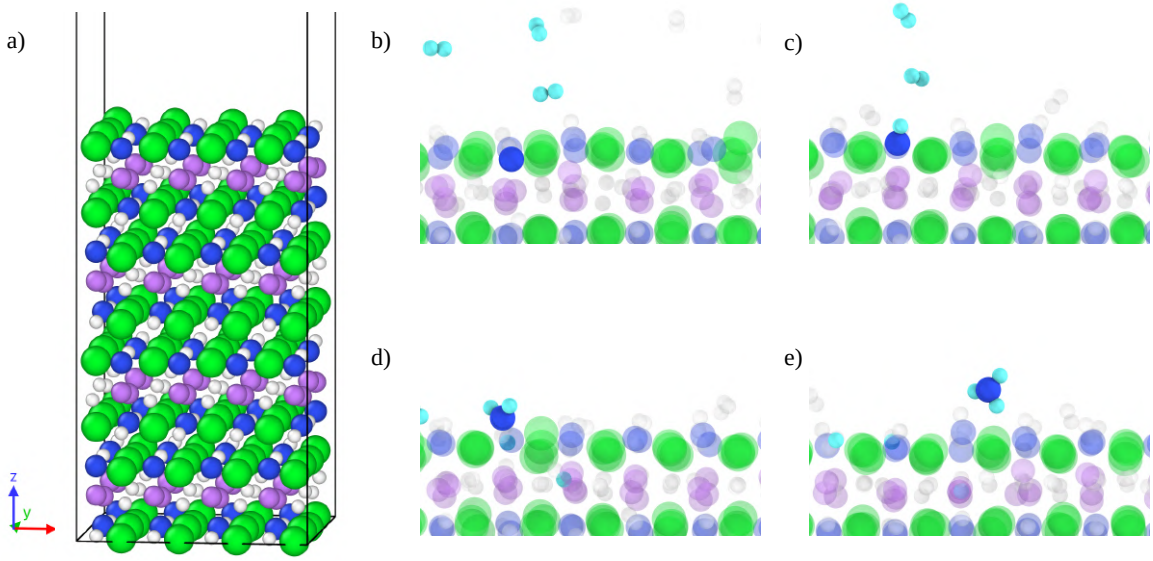

Figure 7.7: **(001) surface and hydrogenation.** a) Structure of the (001) surface. Representative snapshots for MLP-MD simulation of the hydrogenation of the (001) surface to produce ammonia: a) initial frame with superficial nitrides; b) formation of imide; c) formation of amide; e) formation of ammonia. Sr, Li, N, and H atoms are represented with green, violet, blue, and white spheres, respectively. Reacting H_2 molecules are highlighted in cyan.

Conclusions

The impact of machine learning on atomistic simulations is deepening our understanding of catalysis and reinforcing the perspective that the catalyst and the *operando* conditions cannot be treated as separate entities. Instead, catalytic activity emerges from the complex, dynamic interplay between the material and its environment.

In this thesis, we have shown that incorporating dynamics into computational models and properly treating *operando* variables, offers a more accurate view of catalysis. These approaches allow us to move beyond the traditional concept of static active sites and toward a framework where reactivity is governed by collective, time-dependent system's behavior.

In the case of barium hydride, we demonstrated that the dynamic transformations of the catalyst under *operando* conditions are central to its catalytic activity in ammonia synthesis. At elevated temperatures, BaH_2 undergoes reversible structural changes that enable the formation of an active phase. This behavior is driven by the high mobility of hydride ions, which facilitates key steps in the catalytic process. These findings underscore the importance of considering phase transformations and ion mobility at *operando* conditions for understanding the underlying mechanism of catalysis.

For CO_2 electroreduction, we investigated the role of interfacial water and the effect of the applied potential on the structure and dynamics of the electrode–electrolyte interface. Our results reveal that the system exhibits markedly different dynamical behavior when transitioning from the potential of zero charge (PZC) to constant-potential (CP) conditions. This shift influences the orientation and mobility of water molecules, which in turn affects the stabilization of key reaction intermediates.

Finally, inspired by our insights into the importance of dynamics, we proposed a novel catalyst for ammonia synthesis based on the liquid Li−Na alloy. In this system, the intrinsic microheterogeneity and dynamical fluctuations of the Li-Na interface create a highly adaptable environment for N_2 activation and hydrogenation to NH_3. Also, we are investigating the ability of other hydrides, such as Sr_2LiH_2N for NH_3 synthesis or BaH_2 for olefine selective reduction or functionalization. These examples highlights how embracing dynamic can lead to unconventional yet highly promising catalytic designs.

Overall, our findings highlight the necessity of embracing a dynamic and collective view of catalysis. Recognizing that the entire catalytic environment participates in the reaction enables the design of more efficient, robust, and adaptable materials. Despite the progress made in these studies, several avenues remain open for future research. In particular, the rational design of catalysts using dynamic descriptors that capture the evolving nature of atomic environments,

represents a promising direction. Such descriptors can provide a more accurate and generalizable basis for understanding and optimizing catalytic performance under realistic conditions.
As computational chemistry continues to evolve, it is no longer a supporting tool in catalysis, but a central driver of discovery, pushing the boundaries of what we can simulate, understand, and design. Machine learning, in this context, should not be seen as an endpoint but as a powerful tool, one that complements, rather than replaces, physical insight, human intuition, and rigorous theoretical understanding. Machine learning continues to hold great potential for advancing the study of catalytic processes. Future efforts should focus on developing more sophisticated algorithms capable of autonomously exploring chemical space, identifying reaction mechanisms, and optimizing catalyst structures in a data-driven yet physically grounded manner.
Yet, with each advancement, the frontier of knowledge keeps shifting. The task of the scientific community, of which I am a small part, remains clear: to keep pushing boundaries, daring, exploring, and innovating. The challenge never ends.

Publications

Articles published in peer-reviewed journals

1 Pengchao Zhang, Axel Tosello Gardini, Xuefei Xu, and Michele Parrinello. Intramolecular and water-mediated tautomerism of solvated glycine. *Journal of Chemical Information and Modeling*, 64(8):3599–3604, 2024.

2 Axel Tosello Gardini, Umberto Raucci, and Michele Parrinello. Machine learning-driven molecular dynamics unveils a bulk phase transformation driving ammonia synthesis on barium hydride. *Nature Communications*, 16(1):2475, 2025.

3 Xiongwei Tian, Axel Tosello Gardini, Umberto Raucci, Hai Xiao, Yuqun Zhuo, and Michele Parrinello. Electrochemical potential-driven water dynamics control CO_2 electroreduction at the Ag/H_2O interface. Nature Communications, 16(1):10636, 2025.

Articles in preparation

1 Axel Tosello Gardini, Umberto Raucci, Michele Parrinello. Ammonia synthesis catalyzed by liquid lithium-sodium alloys. In preparation (2026).

2 Axel Tosello Gardini, Umberto Raucci, Michele Parrinello. A Unifying Timescale Language for Heterogeneous Catalysis. In preparation (2026).

Patents

1 Axel Tosello Gardini, Umberto Raucci, Michele Parrinello, 2025, Metodo per la sintesi catalitica di ammoniaca. IT 102025000020656, filed August 05, 2025. Patent Pending.

Bibliography

[1] Jöns Jacob Berzelius. *Årsberättelse om framstegen i fysik och kemi: 1835*. Norstedt, 1835.

[2] Wilhelm Ostwald. Definition der katalyse. *Zeitschrift für physikalische Chemie*, 15:705–706, 1894.

[3] Fanran Meng, Andreas Wagner, Alexandre B Kremer, Daisuke Kanazawa, Jane J Leung, Peter Goult, Min Guan, Sophie Herrmann, Eveline Speelman, Pim Sauter, et al. Planet-compatible pathways for transitioning the chemical industry. *Proceedings of the National Academy of Sciences*, 120(8):e2218294120, 2023.

[4] C Richard Catlow, Matthew Davidson, Christopher Hardacre, and Graham J Hutchings. Catalysis making the world a better place, 2016.

[5] Irving Langmuir. The adsorption of gases on plane surfaces of glass, mica and platinum. *Journal of the American Chemical society*, 40(9):1361–1403, 1918.

[6] Hugh Stott Taylor. A theory of the catalytic surface. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 108(745):105–111, 1925.

[7] Bert M Weckhuysen. Chemical imaging of spatial heterogeneities in catalytic solids at different length and time scales. *Angewandte Chemie International Edition*, 48(27):4910–4943, 2009.

[8] MS Spencer. Stable and metastable metal surfaces in heterogeneous catalysis. *Nature*, 323(6090):685–687, 1986.

[9] Robert Schlögl. Heterogeneous catalysis. *Angewandte Chemie International Edition*, 54(11):3465–3520, 2015.

[10] Bert M Weckhuysen. Determining the active site in a catalytic process: Operando spectroscopy is more than a buzzword. *Physical Chemistry Chemical Physics*, 5(20):4351–4360, 2003.

[11] Pieter Van der Avert and Bert M Weckhuysen. Low-temperature destruction of chlorinated hydrocarbons over lanthanide oxide based catalysts. *Angewandte Chemie*, 114(24):4924–4926, 2002.

[12] Bert M Weckhuysen and Robert A Schoonheydt. Alkane dehydrogenation over supported chromium oxide catalysts. *Catalysis today*, 51(2):223–232, 1999.

[13] Bert M Weckhuysen, Israel E Wachs, and Robert A Schoonheydt. Surface chemistry and spectroscopy of chromium in inorganic oxides. *Chemical Reviews*, 96(8):3327–3350, 1996.

[14] Bert M Weckhuysen. Studying birth, life and death of catalytic solids with operando spectroscopy. *National Science Review*, 2(2):147–149, 2015.

[15] Frank C Hendriks, Florian Meirer, Alexey V Kubarev, Zoran Ristanovic, Maarten BJ Roeffaers, Eelco TC Vogt, Pieter CA Bruijnincx, and Bert M Weckhuysen. Single-molecule fluorescence microscopy reveals local diffusion coefficients in the pore network of an individual catalyst particle. *Journal of the American Chemical Society*, 139(39):13632–13635, 2017.

[16] Florian Meirer and Bert M Weckhuysen. Spatial and temporal exploration of heterogeneous catalysts with synchrotron radiation. *Nature Reviews Materials*, 3(9):324–340, 2018.

[17] Joel E Schmidt, Ramon Oord, Wei Guo, Jonathan D Poplawsky, and Bert M Weckhuysen. Nanoscale tomography reveals the deactivation of automotive copper-exchanged zeolite catalysts. *Nature communications*, 8(1):1666, 2017.

[18] Michel Boudart. Turnover rates in heterogeneous catalysis. *Chemical reviews*, 95(3):661–666, 1995.

[19] Charlotte Vogt, Esther Groeneveld, Gerda Kamsma, Maarten Nachtegaal, Li Lu, Christopher J Kiely, Peter H Berben, Florian Meirer, and Bert M Weckhuysen. Unravelling structure sensitivity in co2 hydrogenation over nickel. *Nature Catalysis*, 1(2):127–134, 2018.

[20] Walter Thiel. Computational catalysis—past, present, and future. *Angewandte Chemie International Edition*, 53(33):8605–8613, 2014.

[21] Joachim Sauer. The future of computational catalysis. *Journal of Catalysis*, 433:115482, 2024.

[22] Greg Collinge, Simuck F Yuk, Manh-Thuong Nguyen, Mal-Soon Lee, Vassiliki-Alexandra Glezakou, and Roger Rousseau. Effect of collective dynamics and anharmonicity on entropy in heterogenous catalysis: Building the case for advanced molecular simulations. *ACS Catalysis*, 10(16):9236–9260, 2020.

[23] GiovanniMaria Piccini, Mal-Soon Lee, Simuck F Yuk, Difan Zhang, Greg Collinge, Loukas Kollias, Manh-Thuong Nguyen, Vassiliki-Alexandra Glezakou, and Roger Rousseau. Ab initio molecular dynamics with enhanced sampling in heterogeneous catalysis. *Catalysis Science & Technology*, 12(1):12–37, 2022.

[24] Yang-Gang Wang, Donghai Mei, Vassiliki-Alexandra Glezakou, Jun Li, and Roger Rousseau. Dynamic formation of single-atom catalytic active sites on ceria-supported gold nanoparticles. *Nature communications*, 6(1):6511, 2015.

[25] Zachary K Goldsmith, Marcos F Calegari Andrade, and Annabella Selloni. Effects of applied voltage on water at a gold electrode interface from ab initio molecular dynamics. *Chemical Science*, 12(16):5865–5873, 2021.

[26] Emanuele Grifoni, GiovanniMaria Piccini, Johannes A Lercher, Vassiliki-Alexandra Glezakou, Roger Rousseau, and Michele Parrinello. Confinement effects and acid strength in zeolites. *Nature communications*, 12(1):2630, 2021.

[27] Veronique Van Speybroeck, Massimo Bocus, Pieter Cnudde, and Louis Vanduyfhuys. Operando modeling of zeolite-catalyzed reactions using first-principles molecular dynamics simulations. *ACS catalysis*, 13(17):11455–11493, 2023.

[28] Axel Tosello Gardini, Umberto Raucci, and Michele Parrinello. Machine learning-driven molecular dynamics unveils a bulk phase transformation driving ammonia synthesis on barium hydride. *Nature Communications*, 16(1):2475, 2025.

[29] Xiongwei Tian, Axel Tosello Gardini, Umberto Raucci, Hai Xiao, Yuqun Zhuo, and Michele Parrinello. Electrochemical potential-driven water dynamics control co2 electroreduction at the ag/h2o interface. *Nature Communications*, 16(1):10636, 2025.

[30] Axel Tosello Gardini, Umberto Raucci, and Michele Parrinello. Metodo per la sintesi catalitica di ammoniaca, August 2025. IT102025000020656.

[31] Collin Smith, Alfred K Hill, and Laura Torrente-Murciano. Current and future role of haber–bosch ammonia in a carbon-free energy landscape. *Energy & environmental science*, 13(2):331–344, 2020.

[32] John Humphreys, Rong Lan, and Shanwen Tao. Development and recent progress on ammonia synthesis catalysts for haber-bosch process. *Advanced Energy and Sustainability Research*, 2(1):2000043, 2021.

[33] M-C Tsai, U Ship, IC Bassignana, J Küppers, and G Ertl. A vibrational spectroscopy study on the interaction of n2 with clean and k-promoted fe (111) surfaces: π-bonded dinitrogen as precursor for dissociation. *Surface science*, 155(2-3):387–399, 1985.

[34] M Grunze, M Golze, W Hirschwald, H-J Freund, H Pulm, U Seip, MC Tsai, G Ertl, and J Küppers. π-bonded n 2 on fe (111): the precursor for dissociation. *Physical review letters*, 53(8):850, 1984.

[35] H-J Freund, B Bartos, RP Messmer, H Grunze, H Kuhlenbeck, and M Neumann. The adsorption of n2 on fe (111): Angle resolved photoemission and theoretical model studies. *Surface Science*, 185(1-2):187–202, 1987.

[36] H Itoh, G Ertl, and AB Kunz. A molecular orbital study on the interaction of dinitrogen with transition metal atoms. *Chemical Physics*, 59(1-2):149–156, 1981.

[37] Jens Jørgen Mortensen, Lars Bruno Hansen, Bjørk Hammer, and Jens Kehlet Nørskov. Nitrogen adsorption and dissociation on fe (111). *Journal of Catalysis*, 182(2):479–488, 1999.

[38] Luigi Bonati, Daniela Polino, Cristina Pizzolitto, Pierdomenico Biasi, Rene Eckert, Stephan Reitmeier, Robert Schlögl, and Michele Parrinello. The role of dynamics in heterogeneous catalysis: Surface diffusivity and n2 decomposition on fe (111). *Proceedings of the National Academy of Sciences*, 120(50):e2313023120, 2023.

[39] Shivam Tripathi, Luigi Bonati, Simone Perego, and Michele Parrinello. How poisoning is avoided in a step of relevance to the haber–bosch catalysis. *ACS Catalysis*, 14(7):4944–4950, 2024.

[40] Douglas R MacFarlane, Pavel V Cherepanov, Jaecheol Choi, Bryan HR Suryanto, Rebecca Y Hodgetts, Jacinta M Bakker, Federico M Ferrero Vallana, and Alexandr N Simonov. A roadmap to the ammonia economy. *Joule*, 4(6):1186–1205, 2020.

[41] Natalia Morlanés, Sai P Katikaneni, Stephen N Paglieri, Aadesh Harale, Bandar Solami, S Mani Sarathy, and Jorge Gascon. A technological roadmap to the ammonia energy economy: Current state and missing technologies. *Chemical Engineering Journal*, 408:127310, 2021.

[42] Chidozie Eluwah and Paul S Fennell. Novel onboard ammonia cracker for light-duty automotive fuel cell vehicles. *Energy Advances*, 2025.

[43] Simone Perego, Luigi Bonati, Shivam Tripathi, and Michele Parrinello. How dynamics changes ammonia cracking on iron surfaces. *ACS Catalysis*, 14(19):14652–14664, 2024.

[44] M Purcel, S Berendts, L Bonati, S Perego, A Muller, M Lerch, M Parrinello, and M Muhler. Iron nitride formation and decomposition during ammonia decomposition over a wustite-based bulk iron catalyst. *ACS Catalysis*, 14(18):13947–13957, 2024.

[45] Simone Perego, Maximilian Purcel, Yannick Baum, Shilong Chen, Astrid Sophie Muller, Michele Parrinello, Malte Behrens, Martin Muhler, and Luigi Bonati. No time for nitrides: How cobalt alloying promotes iron catalysts for ammonia decomposition. *ACS Catalysis*, 15:16690–16702, 2025.

[46] Simone Perego and Luigi Bonati. Data efficient machine learning potentials for modeling catalytic reactivity via active learning and enhanced sampling. *npj Computational Materials*, 10(1):291, 2024.

[47] Manyi Yang, Umberto Raucci, and Michele Parrinello. Reactant-induced dynamics of lithium imide surfaces during the ammonia decomposition process. *Nature Catalysis*, 6(9):829–836, 2023.

[48] Francesco Mambretti, Umberto Raucci, Manyi Yang, and Michele Parrinello. How does structural disorder impact heterogeneous catalysts? the case of ammonia decomposition on non-stoichiometric lithium imide. *ACS Catalysis*, 14(3):1252–1256, 2024.

[49] Francesco Mambretti, Umberto Raucci, Oscar Gómez-Cápiro, Mirabbos Hojamberdiev, Stefan Berendts, Holger Ruland, Martin Lerch, and Michele Parrinello. Theory meets experiment in

ammonia decomposition on li14cr2n8o: From order to disorder under reaction conditions. 2025.

[50] Axel Groß. Challenges for ab initio molecular dynamics simulations of electrochemical interfaces. *Current Opinion in Electrochemistry*, 40:101345, 2023.

[51] Hendrik H Heenen, Joseph A Gauthier, Henrik H Kristoffersen, Thomas Ludwig, and Karen Chan. Solvation at metal/water interfaces: An ab initio molecular dynamics benchmark of common computational approaches. *The Journal of Chemical Physics*, 152(14), 2020.

[52] Jia-Bo Le, Xiao-Hui Yang, Yong-Bin Zhuang, Mei Jia, and Jun Cheng. Recent progress toward ab initio modeling of electrocatalysis. *The Journal of Physical Chemistry Letters*, 12(37):8924–8931, 2021.

[53] Matteo Farnesi Camellone and Dominik Marx. On the impact of solvation on a au/tio2 nanocatalyst in contact with water. *The journal of physical chemistry letters*, 4(3):514–518, 2013.

[54] Gabriele Tocci and Angelos Michaelides. Solvent-induced proton hopping at a water–oxide interface. *The journal of physical chemistry letters*, 5(3):474–480, 2014.

[55] Jens Kehlet Nørskov, Jan Rossmeisl, Ashildur Logadottir, LRKJ Lindqvist, John R Kitchin, Thomas Bligaard, and Hannes Jonsson. Origin of the overpotential for oxygen reduction at a fuel-cell cathode. *The Journal of Physical Chemistry B*, 108(46):17886–17892, 2004.

[56] Nawras Abidi, Kang Rui Garrick Lim, Zhi Wei Seh, and Stephan N Steinmann. Atomistic modeling of electrocatalysis: Are we there yet? *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 11(3):e1499, 2021.

[57] Antton Curutchet, Pauline Colinet, Carine Michel, Stephan N Steinmann, and Tangui Le Bahers. Two-sites are better than one: revisiting the oer mechanism on coooh by dft with electrode polarization. *Physical Chemistry Chemical Physics*, 22(13):7031–7038, 2020.

[58] Stephan N Steinmann, Carine Michel, Renate Schwiedernoch, Jean-Sebastien Filhol, and Philippe Sautet. Modeling the hcooh/co2 electrocatalytic reaction: When details are key. *ChemPhysChem*, 16(11):2307–2311, 2015.

[59] Christopher D Taylor, Sally A Wasileski, Jean-Sebastien Filhol, and Matthew Neurock. First principles reaction modeling of the electrochemical interface: Consideration and calculation of a tunable surface potential from atomic and electronic structure. *Physical Review B—Condensed Matter and Materials Physics*, 73(16):165402, 2006.

[60] Duy Le. An explicit-implicit hybrid solvent model for grand canonical simulations of the electrochemical environment. *Chemrxiv*, 2023.

[61] Reihaneh Amirbeigiarab, Jing Tian, Antonia Herzog, Canrong Qiu, Arno Bergmann, Beatriz Roldan Cuenya, and Olaf M Magnussen. Atomic-scale surface restructuring of copper electrodes under co2 electroreduction conditions. *Nature Catalysis*, 6(9):837–846, 2023.

[62] Chao Zhan, Federico Dattila, Clara Rettenmaier, Antonia Herzog, Matias Herran, Timon Wagner, Fabian Scholten, Arno Bergmann, Núria López, and Beatriz Roldan Cuenya. Key intermediates and cu active sites for co2 electroreduction to ethylene and ethanol. *Nature Energy*, 9(12):1485–1496, 2024.

[63] Shuoqi Zhang, Qingli Tang, Beien Zhu, and Yi Gao. Reaction-driven varieties of active sites on cu (100) in electrochemical co2 reduction reaction. *ACS Catalysis*, 15(8):6497–6506, 2025.

[64] Zan Lian, Federico Dattila, and Núria López. Stability and lifetime of diffusion-trapped oxygen in oxide-derived copper co2 reduction electrocatalysts. *Nature Catalysis*, 7(4):401–411, 2024.

[65] Apinya Ngoipala, Christian Schott, Valentin Briega-Martos, Minaam Qamar, Matous Mrovec, Sousa Javan Nikkhah, Thorsten O Schmidt, Lewin Deville, Andrea Capogrosso, Lilian Moumaneix, et al. Hydride-induced reconstruction of pd electrode surfaces: A combined computational and experimental study. *Advanced Materials*, 37(4):2410951, 2025.

[66] Kourosh Kalantar-Zadeh, Torben Daeneke, and Junma Tang. The atomic intelligence of liquid metals. *Science*, 385(6707):372–373, 2024.

[67] Torben Daeneke, Khashayar Khoshmanesh, Nasir Mahmood, Isabela Alves De Castro, Dorna Esrafilzadeh, Steven J Barrow, Michael D Dickey, and Kourosh Kalantar-Zadeh. Liquid metals: fundamentals and applications in chemistry. *Chemical Society Reviews*, 47(11):4073–4111, 2018.

[68] Kuikui Zhang, Zeai Huang, Mingkai Yang, Mengying Liu, Yunxiao Zhou, Junjie Zhan, and Ying Zhou. Recent progress in melt pyrolysis: Fabrication and applications of high-value carbon materials from abundant sources. *SusMat*, 3(5):558–580, 2023.

[69] Mark McConnachie, Muxina Konarova, and Simon Smart. Literature review of the catalytic pyrolysis of methane for hydrogen and carbon production. *International Journal of Hydrogen Energy*, 48(66):25660–25682, 2023.

[70] Karma Zuraiqi, Yichao Jin, Caiden J Parker, Jaydon Meilak, Nastaran Meftahi, Andrew J Christofferson, Salvy P Russo, Michelle JS Spencer, Huai Yong Zhu, Lizhuo Wang, et al. Unveiling metal mobility in a liquid cu–ga catalyst for ammonia synthesis. *Nature Catalysis*, 7(9):1044–1052, 2024.

[71] Yan Gong, Da Luo, Myeonggi Choe, Yongchul Kim, Babu Ram, Mohammad Zafari, Won Kyung Seong, Pavel Bakharev, Meihui Wang, In Kee Park, et al. Growth of diamond in liquid metal at 1 atm pressure. *Nature*, 629(8011):348–354, 2024.

[72] Karma Zuraiqi, Ali Zavabeti, Francois-Marie Allioux, Jianbo Tang, Chung Kim Nguyen, Parisa Tafazolymotie, Mohannad Mayyas, Aswin V Ramarao, Michelle Spencer, Kalpit Shah, et al. Liquid metals in catalysis for energy applications. *Joule*, 4(11):2290–2321, 2020.

[73] Stephan Handschuh-Wang, Tao Wang, Tomasz Gancarz, Xiaorui Liu, Ben Wang, Bin He, Michael D Dickey, Georg W Wimmer, and Florian J Stadler. The liquid metal age: a transition from hg to ga. *Advanced Materials*, 36(45):2408466, 2024.

[74] Junma Tang, Andrew J Christofferson, Jing Sun, Qingfeng Zhai, Priyank V Kumar, Jodie A Yuwono, Mohammad Tajik, Nastaran Meftahi, Jianbo Tang, Liming Dai, et al. Dynamic configurations of metallic atoms in the liquid state for selective propylene synthesis. *Nature nanotechnology*, 19(3):306–310, 2024.

[75] Junma Tang, Nastaran Meftahi, Andrew J Christofferson, Jing Sun, Ruohan Yu, Md Arifur Rahim, Jianbo Tang, Guangzhao Mao, Torben Daeneke, Richard B Kaner, et al. Molten sn solvent expands liquid metal catalysis. *Nature Communications*, 16(1):907, 2025.

[76] Fumio Kawamura and Takashi Taniguchi. Synthesis of ammonia using sodium melt. *Scientific Reports*, 7(1):11578, 2017.

[77] Zujian Tang, Xian Meng, Yue Shi, and Xiaofei Guan. Lithium-based loop for ambient-pressure ammonia synthesis in a liquid alloy-salt catalytic system. *ChemSusChem*, 14(21):4697–4707, 2021.

[78] Daan Frenkel and Berend Smit. *Understanding molecular simulation: from algorithms to applications*. Elsevier, 2023.

[79] Mark E Tuckerman. *Statistical mechanics: theory and molecular simulation*. Oxford university press, 2023.

[80] Godehard Sutmann. Classical molecular dynamics, 2002.

[81] William C Swope, Hans C Andersen, Peter H Berens, and Kent R Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of chemical physics*, 76(1):637–649, 1982.

[82] Shuichi Nosé. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of chemical physics*, 81(1):511–519, 1984.

[83] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. *The Journal of chemical physics*, 126(1), 2007.

[84] Hans C Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of chemical physics*, 72(4):2384–2393, 1980.

[85] Herman JC Berendsen, JPM van Postma, Wilfred F Van Gunsteren, ARHJ DiNola, and Jan R Haak. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81(8):3684–3690, 1984.

[86] J v Neumann. Physical applications of the ergodic hypothesis. *Proceedings of the National Academy of Sciences*, 18(3):263–266, 1932.

[87] J v Neumann. Proof of the quasi-ergodic hypothesis. *Proceedings of the National Academy of Sciences*, 18(1):70–82, 1932.

[88] George D Birkhoff. Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences*, 17(12):656–660, 1931.

[89] Furio Ercolessi. A molecular dynamics primer. *Spring college in computational physics, ICTP, Trieste*, 19, 1997.

[90] Dominik Marx and Jurg Hutter. Ab initio molecular dynamics: Theory and implementation. *Modern methods and algorithms of quantum chemistry*, 1(301-449):141, 2000.

[91] Dominik Marx and Jürg Hutter. *Ab initio molecular dynamics: basic theory and advanced methods*. Cambridge University Press, 2009.

[92] Walter Thiel. Semiempirical quantum–chemical methods. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 4(2):145–157, 2014.

[93] John E Lennard-Jones. Cohesion. *Proceedings of the Physical Society*, 43(5):461, 1931.

[94] Philip M Morse. Diatomic molecules according to the wave mechanics. ii. vibrational levels. *Physical review*, 34(1):57, 1929.

[95] Frank H Stillinger and Thomas A Weber. Computer simulation of local order in condensed phases of silicon. *Physical review B*, 31(8):5262, 1985.

[96] RA Johnson. Alloy models with the embedded-atom method. *Physical Review B*, 39(17):12554, 1989.

[97] Bernard R Brooks, Charles L Brooks III, Alexander D Mackerell Jr, Lennart Nilsson, Robert J Petrella, Benoît Roux, Youngdo Won, Georgios Archontis, Christian Bartels, Stefan Boresch, et al. Charmm: the biomolecular simulation program. *Journal of computational chemistry*, 30(10):1545–1614, 2009.

[98] David A Case, H Metin Aktulga, Kellon Belfon, Ido Y Ben-Shalom, Joshua T Berryman, Scott R Brozell, David S Cerutti, Thomas E Cheatham III, G Andrés Cisneros, Vinícius Wilian D Cruzeiro, et al. *Amber 2023*. University of California, San Francisco, 2023.

[99] William L Jorgensen, David S Maxwell, and Julian Tirado-Rives. Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *Journal of the american chemical society*, 118(45):11225–11236, 1996.

[100] Thomas P Senftle, Sungwook Hong, Md Mahbubul Islam, Sudhir B Kylasa, Yuanxia Zheng, Yun Kyung Shin, Chad Junkermeier, Roman Engel-Herbert, Michael J Janik, Hasan Metin Aktulga, et al. The reaxff reactive force-field: development, applications and future directions. *npj Computational Materials*, 2(1):1–14, 2016.

[101] Llewellyn H Thomas. The calculation of atomic fields. In *Mathematical proceedings of the Cambridge philosophical society*, volume 23, pages 542–548. Cambridge University Press, 1927.

[102] Enrico Fermi. Statistical method to determine some properties of atoms. *Rend. Accad. Naz. Lincei*, 6(602-607):5, 1927.

[103] Pierre Hohenberg and Walter Kohn. Inhomogeneous electron gas. *Physical review*, 136(3B):B864, 1964.

[104] Walter Kohn and Lu Jeu Sham. Self-consistent equations including exchange and correlation effects. *Physical review*, 140(4A):A1133, 1965.

[105] John P Perdew and Karla Schmidt. Jacob's ladder of density functional approximations for the exchange-correlation energy. In *AIP conference proceedings*, volume 577, pages 1–20. American Institute of Physics, 2001.

[106] Zhaoming Xia and Hai Xiao. Grand canonical ensemble modeling of electrochemical interfaces made simple. *Journal of Chemical Theory and Computation*, 19(15):5168–5175, 2023.

[107] Nicéphore Bonnet, Tetsuya Morishita, Osamu Sugino, and Minoru Otani. First-principles molecular dynamics at a constant electrode potential. *Physical Review Letters*, 109(26):266101, 2012.

[108] Kai Zhu, Enrico Trizio, Jintu Zhang, Renling Hu, Linlong Jiang, Tingjun Hou, and Luigi Bonati. Enhanced sampling in the age of machine learning: Algorithms and applications. *arXiv preprint arXiv:2509.04291*, 2025.

[109] Daniel Ferreira de Lima Neto, Vagner Fonseca, Ronaldo Jesus, Leonardo Hermes Dutra, Layssa Miranda de Oliveria Portela, Carla Freitas, Eduardo Fillizola, Breno Soares, André Luiz de Abreu, Sandeep Twiari, et al. Molecular dynamics simulations of the sars-cov-2 spike protein and variants of concern: Structural evidence for convergent adaptive evolution. *Journal of Biomolecular Structure and Dynamics*, 41(12):5789–5801, 2023.

[110] Enrico Trizio, Andrea Rizzi, Pablo M Piaggi, Michele Invernizzi, and Luigi Bonati. Advanced simulations with PLUMED: OPES and machine learning collective variables. *arXiv preprint arXiv:2410.18019*, 2024.

[111] Johannes CB Dietschreit, Dennis J Diestler, Andreas Hulm, Christian Ochsenfeld, and Rafael Gómez-Bombarelli. From free-energy profiles to activation free energies. *The Journal of Chemical Physics*, 157(8), 2022.

[112] Glenn M Torrie and John P Valleau. Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. *Journal of computational physics*, 23(2):187–199, 1977.

[113] Alessandro Laio and Michele Parrinello. Escaping free-energy minima. *Proceedings of the national academy of sciences*, 99(20):12562–12566, 2002.

[114] Giovanni Bussi and Alessandro Laio. Using metadynamics to explore complex free-energy landscapes. *Nature Reviews Physics*, 2(4):200–212, 2020.

[115] Michele Invernizzi and Michele Parrinello. Rethinking metadynamics: from bias potentials to probability distributions. *The journal of physical chemistry letters*, 11(7):2731–2736, 2020.

[116] Dhiman Ray, Narjes Ansari, Valerio Rizzi, Michele Invernizzi, and Michele Parrinello. Rare event kinetics from adaptive bias enhanced sampling. *Journal of Chemical Theory and Computation*, 18(11):6500–6509, 2022.

[117] Dhiman Ray and Michele Parrinello. Kinetics from metadynamics: Principles, applications, and outlook. *Journal of Chemical Theory and Computation*, 19(17):5649–5670, 2023.

[118] Matteo Salvalaglio, Pratyush Tiwary, and Michele Parrinello. Assessing the reliability of the dynamics reconstructed from metadynamics. *Journal of chemical theory and computation*, 10(4):1420–1425, 2014.

[119] Kenneth S Kaminsky. Confidence intervals for the exponential scale parameter using optimally selected order statistics. *Technometrics*, 14(2):371–383, 1972.

[120] Emanuele Grifoni, GiovanniMaria Piccini, and Michele Parrinello. Microscopic description of acid–base equilibrium. *Proceedings of the National Academy of Sciences*, 116(10):4054–4057, 2019.

[121] Emanuele Grifoni, GiovanniMaria Piccini, and Michele Parrinello. Tautomeric equilibrium in condensed phases. *Journal of chemical theory and computation*, 16(10):6027–6031, 2020.

[122] Pengchao Zhang, Axel Tosello Gardini, Xuefei Xu, and Michele Parrinello. Intramolecular and water mediated tautomerism of solvated glycine. *Journal of Chemical Information and Modeling*, 64(9):3599–3604, 2024.

[123] Pablo M Piaggi and Michele Parrinello. Predicting polymorphism in molecular crystals using orientational entropy. *Proceedings of the National Academy of Sciences*, 115(41):10251–10256, 2018.

[124] Pablo M Piaggi and Michele Parrinello. Entropy based fingerprint for local crystalline order. *The Journal of chemical physics*, 147(11), 2017.

[125] Umberto Raucci, Valerio Rizzi, and Michele Parrinello. Discover, sample, and refine: Exploring chemistry with enhanced sampling techniques. *The Journal of Physical Chemistry Letters*, 13(6):1424–1430, 2022.

[126] Huan Yang, Umberto Raucci, Siddharth Iyer, Galib Hasan, Thomas Golin Almeida, Shawon Barua, Anni Savolainen, Juha Kangasluoma, Matti Rissanen, Hanna Vehkamäki, et al. Molecular dynamics-guided reaction discovery reveals endoperoxide-to-alkoxy radical isomerization as key branching point in α-pinene ozonolysis. *Nature Communications*, 16(1):661, 2025.

[127] Gareth A Tribello, Massimiliano Bonomi, Davide Branduardi, Carlo Camilloni, and Giovanni Bussi. Plumed 2: New feathers for an old bird. *Computer physics communications*, 185(2):604–613, 2014.

[128] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.

[129] Christopher M Bishop and Hugh Bishop. *Deep learning: Foundations and concepts*. Springer Nature, 2023.

[130] Pádraig Cunningham, Matthieu Cord, and Sarah Jane Delany. Supervised learning. In *Machine learning techniques for multimedia: case studies on organization and retrieval*, pages 21–49. Springer, 2008.

[131] Horace B Barlow. Unsupervised learning. *Neural computation*, 1(3):295–311, 1989.

[132] Michele Ceriotti. Unsupervised machine learning in atomistic simulations, between predictions and understanding. *The Journal of chemical physics*, 150(15), 2019.

[133] Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. Reinforcement learning: A survey. *Journal of artificial intelligence research*, 4:237–285, 1996.

[134] Kevin Gurney. *An introduction to neural networks*. CRC press, 2018.

[135] Herve Abdi. A neural network primer. *Journal of Biological Systems*, 2(03):247–281, 1994.

[136] Sagar Sharma, Simone Sharma, and Anidhya Athaiya. Activation functions in neural networks. *Towards Data Sci*, 6(12):310–316, 2017.

[137] Andrinandrasana David Rasamoelina, Fouzia Adjailia, and Peter Sinčák. A review of activation function for artificial neural network. In *2020 IEEE 18th world symposium on applied machine intelligence and informatics (SAMI)*, pages 281–286. IEEE, 2020.

[138] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5):359–366, 1989.

[139] Ronen Eldan and Ohad Shamir. The power of depth for feedforward neural networks. In *Conference on learning theory*, pages 907–940. PMLR, 2016.

[140] Murat H Sazlı. A brief review of feed-forward neural networks. *Communications Faculty of Sciences University of Ankara Series A2-A3 Physical Sciences and Engineering*, 50(01), 2006.

[141] Thomas B Blank, Steven D Brown, August W Calhoun, and Douglas J Doren. Neural network models of potential energy surfaces. *The Journal of chemical physics*, 103(10):4129–4137, 1995.

[142] Steven Hobday, Roger Smith, and Joe Belbruno. Applications of neural networks to fitting interatomic potential functions. *Modelling and Simulation in Materials Science and Engineering*, 7(3):397, 1999.

[143] Jörg Behler and Michele Parrinello. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Physical review letters*, 98(14):146401, 2007.

[144] Han Wang, Linfeng Zhang, Jiequn Han, and E Weinan. Deepmd-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Computer Physics Communications*, 228:178–184, 2018.

[145] Jinzhe Zeng, Duo Zhang, Anyang Peng, Xiangyu Zhang, Sensen He, Yan Wang, Xinzijian Liu, Hangrui Bi, Yifan Li, Chun Cai, et al. Deepmd-kit v3: a multiple-backend framework for machine learning potentials. *Journal of Chemical Theory and Computation*, 21(9):4375–4385, 2025.

[146] Gabriele Corso, Hannes Stark, Stefanie Jegelka, Tommi Jaakkola, and Regina Barzilay. Graph neural networks. *Nature Reviews Methods Primers*, 4(1):17, 2024.

[147] Alexandre Duval, Simon V Mathis, Chaitanya K Joshi, Victor Schmidt, Santiago Miret, Fragkiskos D Malliaros, Taco Cohen, Pietro Lio, Yoshua Bengio, and Michael Bronstein. A hitchhiker's guide to geometric gnns for 3d atomic systems. *arXiv preprint arXiv:2312.07511*, 2023.

[148] Ilyes Batatia, David P Kovacs, Gregor Simm, Christoph Ortner, and Gábor Csányi. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields. *Advances in neural information processing systems*, 35:11423–11436, 2022.

[149] I. Batatia, S. Batzner, D. P. Kovács, A. Musaelian, G. N. Simm, R. Drautz, C. Ortner, B. Kozinsky, and G. Csányi. The design space of e (3)-equivariant atom-centered interatomic potentials. *arXiv preprint arXiv:2205.06643*, May 2022.

[150] Ilyes Batatia, Philipp Benner, Yuan Chiang, Alin M Elena, Dávid P Kovács, Janosh Riebesell, Xavier R Advincula, Mark Asta, Matthew Avaylon, William J Baldwin, et al. A foundation model for atomistic materials chemistry. *arXiv preprint arXiv:2401.00096*, 2023.

[151] Dávid Péter Kovács, J Harry Moore, Nicholas J Browning, Ilyes Batatia, Joshua T Horton, Yixuan Pu, Venkat Kapil, William C Witt, Ioan-Bogdan Magdau, Daniel J Cole, et al. Mace-off: Short-range transferable machine learning force fields for organic molecules. *Journal of the American Chemical Society*, 147(21):17598–17611, 2025.

[152] Yinuo Yang, Shuhao Zhang, Kavindri D Ranasinghe, Olexandr Isayev, and Adrian E Roitberg. Machine learning of reactive potentials. *Annual review of physical chemistry*, 75(1):371–395, 2024.

[153] Ryan Jacobs, Dane Morgan, Siamak Attarian, Jun Meng, Chen Shen, Zhenghao Wu, Clare Yijia Xie, Julia H Yang, Nongnuch Artrith, Ben Blaiszik, et al. A practical guide to machine learning interatomic potentials–status and future. *Current Opinion in Solid State and Materials Science*, 35:101214, 2025.

[154] Emir Kocer, Tsz Wai Ko, and Jörg Behler. Neural network potentials: A concise overview of methods. *Annual review of physical chemistry*, 73:163–186, 2022.

[155] Amir Omranpour, Jan Elsner, K Nikolas Lausch, and Jorg Behler. Machine learning potentials for heterogeneous catalysis. *ACS Catalysis*, 15(3):1616–1634, 2025.

[156] Maksim Kulichenko, Benjamin Nebgen, Nicholas Lubbers, Justin S Smith, Kipton Barros, Alice EA Allen, Adela Habib, Emily Shinkle, Nikita Fedik, Ying Wai Li, et al. Data generation for machine learning interatomic potentials and beyond. *Chemical Reviews*, 124(24):13681–13714, 2024.

[157] Pieter Dobbelaere, Sander Vandenhaute, and Veronique Van Speybroeck. Cluster-based machine learning potentials to describe disordered metal-organic frameworks up to the mesoscale. *arXiv preprint arXiv:2504.03881*, 2025.

[158] Jie Wang. An intuitive tutorial to gaussian process regression. *Computing in Science & Engineering*, 25(4):4–11, 2023.

[159] Jonathan Vandermause, Steven B Torrisi, Simon Batzner, Yu Xie, Lixin Sun, Alexie M Kolpak, and Boris Kozinsky. On-the-fly active learning of interpretable bayesian force fields for atomistic rare events. *npj Computational Materials*, 6(1):20, 2020.

[160] Bowen Deng, Peichen Zhong, KyuJung Jun, Janosh Riebesell, Kevin Han, Christopher J Bartel, and Gerbrand Ceder. Chgnet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nature Machine Intelligence*, 5(9):1031–1041, 2023.

[161] Chi Chen and Shyue Ping Ong. A universal graph deep learning interatomic potential for the periodic table. *Nature Computational Science*, 2(11):718–728, 2022.

[162] Han Yang, Chenxi Hu, Yichi Zhou, Xixian Liu, Yu Shi, Jielan Li, Guanzhi Li, Zekun Chen, Shuizhou Chen, Claudio Zeni, et al. Mattersim: A deep learning atomistic model across elements, temperatures and pressures. *arXiv preprint arXiv:2405.04967*, 2024.

[163] Pietro Novelli, PJ Buigues, G Meanti, L Rosasco, M Parrinello, and M Pontil. Fine-tuning foundation models for molecular dynamics: A data-efficient approach with random features. *Proc. Adv. Neural Inf. Process. Syst.(NeurIPS)*, 2024.

[164] Pietro Novelli, Giacomo Meanti, Pedro J Buigues, Lorenzo Rosasco, Michele Parrinello, Massimiliano Pontil, and Luigi Bonati. Fast and fourier features for transfer learning of interatomic potentials. *npj Computational Materials*, 11(1):293, 2025.

[165] Ian T Jolliffe and Jorge Cadima. Principal component analysis: a review and recent developments. *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202, 2016.

[166] Dan Mendels, GiovanniMaria Piccini, and Michele Parrinello. Collective variables from local fluctuations. *The journal of physical chemistry letters*, 9(11):2776–2781, 2018.

[167] GiovanniMaria Piccini, Dan Mendels, and Michele Parrinello. Metadynamics with discriminants: A tool for understanding chemistry. *Journal of chemical theory and computation*, 14(10):5040–5044, 2018.

[168] Guillermo Pérez-Hernández, Fabian Paul, Toni Giorgino, Gianni De Fabritiis, and Frank Noé. Identification of slow molecular order parameters for markov model construction. *The Journal of chemical physics*, 139(1), 2013.

[169] Luigi Bonati, Valerio Rizzi, and Michele Parrinello. Data-driven collective variables for enhanced sampling. *The journal of physical chemistry letters*, 11(8):2998–3004, 2020.

[170] Manyi Yang, Enrico Trizio, and Michele Parrinello. Structure and polymerization of liquid sulfur across the λ-transition. *Chemical Science*, 15(9):3382–3392, 2024.

[171] Peilin Kang, Enrico Trizio, and Michele Parrinello. Computing the committor with the committor to study the transition state ensemble. *Nature Computational Science*, 4(6):451–460, 2024.

[172] Enrico Trizio, Peilin Kang, and Michele Parrinello. Everything everywhere all at once: a probability-based enhanced sampling approach to rare events. *Nature Computational Science*, pages 1–10, 2025.

[173] Wenbo Gao, Jianping Guo, Peikun Wang, Qianru Wang, Fei Chang, Qijun Pei, Weijin Zhang, Lin Liu, and Ping Chen. Production of ammonia via a chemical looping process based on metal imides as nitrogen carriers. *Nature Energy*, 3(12):1067–1075, 2018.

[174] Yeqin Guan, Chuangwei Liu, Qianru Wang, Wenbo Gao, Heine Anton Hansen, Jianping Guo, Tejs Vegge, and Ping Chen. Transition-metal-free barium hydride mediates dinitrogen fixation and ammonia synthesis. *Angewandte Chemie International Edition*, 61(39):e202205805, 2022.

[175] Sai A Varanasi, Carlos A Fernández, and Marta C Hatzell. Energy management and economic considerations of intermittent photovoltaic-driven electrochemical ammonia production. *Energy & Fuels*, 37(19):15222–15230, 2023.

[176] Behdad Moghtaderi. Review of the recent chemical looping process developments for novel energy and fuel applications. *Energy & fuels*, 26(1):15–40, 2012.

[177] Lance C Seefeldt, Zhi-Yong Yang, Dmitriy A Lukoyanov, Derek F Harris, Dennis R Dean, Simone Raugei, and Brian M Hoffman. Reduction of substrates by nitrogenases. *Chemical reviews*, 120(12):5082–5106, 2020.

[178] Maarten C Verbraeken, Chaksum Cheung, Emmanuelle Suard, and John TS Irvine. High h-ionic conductivity in barium hydride. *Nature materials*, 14(1):95–100, 2015.

[179] David T Peterson and M Indig. The barium-barium hydride phase system1. *Journal of the American Chemical Society*, 82(21):5645–5646, 1960.

[180] W Bronger, Chi-Chien Scha, and P.+ Mueller. Crystal structure of barium hydride, determined by neutron diffraction experiments on bad 2. *Zeitschrift fuer Anorganische und Allgemeine Chemie (1950)*, 545, 1987.

[181] Eric Novak, Luke Daemen, Anibal Javier Ramirez-Cuesta, Yongqiang Cheng, Robert Smith, Takeshi Egami, and Niina Jalarvo. Uncovering the hydride ion diffusion pathway in barium hydride via neutron spectroscopy. *Scientific Reports*, 12(1):6194, 2022.

[182] GJ Snyder, H Borrmann, and A Simon. Crystal structure of barium dihydride, bah2. *Zeitschrift für Kristallographie-Crystalline Materials*, 209(5):458–458, 1994.

[183] Jun Lu, Zhigang Zak Fang, and Hong Yong Sohn. A dehydrogenation mechanism of metal hydrides based on interactions between hδ+ and h-. *Inorganic chemistry*, 45(21):8749–8754, 2006.

[184] Andrew JE Rowberg, Leigh Weston, and Chris G Van de Walle. Ion-transport engineering of alkaline-earth hydrides for hydride electrolyte applications. *Chemistry of Materials*, 30(17):5878–5885, 2018.

[185] Hideo Hosono and Masaaki Kitano. Advances in materials and applications of inorganic electrides. *Chemical Reviews*, 121(5):3121–3185, 2021.

[186] Alexander Stukowski. Visualization and analysis of atomistic simulation data with ovito–the open visualization tool. *Modelling and simulation in materials science and engineering*, 18(1):015012, 2009.

[187] Masaaki Kitano, Yasunori Inoue, Youhei Yamazaki, Fumitaka Hayashi, Shinji Kanbara, Satoru Matsuishi, Toshiharu Yokoyama, Sung-Wng Kim, Michikazu Hara, and Hideo Hosono. Ammonia synthesis using a stable electride as an electron donor and reversible hydrogen store. *Nature chemistry*, 4(11):934–940, 2012.

[188] Masaaki Kitano, Shinji Kanbara, Yasunori Inoue, Navaratnarajah Kuganathan, Peter V Sushko, Toshiharu Yokoyama, Michikazu Hara, and Hideo Hosono. Electride support boosts nitrogen dissociation over ruthenium catalyst and shifts the bottleneck in ammonia synthesis. *Nature communications*, 6(1):6731, 2015.

[189] Jiazhen Wu, Jiang Li, Yutong Gong, Masaaki Kitano, Takeshi Inoshita, and Hideo Hosono. Intermetallic electride catalyst as a platform for ammonia synthesis. *Angewandte Chemie International Edition*, 58(3):825–829, 2019.

[190] Wei Chen, Christoph Tegenkamp, Herbert Pfnür, and Thomas Bredow. Tailoring band gaps of insulators by adsorption at surface defects: Benzoic acids on nacl surfaces. *Physical Review B—Condensed Matter and Materials Physics*, 79(23):235419, 2009.

[191] Manuel Hochheim and Thomas Bredow. Adsorption of ptcda on nacl surfaces with color centers: Charge transfer and formation of radical ions. *The Journal of Physical Chemistry C*, 122(51):29426–29434, 2018.

[192] Michael Hafner, Manuel Hochheim, and Thomas Bredow. Chemistry with f centers: Reduction of organic molecules on the defective potassium chloride (100) surface. *The Journal of Physical Chemistry C*, 124(23):12606–12616, 2020.

[193] Inigo Marcos-Alcalde, Javier Setoain, Jesus I Mendieta-Moreno, Jesus Mendieta, and Paulino Gomez-Puertas. Mepsa: minimum energy pathway analysis for energy landscapes. *Bioinformatics*, 31(23):3853–3855, 2015.

[194] Alexander Stukowski. Computational analysis methods in atomistic modeling of crystals. *Jom*, 66(3):399–407, 2014.

[195] Herbert Edelsbrunner and Ernst P Mücke. Three-dimensional alpha shapes. *ACM Transactions On Graphics (TOG)*, 13(1):43–72, 1994.

[196] Yeqin Guan, Hong Wen, Kaixun Cui, Qianru Wang, Wenbo Gao, Yongli Cai, Zibo Cheng, Qijun Pei, Zhao Li, Hujun Cao, et al. Light-driven ammonia synthesis under mild conditions using lithium hydride. *Nature Chemistry*, pages 1–7, 2024.

[197] RFW Bader, W H_ Henneker, and Paul E Cade. Molecular charge distributions and chemical binding. *The Journal of Chemical Physics*, 46(9):3341–3363, 1967.

[198] Richard FW Bader and Preston J MacDougall. Toward a theory of chemical reactivity based on the charge density. *Journal of the American Chemical Society*, 107(24):6788–6795, 1985.

[199] Qianru Wang, Jaysree Pan, Jianping Guo, Heine Anton Hansen, Hua Xie, Ling Jiang, Lei Hua, Haiyang Li, Yeqin Guan, Peikun Wang, et al. Ternary ruthenium complex hydrides for ammonia synthesis via the associative mechanism. *Nature Catalysis*, 4(11):959–967, 2021.

[200] Michele Parrinello and Aneesur Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, 52(12):7182–7190, 1981.

[201] Wenbo Gao, Peikun Wang, Jianping Guo, Fei Chang, Teng He, Qianru Wang, Guotao Wu, and Ping Chen. Barium hydride-mediated nitrogen transfer and hydrogenation for ammonia synthesis: a case study of cobalt. *ACS Catalysis*, 7(5):3654–3661, 2017.

[202] Qianru Wang, Yeqin Guan, Jianping Guo, and Ping Chen. Hydrides mediate nitrogen fixation. *Cell Reports Physical Science*, 3(3), 2022.

[203] Phuong-Vu Ong, Lewis E Johnson, Hideo Hosono, and Peter V Sushko. Structure and stability of cah 2 surfaces: on the possibility of electron-rich surfaces in metal hydrides for catalysis. *Journal of materials chemistry A*, 5(11):5550–5558, 2017.

[204] Masaaki Kitano, Yasunori Inoue, Hiroki Ishikawa, Kyosuke Yamagata, Takuya Nakao, Tomofumi Tada, Satoru Matsuishi, Toshiharu Yokoyama, Michikazu Hara, and Hideo Hosono. Essential role of hydride ion in ruthenium-based ammonia synthesis catalysts. *Chemical science*, 7(7):4036–4043, 2016.

[205] Wenbo Gao, Qianru Wang, Yeqin Guan, Hanxue Yan, Jianping Guo, and Ping Chen. Barium hydride activates ni for ammonia synthesis catalysis. *Faraday Discussions*, 243:27–37, 2023.

[206] Joost VandeVondele, Matthias Krack, Fawzi Mohamed, Michele Parrinello, Thomas Chassaing, and Jürg Hutter. Quickstep: Fast and accurate density functional calculations using a mixed gaussian and plane waves approach. *Computer Physics Communications*, 167(2):103–128, 2005.

[207] Thomas D Kühne, Marcella Iannuzzi, Mauro Del Ben, Vladimir V Rybkin, Patrick Seewald, Frederick Stein, Teodoro Laino, Rustam Z Khaliullin, Ole Schütt, Florian Schiffmann, et al. Cp2k: An electronic structure and molecular dynamics software package-quickstep: Efficient and accurate electronic structure calculations. *The Journal of Chemical Physics*, 152(19), 2020.

[208] John P Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 77(18):3865, 1996.

[209] Stefan Goedecker, Michael Teter, and Jürg Hutter. Separable dual-space gaussian pseudopotentials. *Physical Review B*, 54(3):1703, 1996.

[210] Christian Hartwigsen, Sephen Gœdecker, and Jürg Hutter. Relativistic separable dual-space gaussian pseudopotentials from h to rn. *Physical Review B*, 58(7):3641, 1998.

[211] Paolo Giannozzi, Stefano Baroni, Nicola Bonini, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, Guido L Chiarotti, Matteo Cococcioni, Ismaila Dabo, et al. Quantum espresso: a modular and open-source software project for quantum simulations of materials. *Journal of physics: Condensed matter*, 21(39):395502, 2009.

[212] Paolo Giannozzi, Oliviero Andreussi, Thomas Brumme, Oana Bunau, M Buongiorno Nardelli, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, Matteo Cococcioni, et al. Advanced capabilities for materials modelling with quantum espresso. *Journal of physics: Condensed matter*, 29(46):465901, 2017.

[213] Paolo Giannozzi, Oscar Baseggio, Pietro Bonfà, Davide Brunato, Roberto Car, Ivan Carnimeo, Carlo Cavazzoni, Stefano De Gironcoli, Pietro Delugas, Fabrizio Ferrari Ruffino, et al. Quantum espresso toward the exascale. *The Journal of chemical physics*, 152(15), 2020.

[214] Andrew M Rappe, Karin M Rabe, Efthimios Kaxiras, and JD Joannopoulos. Optimized pseudopotentials. *Physical Review B*, 41(2):1227, 1990.

[215] Stephen D Bond and Benedict J Leimkuhler. Stabilized integration of hamiltonian systems with hard-sphere inequality constraints. *SIAM Journal on Scientific Computing*, 30(1):134–147, 2008.

[216] Aidan P Thompson, H Metin Aktulga, Richard Berger, Dan S Bolintineanu, W Michael Brown, Paul S Crozier, Pieter J In't Veld, Axel Kohlmeyer, Stan G Moore, Trung Dac Nguyen, et al. Lammps-a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications*, 271:108171, 2022.

[217] Qiyu Zeng, Bo Chen, Xiaoxiang Yu, Shen Zhang, Dongdong Kang, Han Wang, and Jiayu Dai. Towards large-scale and spatiotemporally resolved diagnosis of electronic density of states by deep learning. *Physical Review B*, 105(17):174109, 2022.

[218] Duo Zhang, Hangrui Bi, Fu-Zhi Dai, Wanrun Jiang, Linfeng Zhang, and Han Wang. Dpa-1: Pretraining of attention-based deep potential model for molecular simulation. *arXiv preprint arXiv:2208.08236*, 2022.

[219] Graeme Henkelman, Andri Arnaldsson, and Hannes Jónsson. A fast and robust algorithm for bader decomposition of charge density. *Computational Materials Science*, 36(3):354–360, 2006.

[220] Leyu Liu and Hai Xiao. Inverted Region in Electrochemical Reduction of CO2 Induced by Potential-Dependent Pauli Repulsion. *Journal of the American Chemical Society*, 145(26):14267–14275, 2023.

[221] Yipeng Zhou, Yixin Ouyang, Yehui Zhang, Qiang Li, and Jinlan Wang. Machine Learning Assisted Simulations of Electrochemical Interfaces: Recent Progress and Challenges. *The Journal of Physical Chemistry Letters*, 14(9):2308–2316, 2023.

[222] J. B. Le, X. H. Yang, Y. B. Zhuang, M. Jia, and J. Cheng. Recent Progress toward Ab Initio Modeling of Electrocatalysis. *J Phys Chem Lett*, 12(37):8924–8931, 2021.

[223] Jiali Wang, Hui-Ying Tan, Ming-Yu Qi, Jing-Yu Li, Zi-Rong Tang, Nian-Tzu Suen, Yi-Jun Xu, and Hao Ming Chen. Spatially and temporally understanding dynamic solid–electrolyte interfaces in carbon dioxide electroreduction. *Chemical Society Reviews*, 52(15):5013–5050, 2023.

[224] Stefan Ringe, Ezra L. Clark, Joaquin Resasco, Amber Walton, Brian Seger, Alexis T. Bell, and Karen Chan. Understanding cation effects in electrochemical CO $_2$ reduction. *Energy & Environmental Science*, 12(10):3001–3014, 2019.

[225] Michele Re Fiorentin, Francesca Risplendi, Clara Salvini, Juqin Zeng, Giancarlo Cicero, and Hannes Jónsson. Silver Electrodes Are Highly Selective for CO in CO2 Electroreduction due to Interplay between Voltage Dependent Kinetics and Thermodynamics. *The Journal of Physical Chemistry Letters*, 15(46):11538–11545, 2024.

[226] Leanne D. Chen, Makoto Urushihara, Karen Chan, and Jens K. Nørskov. Electric Field Effects in Electrochemical CO2 Reduction. *ACS Catalysis*, 6(10):7133–7139, 2016.

[227] Kiran Mathew, VS Kolluru, Srinidhi Mula, Stephan N Steinmann, and Richard G Hennig. Implicit self-consistent electrolyte model in plane-wave density-functional theory. *The Journal of chemical physics*, 151(23), 2019.

[228] Karen Chan and Jens K. Nørskov. Potential Dependence of Electrochemical Barriers from ab Initio Calculations. *The Journal of Physical Chemistry Letters*, 7(9):1686–1690, May 2016.

[229] Saerom Yu, Zachary Levell, Zhou Jiang, Xunhua Zhao, and Yuanyue Liu. What Is the Rate-Limiting Step of Oxygen Reduction Reaction on Fe–N–C Catalysts? *Journal of the American Chemical Society*, 145(46):25352–25356, November 2023.

[230] Balázs Endrődi, A Samu, E Kecsenovity, T Halmágyi, D Sebők, and Cs Janáky. Operando cathode activation with alkali metal cations for high current density operation of water-fed zero-gap carbon dioxide electrolysers. *Nature energy*, 6(4):439–448, 2021.

[231] Jiayi Chen and Lei Wang. Effects of the catalyst dynamic changes and influence of the reaction environment on the performance of electrochemical co2 reduction. *Advanced Materials*, 34(25):2103900, 2022.

[232] Md Golam Kibria, Jonathan P Edwards, Christine M Gabardo, Cao-Thang Dinh, Ali Seifitokaldani, David Sinton, and Edward H Sargent. Electrochemical co2 reduction into chemical feedstocks: from mechanistic electrocatalysis models to system design. *Advanced Materials*, 31(31):1807166, 2019.

[233] Kendra P Kuhl, Toru Hatsukade, Etosha R Cave, David N Abram, Jakob Kibsgaard, and Thomas F Jaramillo. Electrocatalytic conversion of carbon dioxide to methane and methanol on transition metal surfaces. *Journal of the American Chemical Society*, 136(40):14107–14113, 2014.

[234] Ming Ma, Bartek J Trześniewski, Jie Xie, and Wilson A Smith. Selective and efficient reduction of carbon dioxide to carbon monoxide on oxide-derived nanostructured silver electrocatalysts. *Angewandte Chemie*, 128(33):9900–9904, 2016.

[235] AN Frumkin, OA Petrii, and BB Damaskin. Potentials of zero charge. *Comprehensive Treatise of Electrochemistry: The Double Layer*, pages 221–289, 1980.

[236] Chao-Yu Li, Jia-Bo Le, Yao-Hui Wang, Shu Chen, Zhi-Lin Yang, Jian-Feng Li, Jun Cheng, and Zhong-Qun Tian. In situ probing electrified interfacial water structures at atomically flat surfaces. *Nature materials*, 18(7):697–701, 2019.

[237] Nicholas R Larson, Yangjie Wei, Thayana Araújo Cruz, Reza Esfandiary, Cavan K Kalonia, M Laird Forrest, and C Russell Middaugh. Conformational changes and drivers of monoclonal antibody liquid-liquid phase separation. *Journal of pharmaceutical sciences*, 112(3):680–690, 2023.

[238] Mohammed Ahmed, V Namboodiri, Ajay K Singh, and Jahur A Mondal. On the intermolecular vibrational coupling, hydrogen bonding, and librational freedom of water in the hydration shell of mono-and bivalent anions. *The Journal of Chemical Physics*, 141(16), 2014.

[239] Jia-Bo Le, Qi-Yuan Fan, Jie-Qiong Li, and Jun Cheng. Molecular origin of negative component of helmholtz capacitance at electrified pt (111)/water interface. *Science advances*, 6(41):eabb1219, 2020.

[240] Yao-Hui Wang, Shisheng Zheng, Wei-Min Yang, Ru-Yu Zhou, Quan-Feng He, Petar Radjenovic, Jin-Chao Dong, Shunning Li, Jiaxin Zheng, Zhi-Lin Yang, et al. In situ raman spectroscopy reveals the structure and dissociation of interfacial water. *Nature*, 600(7887):81–85, 2021.

[241] Mariana CO Monteiro, Federico Dattila, Bellenod Hagedoorn, Rodrigo García-Muelas, Núria López, and Marc TM Koper. Absence of co2 electroreduction on copper, gold and silver electrodes without metal cations in solution. *Nature Catalysis*, 4(8):654–662, 2021.

[242] Xueping Qin, Heine A Hansen, Karoliina Honkala, and Marko M Melander. Cation-induced changes in the inner-and outer-sphere mechanisms of electrocatalytic co2 reduction. *Nature Communications*, 14(1):7607, 2023.

[243] Yousef A. Alsunni, Abdulaziz W. Alherz, and Charles B. Musgrave. Electrocatalytic Reduction of CO2 to CO over Ag(110) and Cu(211) Modeled by Grand-Canonical Density Functional Theory. *The Journal of Physical Chemistry C*, 125(43):23773–23783, November 2021.

[244] Dominik Marx, Amalendu Chandra, and Mark E Tuckerman. Aqueous basic solutions: hydroxide solvation, structural diffusion, and comparison to the hydrated proton. *Chemical reviews*, 110(4):2174–2216, 2010.

[245] Jonathan Rosen, Gregory S. Hutchings, Qi Lu, Sean Rivera, Yang Zhou, Dionisios G. Vlachos, and Feng Jiao. Mechanistic insights into the electrochemical reduction of co2 to co on nanostructured ag surfaces. *ACS Catalysis*, 5(7):4293–4299, 2015.

[246] Xia-Guang Zhang, Xi Jin, De-Yin Wu, and Zhong-Qun Tian. Selective electrocatalytic mechanism of co2 reduction reaction to co on silver electrodes: A unique reaction intermediate. *The Journal of Physical Chemistry C*, 122(44):25447–25455, 2018.

[247] Xiaoxia Chang, Jing Li, Haocheng Xiong, Haochen Zhang, Yifei Xu, Hai Xiao, Qi Lu, and Bingjun Xu. C-C Coupling Is Unlikely to Be the Rate-Determining Step in the Formation of C2+ Products in the Copper-Catalyzed Electrochemical Reduction of CO. *Angewandte Chemie International Edition*, 61(2):e202111167, January 2022.

[248] Yuliang Yuan, Jin Li, Yiting Zhu, Yan Qiao, Zhenye Kang, Zhitong Wang, Xinlong Tian, Hongwen Huang, and Wenchuan Lai. Water in electrocatalysis. *Angewandte Chemie International Edition*, 64(18):e202425590, 2025.

[249] Matthias M Waegele, Charuni M Gunathunge, Jingyi Li, and Xiang Li. How cations affect the electric double layer and the rates and selectivity of electrocatalytic processes. *The Journal of chemical physics*, 151(16), 2019.

[250] Georg Kresse and Jürgen Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical review B*, 54(16):11169, 1996.

[251] Georg Kresse and Jürgen Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational materials science*, 6(1):15–50, 1996.

[252] Georg Kresse and Daniel Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical review b*, 59(3):1758, 1999.

[253] Peter E Blöchl. Projector augmented-wave method. *Physical review B*, 50(24):17953, 1994.

[254] James W Furness, Aaron D Kaplan, Jinliang Ning, John P Perdew, and Jianwei Sun. Accurate and numerically efficient r2scan meta-generalized gradient approximation. *The journal of physical chemistry letters*, 11(19):8208–8215, 2020.

[255] Jinliang Ning, Manish Kothakonda, James W Furness, Aaron D Kaplan, Sebastian Ehlert, Jan Gerit Brandenburg, John P Perdew, and Jianwei Sun. Workhorse minimally empirical dispersion-corrected density functional with tests for weakly bound systems: r 2 scan+ rvv 10. *Physical Review B*, 106(7):075422, 2022.

[256] Kiran Mathew, Ravishankar Sundararaman, Kendra Letchworth-Weaver, TA Arias, and Richard G Hennig. Implicit solvation model for density-functional study of nanocrystal surfaces and reaction pathways. *The Journal of chemical physics*, 140(8), 2014.

[257] Hao Cao, Xinmao Lv, Shengjie Qian, Jun Li, and Yang-Gang Wang. Constant potential thermodynamic integration for obtaining the free energy profile of electrochemical reaction. *The Journal of Physical Chemistry Letters*, 15(5):1314–1320, 2024.

[258] Jan Vavra, Gaétan PL Ramona, Federico Dattila, Attila Kormányos, Tatiana Priamushko, Petru P Albertini, Anna Loiudice, Serhiy Cherevko, Núria Lopéz, and Raffaella Buonsanti. Solution-based cu+ transient species mediate the reconstruction of copper electrocatalysts for co2 reduction. *Nature Catalysis*, 7(1):89–97, 2024.

[259] Dávid Péter Kovács, Ilyes Batatia, Eszter Sára Arany, and Gábor Csányi. Evaluation of the mace force field architecture: From medicinal chemistry to materials science. *The Journal of Chemical Physics*, 159(4), 2023.

[260] Md Arifur Rahim, Jianbo Tang, Andrew J Christofferson, Priyank V Kumar, Nastaran Meftahi, Franco Centurion, Zhenbang Cao, Junma Tang, Mahroo Baharfar, Mohannad Mayyas, et al. Low-temperature liquid platinum catalyst. *Nature Chemistry*, 14(8):935–941, 2022.

[261] Zujian Tang and Xiaofei Guan. Method for catalytic synthesis of ammonia under normal pressures, October 2024. Assigned to ShanghaiTech University.

[262] TL Bush, DO Hayward, and TS Jones. The sodium promoted nitridation of Si (100)-2$\times$ 1 using N2 molecular beams. *Surface science*, 313(1-2):179–187, 1994.

[263] TL Bush, DO Hayward, and TS Jones. A molecular beam study of the reaction of N2 at clean and sodium covered Si (100) surfaces. *Surface science*, 331:306–310, 1995.

[264] F Kawamura, M Morishita, K Omae, M Yoshimura, Y Mori, and T Sasaki. The effects of Na and some additives on nitrogen dissolution in the Ga-Na system: A growth mechanism of GaN in the Na flux method. *Journal of Materials Science: Materials in Electronics*, 16:29–34, 2005.

[265] Fumio Kawamura and Takashi Taniguchi. Method for synthesizing ammonia, and apparatus for said method, August 2022. Assigned to National Institute for Materials Science.

[266] Caian Qiu, Susanne M Opalka, Gregory B Olson, and Donald L Anton. The Na-H system: from first-principles calculations to thermodynamic modeling. *International journal of materials research*, 97(6):845–853, 2022.

[267] EF McFarlane and FC Tompkins. Nitridation of lithium. *Transactions of the Faraday Society*, 58:997–1007, 1962.

[268] Thomas Ludwig, Aayush R Singh, and Jens K Nørskov. Subsurface nitrogen dissociation kinetics in lithium metal from metadynamics. *The Journal of Physical Chemistry C*, 124(48):26368–26378, 2020.

[269] Charles E Messer, Edwin B Damon, P Calvin Maybury, John Mellor, and Regina A Seales. Solid–liquid equilibrium in the lithium lithium hydride system. *The Journal of Physical Chemistry*, 62(2):220–222, 1958.

[270] Ross T Howie, Olga Narygina, Christophe L Guillaume, Shaun Evans, and Eugene Gregoryanz. High-pressure synthesis of lithium hydride. *Physical Review B—Condensed Matter and Materials Physics*, 86(6):064108, 2012.

[271] RA Khairulin, RN Abdullaev, and SV Stankus. Phase equilibria and mutual diffusion in liquid lithium-sodium alloys. *Journal of Engineering Thermophysics*, 28(4):472–483, 2019.

[272] Henner Ruppersberg and Walter Knoll. Short range order in liquid lithium-sodium alloys. *Zeitschrift für Naturforschung A*, 32(12):1374–1382, 1977.

[273] ON Salmon and Donald H Ahmann. The lithium-sodium liquid metal system. *The Journal of Physical Chemistry*, 60(1):13–14, 1956.

[274] CW Bale. The li-na (lithium-sodium) system. *Bulletin of Alloy Phase Diagrams*, 10(3):265–268, 1989.

[275] Yang Huang, Michael Widom, and Michael C Gao. Ab initio free energies of liquid metal alloys: Application to the phase diagrams of li-na and k-na. *Physical Review Materials*, 6(1):013802, 2022.

[276] Manel Canales, Antoni Giro, and Joan Ángel Padró. Molecular dynamics simulation of Li-Mg and Li-Na alloys. *Journal of non-crystalline solids*, 205:907–910, 1996.

[277] AB Bhatia and D Eo Thornton. Structural aspects of the electrical resistivity of binary alloys. *Physical Review B*, 2(8):3004, 1970.

[278] Haoyan Huo and Matthias Rupp. Unified representation of molecules and crystals for machine learning. *Machine Learning: Science and Technology*, 3(4):045017, 2022.

[279] Jarno Laakso, Lauri Himanen, Henrietta Homm, Eiaki V Morooka, Marc OJ Jäger, Milica Todorović, and Patrick Rinke. Updates to the dscribe library: New descriptors and derivatives. *The Journal of Chemical Physics*, 158(23), 2023.

[280] Roel Prins. Hydrogen spillover. facts and fiction. *Chemical reviews*, 112(5):2714–2738, 2012.

[281] Philipp Stegner, Christian Färber, Ulrich Zenneck, Christian Knüpfer, Jonathan Eyselein, Michael Wiesinger, and Sjoerd Harder. Metallic barium: a versatile and efficient hydrogenation catalyst. *Angewandte Chemie*, 133(8):4298–4304, 2021.

[282] Masaru Yamaguchi, Yozaburo Hiraki, Masatada Wada, and Kimio Tarama. The catalytic hydrogenation of ethylene over calcium hydride and the role of the hydrogen of the hydride. *Bulletin of the Chemical Society of Japan*, 48(5):1345–1348, 1975.

[283] Leon Wright and Sol Weller. The catalytic activity of barium and calcium hydrides. i. an exploratory study. *Journal of the American Chemical Society*, 76(21):5305–5308, 1954.

[284] Sol Weller and Leon Wright. The catalytic activity of barium and calcium hydrides. ii. the nature of the catalyst. *Journal of the American Chemical Society*, 76(21):5302–5305, 1954.

[285] Leon Wright and Sol Weller. The catalytic activity of barium and calcium hydrides. iii. hydrogen exchange with some c4 hydrocarbons. *Journal of the American Chemical Society*, 76(23):5948–5950, 1954.

[286] Xu Zhang and Zhen Zhou. Perspective on theoretical models for co2 electrochemical reduction. *The Journal of Physical Chemistry C*, 126(8):3820–3829, 2022.

[287] Federico Dattila and Núria López. The potential of cations for carbon dioxide electroreduction. *ChemCatChem*, page e00749, 2025.

[288] Vivek Sinha, Elena Khramenkova, and Evgeny A Pidko. Solvent-mediated outer-sphere co 2 electro-reduction mechanism over the ag111 surface. *Chemical Science*, 13(13):3803–3808, 2022.

[289] R Michalsky, AM Avram, BA Peterson, Peter H Pfromm, and AA Peterson. Chemical looping of metal nitride catalysts: low-pressure ammonia synthesis for energy storage. *Chemical science*, 6(7):3965–3974, 2015.

[290] A Daisley and JSJ Hargreaves. Metal nitrides, the mars-van krevelen mechanism and heterogeneously catalysed ammonia synthesis. *Catalysis Today*, 423:113874, 2023.

[291] Tian-Nan Ye, Sang-Won Park, Yangfan Lu, Jiang Li, Masato Sasase, Masaaki Kitano, and Hideo Hosono. Contribution of nitrogen vacancies to ammonia synthesis over metal nitride catalysts. *Journal of the American Chemical Society*, 142(33):14374–14383, 2020.

[292] Constantinos D Zeinalipour-Yazdi, Justin SJ Hargreaves, and C Richard A Catlow. Dft-d3 study of molecular n2 and h2 activation on co3mo3n surfaces. *The Journal of Physical Chemistry C*, 120(38):21390–21398, 2016.

[293] Constantinos D Zeinalipour-Yazdi, Justin SJ Hargreaves, and C Richard A Catlow. Low-t mechanisms of ammonia synthesis on co3mo3n. *The Journal of Physical Chemistry C*, 122(11):6078–6082, 2018.

[294] Peikun Wang, Hua Xie, Jianping Guo, Zhi Zhao, Xiangtao Kong, Wenbo Gao, Fei Chang, Teng He, Guotao Wu, Mingshu Chen, et al. The formation of surface lithium–iron ternary hydride and its function on catalytic ammonia synthesis at low temperatures. *Angewandte Chemie*, 129(30):8842–8846, 2017.

[295] Yeqin Guan, Weijin Zhang, Qianru Wang, Claudia Weidenthaler, Anan Wu, Wenbo Gao, Qijun Pei, Hanxue Yan, Jirong Cui, Han Wu, et al. Barium chromium nitride-hydride for ammonia synthesis. *Chem Catalysis*, 1(5):1042–1054, 2021.

[296] Yu Cao, Maria A Kirsanova, Masayuki Ochi, Walid Al Maksoud, Tong Zhu, Rohit Rai, Shenghan Gao, Tatsuya Tsumori, Shintaro Kobayashi, Shogo Kawaguchi, et al. Topochemical synthesis of ca3crn3h involving a rotational structural transformation for catalytic ammonia synthesis. *Angewandte Chemie International Edition*, 61(39):e202209187, 2022.

[297] Matthew K Horton, Patrick Huck, Ruo Xi Yang, Jason M Munro, Shyam Dwaraknath, Alex M Ganose, Ryan S Kingsbury, Mingjian Wen, Jimmy X Shen, Tyler S Mathis, et al. Accelerated data-driven materials science with the materials project. *Nature Materials*, pages 1–11, 2025.

[298] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, et al. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL materials*, 1(1), 2013.