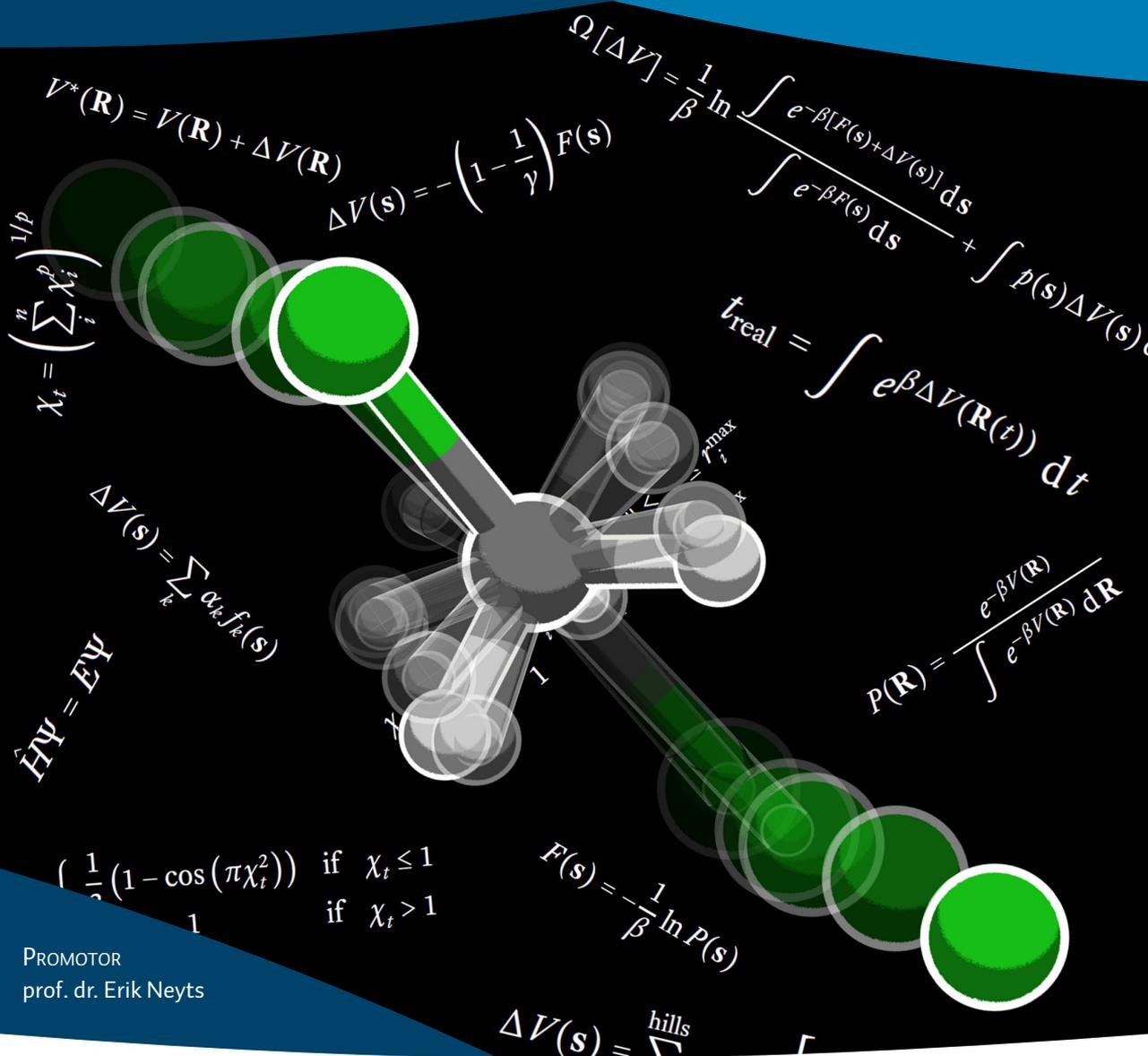


# New ways to bridge the gap between microscopic simulations and macroscopic chemistry

Proefschrift voorgelegd tot het behalen van de graad van doctor  
in de wetenschappen: chemie aan de Universiteit Antwerpen  
te verdedigen door

Kristof Bal



PROMOTOR  
prof. dr. Erik Neyts



# New ways to bridge the gap between microscopic simulations and macroscopic chemistry

Proefschrift voorgelegd tot het behalen  
van de graad doctor in de wetenschappen: chemie  
aan de Universiteit Antwerpen  
te verdedigen door  
Kristof Bal

Promotor: prof. dr. Erik Neyts



*Für Helena & Philippe,  
meine Oma und mein Opa,  
die immer so stolz auf mich waren*



# Contents

<b>List of Figures</b>	<b>vii</b>
<b>List of Tables</b>	<b>xi</b>
<b>Acknowledgments</b>	<b>xiii</b>
<b>Model or reality?</b>	<b>1</b>
<b>I Macroscopic time scales from microscopic simulations</b>	<b>5</b>
<b>1 Simulating slow processes</b>	<b>7</b>
1.1 The time scale problem in atomistic simulations . . . . .	7
1.2 Accelerated MD and related methods . . . . .	8
1.2.1 Infrequent event systems . . . . .	8
1.2.2 Full coarse-graining . . . . .	9
1.2.3 Temperature elevation . . . . .	11
1.2.4 Bias potentials . . . . .	12
1.2.5 Parallelization . . . . .	13
1.3 Enhanced sampling: metadynamics . . . . .	14
1.3.1 Sampling in collective variable space . . . . .	14
1.3.2 How compute the free energy? . . . . .	16
1.3.3 Metadynamics . . . . .	17
1.3.4 Dynamics from metadynamics . . . . .	19
1.3.5 Choosing the right CVs . . . . .	20
1.4 Reconciling the two worlds . . . . .	22
<b>2 The CVHD method</b>	<b>25</b>
2.1 Objectives . . . . .	25

## CONTENTS

2.2	Description of the algorithm . . . . .	26
2.2.1	The CV . . . . .	26
2.2.2	The bias . . . . .	29
2.2.3	Critical aspects . . . . .	32
2.2.4	Related methods . . . . .	34
2.3	Applications . . . . .	35
2.3.1	General methodology . . . . .	35
2.3.2	Bond-based CV: diffusion on Cu(001) . . . . .	35
2.3.3	Dihedral-based CV: folding of a helix . . . . .	39
2.3.4	Methane dissociation on Ni(111) . . . . .	44
2.4	Conclusions . . . . .	47
<b>3</b>	<b>Advanced applications of CVHD</b>	<b>49</b>
3.1	Pyrolysis and combustion as a case study . . . . .	49
3.2	Methodology . . . . .	51
3.3	Results . . . . .	53
3.3.1	Accessible time scale . . . . .	53
3.3.2	Pyrolysis . . . . .	54
3.3.3	Oxidation . . . . .	55
3.3.4	Comparison with experiments and unbiased MD . . . . .	58
3.4	Conclusions . . . . .	60
3.5	Further applications of CVHD . . . . .	61
<b>4</b>	<b>A different approach to biasing</b>	<b>63</b>
4.1	Towards even more challenging applications . . . . .	63
4.2	A variational approach to CVHD . . . . .	64
4.2.1	Variationally enhanced sampling . . . . .	64
4.2.2	Integrating VES in CVHD . . . . .	65
4.3	Application . . . . .	67
4.4	Conclusions . . . . .	69

<b>II</b>	<b>Untangling the experiment through modeling: the case of surface charges in plasma catalysis</b>	<b>71</b>
<b>5</b>	<b>Computational models of charged surfaces</b>	<b>73</b>
5.1	The plasma-catalytic enigma . . . . .	73
5.1.1	The virtues of plasma catalysis . . . . .	73
5.1.2	Constructing models of plasma-catalyst interactions . . . . .	74
5.1.3	Understanding the effect of charging . . . . .	76
5.2	Methodology . . . . .	79
5.3	Results . . . . .	80
5.4	Conclusions . . . . .	84
<b>6</b>	<b>Catalytic effects in plasma catalysis</b>	<b>87</b>
6.1	Surface charging and plasma catalysis . . . . .	87
6.2	Methodology . . . . .	88
6.2.1	Model systems . . . . .	88
6.2.2	Modeling correct charge distributions . . . . .	89
6.3	Results . . . . .	93
6.3.1	Transition metal atom adsorption on the Al <sub>2</sub> O <sub>3</sub> support . . . . .	93
6.3.2	CO <sub>2</sub> adsorption . . . . .	94
6.3.3	Adsorption of other molecules on the support . . . . .	98
6.3.4	Impact on surface reactions . . . . .	99
6.3.5	Larger clusters . . . . .	103
6.4	Conclusions . . . . .	105
6.5	Additional information . . . . .	107
6.5.1	Transition metal adsorption . . . . .	107
6.5.2	A reference van der Waals-corrected functional . . . . .	108
6.5.3	Computational cross-checks . . . . .	110
6.5.4	Should cell neutrality always be enforced? . . . . .	112
<b>7</b>	<b>A brave new world for DFT?</b>	<b>115</b>
7.1	The density functional zoo . . . . .	115
7.2	Methodology . . . . .	117
7.3	Results . . . . .	119
7.3.1	Wide-gap material: <i>h</i> -BN . . . . .	119

## CONTENTS

7.3.2	Narrow-gap material: graphene/ <i>h</i> -BN heterojunction .	123
7.3.3	Revisiting plasma catalysis . . . . .	127
7.4	Conclusions . . . . .	127
<b>8</b>	<b>General conclusions</b>	<b>129</b>
	<b>Summary</b>	<b>135</b>
	<b>Samenvatting</b>	<b>141</b>
	<b>List of publications</b>	<b>147</b>
	<b>List of presentations</b>	<b>149</b>
	<b>Awards and grants</b>	<b>151</b>
<b>A</b>	<b>Timings &amp; time scales: a time capsule</b>	<b>153</b>
	<b>Bibliography</b>	<b>157</b>

# List of Figures

1.1	The different objectives of accelerated MD and enhanced sampling methods. . . . .	8
1.2	Schematic depiction of the four main classes of atomistic simulation methods applicable to infrequent event systems. . . . .	9
1.3	Schematic depiction of a metadynamics procedure. . . . .	18
1.4	Contrasting standard direct metadynamics and its infrequent variant. . . . .	20
2.1	Graphical representation of the key steps occurring in a CVHD simulation of a single transition. . . . .	30
2.2	Schematic depiction of the structure of the CVHD algorithms. . . . .	32
2.3	Boost factor as function of the temperature for diffusion on Cu(001). . . . .	38
2.4	Some possible states of the helix model system. . . . .	42
2.5	Number of transitions and hypertime as a function of simulation time in the helix folding test system. . . . .	42
2.6	Hypertime evolution in a dynamically biased CVHD simulation of nickel-catalyzed methane decomposition. . . . .	46
3.1	Applied maximal bias potential during the initial steps of a 1000 K CVHD pyrolysis simulation. . . . .	52
3.2	Products of the <i>n</i> -dodecane pyrolysis simulations at different temperatures. . . . .	54
3.3	CVHD simulations of <i>n</i> -dodecane oxidation. . . . .	56
3.4	Pressure-dependence of <i>n</i> -dodecane oxidation at 1000 K. . . . .	57
3.5	Comparison of experimental and simulated product distributions of <i>n</i> -dodecane pyrolysis at 1000 K. . . . .	58
3.6	Arrhenius plots of the apparent first order rate constants of <i>n</i> -dodecane pyrolysis and combustion. . . . .	60

LIST OF FIGURES

4.1	Arrhenius plot for the symmetric gas phase $S_N2$ reaction of methyl chloride. . . . .	68
5.1	Possible plasma-catalyst cross-interactions, as seen from the perspective of the catalyst. . . . .	74
5.2	Investigated $CO_2$ adsorption configurations on 2D nanomaterials.	80
5.3	Convergence dependence of $CO_2$ adsorption energies on charged $g-C_4N_3$ on the vacuum size. . . . .	81
5.4	Absolute signed errors for adsorption energies in all systems for various vacuum sizes. . . . .	82
5.5	Absolute signed errors for adsorption energies on $h$ -BN in different charge states in a fully periodic box. . . . .	83
6.1	Convergence of computed $CO_2$ adsorption energies on a negatively charged slab with respect to the position of the neutralizing countercharge. . . . .	90
6.2	Transition metal adsorption on neutral and negatively charged alumina surfaces. . . . .	92
6.3	Projected densities of states (PDOS) for Ni supported on the dry surface. . . . .	93
6.4	Effect of surface charging on $CO_2$ adsorption . . . . .	95
6.5	CVHD simulation of $CO_2$ splitting by supported metal atoms. .	99
6.6	Effect of an excess electron on the reaction energies of the $CO_2$ splitting process. . . . .	101
6.7	Contrasting adsorption behavior of $CO_2$ and CO. . . . .	102
6.8	$CO_2$ splitting on a supported $Cu_{13}$ cluster. . . . .	104
6.9	Metal adsorption configurations on the dry and hydrated surface.	107
6.10	Relative energies of the metal adsorption configurations on the dry and hydrated surface. . . . .	108
6.11	Ar dimer binding curve. . . . .	109
7.1	Graphene/ $h$ -BN heterojunction used as narrow-gap derivative of $h$ -BN. . . . .	117
7.2	Correlation between predicted gap and $CO_2$ adsorption energies on a wide-gap $h$ -BN sheet. . . . .	122

7.3	Correlation between predicted gap and charge-enhanced CO <sub>2</sub> adsorption energies on a narrow-gap graphene/ <i>h</i> -BN heterojunction. . . . .	125
7.4	Spread of predicted CO <sub>2</sub> chemisorption energies by various types of DFAs for the three different cases of charged and neutral materials considered in this work. . . . .	126



# List of Tables

2.1	Kinetic parameters for elementary diffusion processes on the Cu(001) surface. . . . .	37
2.2	Average folding times for the helix model system. . . . .	44
2.3	Average reaction time for all elementary reaction steps of the full methane dehydrogenation process $\text{CH}_4(\text{g}) \rightarrow \text{C}(\text{ad}) + 4\text{H}(\text{ad})$ , on Ni(111) at 800 K . . . . .	45
3.1	Lowest temperatures achieved in simulations of <i>n</i> -dodecane pyrolysis and combustion, and corresponding physical times and boost factors. . . . .	53
3.2	Kinetic parameters of <i>n</i> -dodecane pyrolysis and combustion as obtained from fitting apparent first order Arrhenius and Eyring equations. . . . .	59
4.1	Rate calculations for the symmetric gas phase $\text{S}_{\text{N}}2$ reaction of methyl chloride from variational CVHD . . . . .	68
6.1	Influence of surface charging on molecular adsorption energies at various sites on the $\text{Al}_2\text{O}_3$ support. . . . .	98
6.2	Charge effect on $\text{CO}_2$ binding by a supported $\text{Cu}_{13}$ cluster. . .	104
6.3	Computational consistency checks of the effect of the exchange-correlation functional. . . . .	110
6.4	Computational consistency checks for transition metal adsorption. . . . .	111
6.5	Computational consistency checks for $\text{CO}_2$ adsorption. . . . .	112
7.1	Computed gaps and charge- and field-enhanced $\text{CO}_2$ adsorption energies on the wide-gap <i>h</i> -BN sheet using different density functional approximations. . . . .	120

## LIST OF TABLES

7.2	Computed gaps and charge-enhanced CO <sub>2</sub> adsorption energies on the narrow-gap graphene/ <i>h</i> -BN heterojunction using different density functional approximations. . . . .	124
7.3	Computed adsorption energies of CO <sub>2</sub> on an Al <sub>2</sub> O <sub>3</sub> -supported Cu atom, comparing the plain PBE GGA and its hybrid HSE06 in the case of neutral and charged hydrated surfaces. . . . .	127
A.1	Time scales and computational parameters of the various MD simulation studies undertaken in this thesis. . . . .	154
A.2	SCF times and computational parameters of the various DFT calculations carried out in this thesis. . . . .	154

# Acknowledgments

It's been four years since I started my PhD project, which is quite a long time—about 15% of my life, even! I'm not the same person I was before I embarked on this journey. Even the world has changed quite a bit. Some things changed for the better, some didn't, but this PhD has most definitely been a valuable and all-round positive experience. I've learned a lot, and I believe I've made a useful scientific contribution. It wasn't an easy ride, though, and I couldn't have done it without the direct or indirect help of many individuals.

First of all, I would like to express my gratitude to my adviser, Professor Erik Neyts. He was the one who first introduced me to atomistic simulations back in my bachelor years, then gave me the opportunity to work on an interesting master's thesis project, helped me apply for a PhD fellowship, and has been my adviser in these four years. Erik, I think we've made a pretty great team: I always stubbornly insisted on doing things my way (sorry!), but anytime I'd hit a wall and realized I could use some guidance, you were an infinite supply of great ideas and useful advice. I always greatly enjoyed our conversations and discussions, which meant a lot to me. Thank you for your unceasing belief in me, and for supporting me in (most of) my crazy endeavors.

Many other people in the PLASMANT research group directly or indirectly contributed to this thesis. Of course there is the head of the group, Professor Annemie Bogaerts, who encouraged me to join the group in the first place. Her positive and enthusiastic attitude has been inspiring, and she has offered a valuable different perspective on many (scientific) questions at various stages of my academic career.

There is no doubt that office B2.28 is the most fun place in the hallway (or building? campus? university? point is, it's really fun). My office mates Stijn, Yannick, and Jonas are the greatest colleagues I could have wished for. They have become good friends as well. We started out as a haphazardly cobbled-together group of misfits thrown into a newly created office (“the special cases,”

according to Erik), but have quickly become the biggest and best group act of all. Our many activities have included political debates, general loud talking, fashion, dating advice, musical interludes, interior decoration, teaching, coffee drinking, pool parties, photo shoots, and *even* extensive scientific discussions. We've all been taking our respective PhD journeys together, and you guys have helped make it a much more bearable ride overall. Our office has been so much fun, even, that other people were basically adopted into it: Inne, Yannick (the other one), Vincent, Stijn (again, the other one), Marleen, Charlotte, Antonin, Georgi, Chris, Jeroen, Qi, Sanne, and many others have made my time in the B building much more enjoyable. Special thanks also to Damien, the first user of my CVHD method, with whom it was a pleasure to collaborate. And yes Inne, I *will* write that one erotic story I once promised you to write. (At some point. I think?)

Furthermore, it was great to regularly meet up with my former classmates at the other campus, Evelien, Robert, Julie, and Mats. Our conversations at the campus or during commutes were a great way to take away some of the daily stresses.

I would also like to thank Luc, Nelly, Ingrid, Fabiana, and Tania for their assistance with all sorts of technical and administrative hurdles. Moreover, I'm grateful for the help I received from our local supercomputing staff, Stefan and Franky, with various HPC-related issues.

I was fortunate enough to visit the group of Professor Michele Parrinello in Lugano. Michele is a truly inspiring person and a brilliant scientist. During my time in his group, I even got the opportunity to see him accept the Dreyfus Prize, yet another award in his distinguished career! It was an amazing experience to discuss research with him, and to witness him gleefully make jokes about my last name (and, simultaneously, express his love for Lucille Ball). I had a great time also thanks to the hospitality of many group members, such as Giovanni, Faidon, Daniela, Michele, Dan, Emanuele, and others. Many thanks also to Daniela for arranging all the local paperwork and being so welcoming when I finally arrived.

Professor Fabio Pietrucci was very welcoming when I paid a short visit to his group, and he gave some excellent advice, for which I'm grateful.

Of course, I was lucky to meet many more interesting people at conferences, or even through email (it's a brave new world out there!). Especially Art Voter,

Pratyush Tiwary, Gareth Tribello, Hannes Jónsson, and David Graves have impacted my thinking.

All of my activities were funded through the *Fonds Wetenschappelijk Onderzoek - Vlaanderen* (FWO), in the form of a PhD fellowship and a travel grant.

Also beyond academia I have received various kinds of support. In the music lessons by Tom and Katelijne, I could make new discoveries and be creative outside of my actual research project. Music has been a great way to keep my mind occupied while still recharging my batteries, and find other ways to express myself.

My parents and their respective partners have provided me with two safe havens. My father and Vera invariably made an effort to make me feel at home with them whenever I came by, cooked me a nice meal, and listened patiently whenever I had to get something off my chest. They always supported me in all my choices.

Theo is one of the most kind and patient people I've ever met. Perpetually prepared to help whenever I needed a hand, be it to fix up my bike (how would I otherwise get to campus?) or to give driving lessons (how would I other—eh, nevermind). Thanks to him, I now even have a whole new extra family, who have all been just as welcoming and refreshingly unpretentious like him.

Words cannot express what my mother has meant to me, not only in these PhD years, but all my life. She fought tooth and nail for me to have safe and stable home. She remained strong, even when no one else could. She always knew how to cut through bullshit and indecisiveness. Her support was relentless and unconditional. She can be efficiently practical and supremely goofy in equal measure. Without her, I wouldn't have achieved any of this. *Mama, ich habe es geschafft!*



# Model or reality?

Atomistic modeling techniques have had an enormous impact on chemistry, materials science, and molecular biology. The continuous development of electronic structure theory has made it possible to accurately describe the interactions between atoms in arbitrary environments—properties of molecules and materials can be derived directly from fundamental principles, and chemical reactions studied with atomic resolution. Such methodologies are not only limited to “static” calculations of energies or structures. Simulations based on molecular dynamics (MD) or Monte Carlo (MC) methods can even explicitly describe the atom-level dynamical evolution of matter.

It might be tempting to conclude that computational chemistry could allow one to fully reproduce or replicate reality. Such a belief—although enticing and widely held—is however fundamentally misguided. Length, time and detail in atomistic models are limited to the nanoscale, and even the smallest isolated systems can only be modeled with a finite accuracy—the best models still rely on some approximations. For most practical purposes, the mismatch in scale and complexity between experiment and atomistic model therefore prevents a perfect one-to-one correspondence. This type of exact relation is also wholly unnecessary for most scientific purposes: if simulation and experiment were to give identical answers, one of them would be superfluous! Indeed, as Anderson already noted back in 1980, one should not be blinded by the perceived accuracy of a model, if the model itself is not physically transparent:

“... a great deal of the physics is concealed inside the machinery of the technique, and that very often once one has the answers that these techniques provide, one is not exactly clear what the source of these answers is. In other words the better the machinery, the more likely it is to conceal the workings of nature, in the sense that it simply gives you the experimental answer without telling you why the experimental answer is true.”

A well-performing model is not a model that reproduces the experiment, but one that explains the experimental observations. For this reason, the very act of model construction is significantly more interesting than routine applications of one existing model to different experiments. Why does a model work? When does it fail? Are there competing approaches and, if yes, what can we learn from their differences? What can we improve? A stimulating activity indeed, and one we believe only must (somewhat) adhere to two key rules.

On one hand, a model must have a suitably limited scope. Simulations and models pare down the complexity of the real world, and unveil underlying truths precisely because of their reductive simplifications: a model can be limited to exactly the processes, molecules, mechanisms, or conditions of interest, excising any distracting real-life complexities. Indeed, the actual construction of the model can be much more revealing than any of its actual applications. Simulations are not bound by many of the experiment's restrictions—one has infinite control over all aspects of the model, and physical realities inaccessible by the experiment can be explored.

On the other hand, any practical model should still in some way intersect with reality. The model's simplifications and assumptions are only acceptable inasmuch that the model itself is still physically sound—*as simple as it can be, but not simpler*. To be of proper value, the model must make predictions that are meaningful and can in some way give insight into the actual properties of the experiment. In other words, while experiment and model do not have to overlap, it should be possible to put them side by side.

This thesis presents two case studies that attempt to rise to these challenges.

In the first part, we *strengthen the link between existing models and experiment*. Assuming that a suitably simplified atomistic simulation model exists, we intend to bring it closer to real-life conditions. In particular, we will develop algorithms to perform microscopically detailed MD simulations over *macroscopic* time scales. The goal of our novel methodology is to produce long time scale chemical transformations involving hundreds of reactions in a complex chemical environment. The improved interplay between the very different beasts that are simulation and experiment does not mean that either is turned into an exact copy of the other, but because they share the same timescale it becomes easier to exchange and combine insights from the two approaches.

The second part describes how modeling can be used to obtain a deeper in-

sight into a complicated process. As a case study, we take plasma catalysis, which consists of a highly complicated web of tightly coupled cross-interactions that are very difficult to properly characterize at a fundamental level. In order to achieve the promising potential of plasma catalysis in important 21st century challenges such as greenhouse gas conversion, a systematic insight is required. In this thesis, *a new model is constructed stepwise* to isolate a single factor of plasma catalysis—the presence of excess plasma-supplied electrons at the catalyst surface—and thoroughly assess its effect. Even for such an ostensibly simplified model, it is of paramount importance to ascertain its relevance and tangible connection to the actual experiment, and we thoroughly dive into the search of a right balance of simplicity, realism, and accuracy.

Both of these goals are rather ambitious but thankfully, we stand on the shoulders of giants, and have come a long way since the first Monte Carlo simulation of a liquid on the Los Alamos MANIAC in 1953. Many highly challenging modeling problems have already been tackled, and have led to the creation of powerful dedicated methods and algorithms. We have even reached the point where computational chemistry can give more information than is strictly necessary. Asking the right questions is becoming more important than ever, and it is useful to take a step back and take a critical look at the way the current state-of-the-art methods are applied to a problem of interest, and how approaches that have been successful in *other* fields might benefit from a new perspective. Innovation driven by this type of lateral thinking has the potential to lead to new surprising applications of existing concepts in novel disciplines. Hence, by recognizing the limitations of existing models, and identifying the untapped potential of others, a level of innovation beyond the traditional scientific method can be achieved. In this thesis, we will draw from many sources, and find new and exciting applications even for well-established models and methods.



**PART I**

Macroscopic time scales from  
microscopic simulations



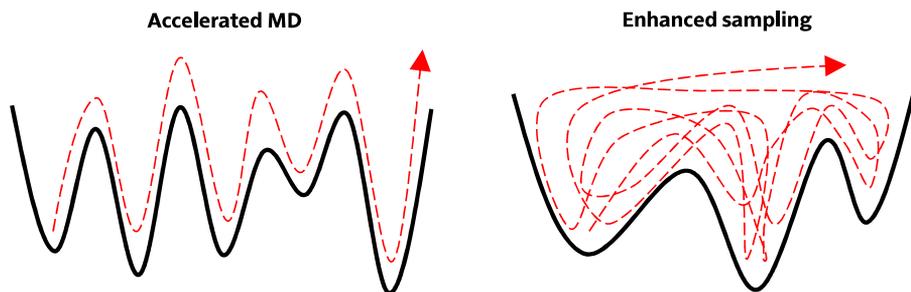
# Simulating slow processes

## 1.1 The time scale problem in atomistic simulations

The molecular dynamics method has become an indispensable tool across many branches of science and engineering. MD simulations can be used to understand the mechanisms of complex processes, calculate properties of materials, explain experimental results, and make predictions—all while providing full atomistic resolution. Such a level of detail, however, comes at a price: the time step used to discretize the equations of motion must be shorter than the fastest motion in the system, putting an upper limit to the time scales that can reasonably be addressed. In general, simulations based on (computationally cheap) classical force fields are restricted to microseconds, whereas more expensive *ab initio* methods cannot even reach beyond nanoseconds. Because many systems spend most of their time in metastable states with long lifetimes of milliseconds or more, the MD time scale is often insufficient to study common chemical processes. The system remains trapped in a single metastable state during the whole MD simulation, and not a single reaction or transition is observed.

Circumventing the MD time scale limitation is an area of active research. Broadly speaking, developments in this area can have one of two specific objectives, schematically depicted in Figure 1.1. *Accelerated MD* methods focus on accelerating the global system evolution, so that the simulation is able to produce the correct sequence of state-to-state transitions. In such an approach, the focus lies on accelerating the escape rate from metastable states while correctly reproducing the probability of all possible escape pathways. As a result, an accelerated MD method should produce a coarse-grained system evolution that is statistically indistinguishable from a (very long) standard MD simulation.

A different viewpoint is taken by *enhanced sampling* methods. Here, MD's



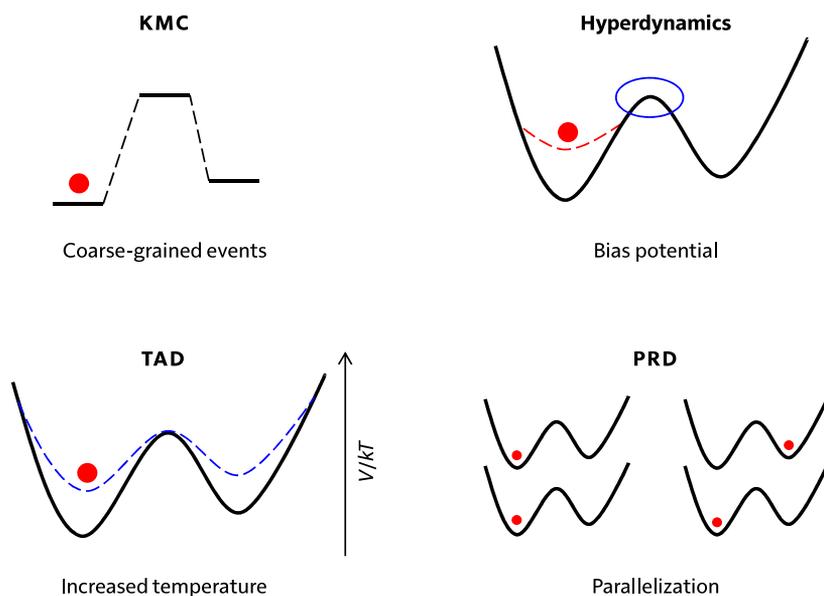
**FIGURE 1.1:** The different objectives of accelerated MD and enhanced sampling methods.

lack of ergodicity is addressed. In principle, thermodynamic properties of the system can be derived from ensemble averages obtained during an MD simulation. However, correct averaging is only possible if the simulation can exactly reproduce the probability distribution associated with the system. That is, the probability  $P_i$  to find the system in a state  $i$  with state energy  $E_i$  at a temperature  $T$  should obey  $P_i \sim e^{-\beta E_i}$ , with  $\beta = (k_B T)^{-1}$ . Correct sampling of this sort is impossible if, for example, the system remains trapped in a single metastable state during the simulation—other relevant metastable states are never reached and  $P_i = 0$  is erroneously predicted for many relevant states. Hence, enhanced sampling methods strive to provide a more balanced sampling of relevant parts of the configuration space, within the framework of MD.

## 1.2 Accelerated MD and related methods

### 1.2.1 Infrequent event systems

At the heart of all accelerated MD methods lies the concept of *events*: the system is trapped in metastable states for extended periods while it are the “events,” i.e., jumps to a different state, that are of interest. If barriers are high, most of the simulation time is spent waiting for an event to occur, and the term *infrequent event system* is used to refer to this type of situation.<sup>1</sup> In such a system, all events can be considered to be decorrelated, because the system spends such a long time in each metastable basis that it loses all memory of how it got there. An accelerated MD methods does therefore not need to provide a perfect description of the dynamics within the basin, in between events. Only the right relative probability of finding each of the possible escape paths must be preserved. The



**FIGURE 1.2:** Schematic depiction of the four main classes of atomistic simulation methods applicable to infrequent event systems. These are the KMC (kinetic Monte Carlo) method, and the three base accelerated MD methods TAD (temperature-accelerated dynamics), hyperdynamics, and PRD (parallel replica dynamics).

degree by which these escapes are accelerated is commonly referred to as the *boost factor* of a method. The true physical time scale of the infrequent event dynamics is the overall MD time scale, multiplied by the boost factor. Voter and coworkers introduced most of the key theoretical developments of infrequent event-based atomistic simulation methods, which we can divide in four main categories, sketched in Figure 1.2.

### 1.2.2 Full coarse-graining

If it is the goal to increase the occurrence of infrequent events in the simulation, the most efficient approach might be to do away with *anything but* these events. That is, if the system evolution can be purely described as the sequence of infrequent events, only these events themselves have to be considered in the model; the time evolution simply occurs as a series of jumps between local minima.

Such a coarse-grained approach to the dynamics forms the basis of the *kinetic Monte Carlo* (KMC) method. In a KMC model, the system is represented only

by minimum energy configurations (local minima) of the system, and a *catalog* of events (and associated rate constants) connecting them.<sup>2</sup> At each simulation step, a list of all possible escape routes out of the current state is extracted from the event catalog, and an escape event is selected by drawing a random number. The probability of drawing an event is made to be proportional to its rate constant  $k_i$ , and the simulation clock is advanced by drawing an escape time from the distribution  $p(t) = k_{\text{tot}} \exp(-k_{\text{tot}}t)$ , with  $k_{\text{tot}}$  overall escape rate, i.e., the sum of all possible  $k_i$ .

In its original form, the KMC approach cannot be used to predict or discover new events, because they must be defined *a priori* in the form of the event catalog. Many atomistic KMC algorithms furthermore map all atomic positions onto a lattice to simplify the storage of configurations and the identification of possible events which, however, means that only processes in crystalline solids can be modeled. By coupling the KMC algorithm to an open-ended saddle point searching method such as the dimer method<sup>3</sup> or activation-relaxation technique (ART)<sup>4</sup> various *adaptive* (or *off-lattice*) KMC methods have been devised that allow to self-consistently predict transformations of materials over long time scales without confining events to a lattice.<sup>5-9</sup>

Common saddle point searching methods as applied in adaptive KMC typically use the curvature of the PES to find transition paths and can only be used in systems where all events are fully characterized just by the minimal energy configurations of two consecutive states, and the saddle point in between. This simplification applies rather well to solids, but is inadequate once configurational entropy by low-barrier diffusional motion becomes important—the constantly changing nearest neighbor interactions in liquids or gasses cannot be included because these do not conform to the more static view of events in typical atomistic KMC methods. Evidently, systems with such “soft” modes of motion also cannot be described by lattice-based KMC.

The limited scope of widely used (atomistic) KMC implementations does not diminish the general utility and power of the KMC method, and dedicated variants can be developed for any kind of problem involving transitions between states. It does, however, illustrate the practical hurdles that prevent the development of an efficient KMC approach to arbitrary chemical transformations.

### 1.2.3 Temperature elevation

High barriers are easier to cross at elevated temperatures. In fact, the use of artificially high temperatures is a rather common trick to circumvent the MD time scale problem. However, while a high-temperature simulation can bring infrequent events within the MD time scale, it also describe a different system altogether. The high-temperature simulation will give more weight to entropically favored higher-barrier transitions. As a result, an incorrect system evolution is predicted, at least if one is actually interested in the low-temperature processes.

The *temperature-accelerated dynamics* (TAD) method<sup>10</sup> can construct an appropriate low-temperature event sequence from high-temperature simulations. This is achieved through a KMC-like handling of events. In TAD, a simulation is conducted at a high temperature  $T_{\text{hi}}$  at which events can occur within MD time scales. Each time an event is detected, its associated barrier  $E_A$  is computed using an interpolating saddle point-searching technique such as the nudged elastic band (NEB) method,<sup>11</sup> and stored along its final configuration and escape time  $t_{\text{hi}}$ . Then, the system is rolled back to the initial state and decorrelated to resume the simulation; a criterion is defined to determine when a sufficiently complete catalog of escape pathways is obtained. Once this happens, all escape times  $t_{\text{hi}}$  can be extrapolated to the target temperature  $T_{\text{lo}}$  by assuming that harmonic transition state theory (HTST)<sup>12</sup> holds, so that the relation  $t \sim e^{\beta E_A}$  can be used:

$$t_{\text{lo}} = t_{\text{hi}} e^{(\beta_{\text{lo}} - \beta_{\text{hi}}) E_A}. \quad (1.1)$$

Finally, the product state corresponding to the shortest  $t_{\text{lo}}$  is selected, and the simulation at  $T_{\text{hi}}$  is resumed from there.

For very low  $T_{\text{lo}}$ , impressive boost factors (up to  $10^9$ ) have been obtained with TAD.<sup>10,13</sup> If  $T_{\text{lo}}$  is fairly high and close to  $T_{\text{hi}}$ , the efficiency of the method decreases. There is also typically an upper limit to  $T_{\text{hi}}$  at which anharmonic effects become unacceptably large.<sup>10</sup> Due to the reliance on PES-based saddle point searching and HTST, TAD largely suffers from the same limitations that restrict KMC to processes in the solid phase. Nevertheless, TAD remains in active development,<sup>14</sup> extending the method to both larger systems and longer timescales, respectively through spatial,<sup>15,16</sup> and temporal parallelization strategies.<sup>14</sup>

### 1.2.4 Bias potentials

Events would be accelerated if the high barriers separating metastable states were lower. In the *hyperdynamics* method,<sup>17</sup> simulations are not carried out on the true potential energy surface (PES)  $V(\mathbf{R})$ , but on a modified potential  $V^*(\mathbf{R})$ , which is obtained by adding a *bias potential*  $\Delta V(\mathbf{R})$ :

$$V^*(\mathbf{R}) = V(\mathbf{R}) + \Delta V(\mathbf{R}), \quad (1.2)$$

in which  $\mathbf{R}$  denotes the system coordinates. If properly designed,  $\Delta V$  can be made to fill deep metastable basins, and lower the effective barriers for escape. If standard transition state theory (TST) is assumed to hold, and  $\Delta V$  *vanishes at all the dividing surfaces* between states, it can be proved that (1) relative rates for different escape pathways out of a basin are preserved, and (2) the boost factor can be computed as

$$\text{boost} = \langle e^{\beta \Delta V} \rangle_b, \quad (1.3)$$

in which  $\langle \dots \rangle_b$  denotes an average over the trajectory on the biased potential.

Designing an accurate, efficient and general expression for the bias potential is rather difficult due to the complexity of the underlying potential energy surface  $V(\mathbf{R})$ , which is a  $3N$ -dimensional function. In the original formulation of hyperdynamics, Voter introduced a bias potential based on the lowest eigenvalue  $\varepsilon_1$  of the Hessian, which naturally allows to detect saddle points, where  $\varepsilon_1 < 0$ .<sup>17</sup> The performance of Hessian-based approaches deteriorates with increasing systems sizes, and although more efficient schemes that avoid diagonalization of the Hessian have been proposed,<sup>18</sup> this class of bias potentials did not find widespread use.

Simplified forms of the bias potential, based on intuitive arguments rather than on strict theoretical foundations, have been quite successful. The simplest class of methods make  $\Delta V$  a direct function of the potential energy of the system: the lower the potential energy, the larger the bias should be. Variations on this approach range from “flat” potentials, setting  $V^* = E$  if  $V$  is below some threshold  $E$ ,<sup>19</sup> to somewhat more elaborate forms such as

$$\Delta V(\mathbf{R}) = \frac{(E - V(\mathbf{R}))^2}{\alpha + (E - V(\mathbf{R}))} \quad \text{if } V(\mathbf{R}) < E, \quad (1.4)$$

used by the rather unfortunately-named “accelerated MD” (aMD) method.<sup>20,21</sup> In all cases, potential energy-based methods only work well

on a simple and smooth PES, where saddle points can be identified purely by the value of  $V(\mathbf{R})$ . Once the thermal contribution to the potential energy ( $\sim 3Nk_B T/2$ ) or the energy difference between states become larger than the smallest barrier, the boost factor becomes negligible, or some accuracy must be sacrificed. Larger systems can be treated by using localized energies, such as the potential energies of single atoms<sup>22</sup> or, as common in aMD, dihedral energy terms in proteins.<sup>20</sup> Nevertheless, with typically applied large boost potentials (in the order of hundreds of kcal/mol in aMD) energy-based biasing methods will always corrupt the transition state, and cannot form the basis of a formally correct hyperdynamics method—in spite of its name, aMD is actually more an enhanced sampling method.

If one has some idea of the mechanism of the to-be-accelerated process, a specifically-tailored bias potential can be developed. Recognizing that solid-state processes typically involve bond breaking, Miron & Fichthorn proposed the *Bond Boost* method, where the bias potential is an explicit function of the  $N_b$  bond lengths  $r_i$  in the system.<sup>23</sup> The Bond Boost bias is designed in such a way that it is largely controlled by the bonds that are compressed or stretched the most relative to the equilibrium distance  $r_i^{\text{eq}}$ , because those are the most likely ones to be involved in a reaction:

$$\Delta V = A(\varepsilon_1, \dots, \varepsilon_{N_b}) \frac{\Delta V^{\text{max}}}{N_b} \sum_i^{N_b} \left[ 1 - \left( \frac{\varepsilon_i}{q} \right)^2 \right] \quad \text{with} \quad \varepsilon_i = \frac{r_i - r_i^{\text{eq}}}{r_i^{\text{eq}}}. \quad (1.5)$$

The bias is multiplied by a so-called “envelope function”  $A$  that ensures the bias is smoothly turned off if  $\varepsilon_{\text{max}} > q$ , before the saddle points are reached. For simple solids, the Bond Boost method requires little parametrization and is cheap to evaluate, and has been applied to various film growth processes.<sup>24–26</sup> Its main limitations are the fact that the bias potential is a “hard-coded” function of bonds, so no other types of mechanisms can be accelerated and, like all hyperdynamics methods, the strength of the bias must be pre-defined (through  $\Delta V^{\text{max}}$ ). The method has a rather narrow time spectrum and will only be efficient for systems in which all barriers are of similar height.

### 1.2.5 Parallelization

The most accurate accelerated MD method is *parallel replica dynamics* (PRD): the only assumption made is that of an infrequent event system obeying first-order kinetics.<sup>27</sup> In PRD, multiple statistically independent replicas of the same

system state are simulated, each on different processors until an event is detected on any replica. Then, the simulation clock is advanced by the total waiting time on all replicas, and the newly discovered state is communicated to all other replicas to initiate a new cycle. Because recent advances in high performance computing have been primarily driven by increased parallelization (rather than higher clock speeds of individual processors), PRD and its variants are an efficient way to harness the abilities of modern hardware resources.<sup>28,29</sup> This fact—and the high generality of the method—makes PRD in many ways the most attractive accelerated MD method. However, PRD will only be feasible if each replica can actually be run on a small number of processors: *ab initio* methods or large classical simulations have much more extensive hardware demands for their force evaluation alone, so that a single replica will already take up tens or hundreds of processors, leaving only little room for additional replicas.

## 1.3 Enhanced sampling: metadynamics

### 1.3.1 Sampling in collective variable space

Because an enhanced sampling method aims to improve the overall sampling of (a region of) the configuration space, a mere event-based breakdown of the dynamics will be insufficient. Of course, analysis of all  $3N$  degrees of freedom of the system will just as well give little physical insight—it becomes impossible to see the forest for the trees once the system contains more than a few atoms. A compromise is given by *collective variables*: a collective variable  $s$  is a function of the system coordinates  $\mathbf{R}$  that captures a relevant aspect of the physics of a problem of interest, and allows to extract physically meaningful insights from a complicated (computer-generated) trajectory. A proper definition of  $s(\mathbf{R})$  is therefore a crucial aspect of representing human intuition and reasoning as machine-readable (and computable) concepts, and *vice versa*. After all, development of better and faster simulation techniques will only make sense if we are able to represent our ideas and results in a meaningful way.

If a small number of CVs  $\mathbf{s} = (s_1(\mathbf{R}), s_2(\mathbf{R}), \dots, s_n(\mathbf{R}))$  is defined, the probability distribution of the CVs can be defined as

$$P(\mathbf{s}) = \int \delta[\mathbf{s} - \mathbf{s}(\mathbf{R})] P(\mathbf{R}) d\mathbf{R}, \quad (1.6)$$

with  $P(\mathbf{R})$  the Boltzmann distribution

$$P(\mathbf{R}) = \frac{e^{-\beta V(\mathbf{R})}}{\int e^{-\beta V(\mathbf{R})} d\mathbf{R}}. \quad (1.7)$$

Hence, all “irrelevant” degrees of freedom are integrated out, and we can focus on the few variables that matter to us.

One particularly interesting property that can be calculated once the distribution  $P(\mathbf{s})$  is defined, is the *free energy surface* (FES):

$$F(\mathbf{s}) = -\frac{1}{\beta} \ln P(\mathbf{s}). \quad (1.8)$$

The FES is typically a smoother function than the rugged, high-dimensional PES, and therefore much easier to analyze. For any process that is well-characterized by the chosen CVs, the nature and relative stabilities of states, and the barriers separating them, can be directly inferred from the FES, with all other degrees of freedom folded into these properties as averaged background effects. For example, a chemical reaction in solution must be treated with full atomistic resolution in order to correctly sample all possible configurations of the surrounding solvent and its interaction with the reactants. However, to identify educt, product, and intermediate states, only the broken and/or formed bonds must be tracked—just a few physically sensible variables! In this sense, the introduction of CVs and the FES can reconcile large-scale MD simulations with conventional chemical intuition. Indeed, such a reduced picture is also used in most chemistry textbooks: only the reacting molecules are depicted, and all other degrees of freedom are implicitly discussed as averaged steric, conformational, or solvation effects.

An enormous number of CVs has been reported in the literature, from simple geometric parameters such as distances, angles or torsions to more elaborate functions such as crystal order parameters, hydrogen bond patterns, protein structure descriptors, and many more, which are implemented in reference codes such as PLUMED<sup>30,31</sup> and Colvars.<sup>32</sup> In essence, identification of appropriate CVs is equivalent to definition of events in accelerated MD simulations because here, too, a simplified representation is used to analyze the state of the system. Of course, event detection is much cruder than using actual CVs—no continuous variables are required. Less sophisticated approaches are therefore

typically applied: usually, simple position-based descriptors suffice to identify events in the solid phase.

### 1.3.2 How compute the free energy?

A direct evaluation of (1.6) through MD simulation simply entails computing this property as a running ensemble average  $\langle \delta [\mathbf{s} - \mathbf{s}(\mathbf{R})] \rangle$ , and obtaining the FES through (1.8). The reason why enhanced sampling methods exist is of course the lack of ergodicity of MD in most systems, as a consequence of the time scale problem: typically only a limited part of the CV space is properly sampled. Various techniques have been developed to improve exploration of CV space, and some key examples are briefly summarized here. It is interesting to note that most of these methods in some aspect resemble one of the accelerated MD methods—all of them either use bias potentials, temperature elevation, or parallelization across multiple replicas—a fact that we will exploit later on.

As outlined above, the construction of a FES typically requires the definition of a few CVs  $\mathbf{s}$ . Some enhanced sampling methods do not depend on  $\mathbf{s}$  during the simulation and only use them for *post factum* analysis, while other techniques attempt to explicitly improve sampling along them.

The most famous member of the former class is the *replica exchange* MD method, which exploits some of the same ideas underpinning TAD and PRD for enhanced sampling.<sup>33,34</sup> A replica exchange simulation consists of several replicas  $i$  with different temperatures  $T_i$  that are independently simulated in parallel. High-temperature replicas can more easily escape metastable states and provide better sampling, while the lowest-temperature replica has the desired temperature and is used for the actual computation of the FES. At fixed steps during the simulation, an attempt is made to swap system coordinates of two replicas  $i$  and  $j$ , accepted with the probability

$$\alpha = \min \left( 1, e^{(\beta_j - \beta_i)(V(\mathbf{R}_j) - V(\mathbf{R}_i))} \right), \quad (1.9)$$

so that the low-temperature replica ends up benefiting from the enhanced sampling in the higher-temperature replicas.

In principle, the aMD method also belongs to the class of CV-free enhanced sampling methods: although the sampling bias always depends on the potential energy, which is a CV, the FES is reconstructed as a function of different CVs specifically chosen for the problem at hand.

The *temperature-accelerated MD* (TAMD) method explicitly depends on the CVs of interest. TAMD simulates an extended system in which the equations of motion of the CVs are evolved at a temperature  $T_{\text{hi}}$  that is higher than the system temperature  $T$ . That way, the probability distribution of the CVs follows  $e^{-\beta_{\text{hi}}F(\mathbf{s},T)}$  — rather than  $e^{-\beta F(\mathbf{s},T)}$  — and sampling along the CVs is improved.<sup>35,36</sup> The distribution  $P(\mathbf{s})$  is hence effectively flattened, the degree of which is controlled by the ratio  $T_{\text{hi}}/T$ , in a manner similar to well-tempered metadynamics (*vide infra*).

Many CV-based free energy sampling techniques are based on a bias potential, and *umbrella sampling* is historically the first of these methods.<sup>37</sup> In umbrella sampling, a bias potential is added to confine the system to a small section of CV space, in which extensive sampling can be carried out. Separate simulations are performed in other (partially overlapping) parts of CVs space, after which all windows are reweighed and combined to produce the unbiased distribution.

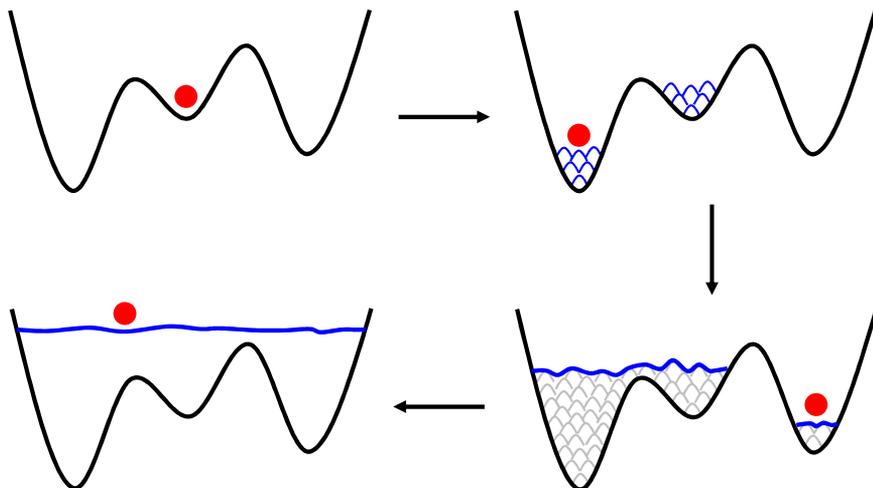
As the name implies, the *adaptive biasing force* (ABF) method is force-based rather than potential-based.<sup>38–40</sup> ABF calculates the average force along  $\mathbf{s}$  in different bins along CV space, and iteratively constructs an estimate of  $\nabla_{\mathbf{s}} F(\mathbf{s})$  which, if applied as an external biasing force, exactly compensates the average force along the CVs, and hence effectively flattens the FES. Once an accurate estimate of the average gradient of  $F(\mathbf{s})$  is known, ABF efficiently samples the full FES, which can be reconstructed through integration of  $\nabla_{\mathbf{s}} F(\mathbf{s})$ .

### 1.3.3 Metadynamics

Most of this work, however, is primarily based on the *metadynamics* method.<sup>41</sup> In metadynamics, a *history-dependent* bias potential is generated during the simulation as a sum of periodically deposited repulsive Gaussians of width  $\sigma$ .

$$\Delta V(\mathbf{s}) = \sum_i^{\text{hills}} w_i \exp \left[ - \sum_j^{\text{CVs}} \frac{(s_j - s_j(t_i))^2}{2\sigma_j^2} \right]. \quad (1.10)$$

Between deposition times  $t_i$ , the simulation is propagated as an MD simulation on the biased potential. In practice, this formulation means that whenever the system ends up in a metastable state, the metadynamics simulation will continue to fill up the basin until the system escapes. After some time (formally  $t \rightarrow \infty$ ), the whole underlying FES has been compensated by the deposited



**FIGURE 1.3:** Schematic depiction of a metadynamics procedure. Initially, the system is trapped in some metastable state, and several other states of relevance would remain unexplored. Bias is therefore slowly added to the system, pushing it over a barrier to another metastable state. After some time, the whole FES is visited, which is eventually fully compensated by the accumulated bias.

$\Delta V(\mathbf{s})$ , so that  $P(\mathbf{s})$  becomes uniform, which means that

$$\Delta V(\mathbf{s}) = -F(\mathbf{s}), \quad (1.11)$$

up to an irrelevant additive constant. This way, metadynamics does not only improve the sampling in collective variable space, but also produces the reconstructed FES as a byproduct. The general evolution of the bias during a metadynamics simulation is illustrated in Figure 1.3.

The original formulation of metadynamics attempts to improve sampling by completely flattening the FES (or, equivalently,  $P(\mathbf{s})$ ). A less aggressive variation on the concept is *well-tempered metadynamics* (WT-metaD),<sup>42</sup> in which the bias is given by

$$\Delta V(\mathbf{s}) = -\left(1 - \frac{1}{\gamma}\right)F(\mathbf{s}). \quad (1.12)$$

The bias factor  $\gamma$  controls the strength of bias potential, and can be viewed as the scaling factor by which the FES is flattened. In other words, the probability distribution of the CVs is made proportional to  $[P(\mathbf{s})]^{1/\gamma}$ , and fluctuations in

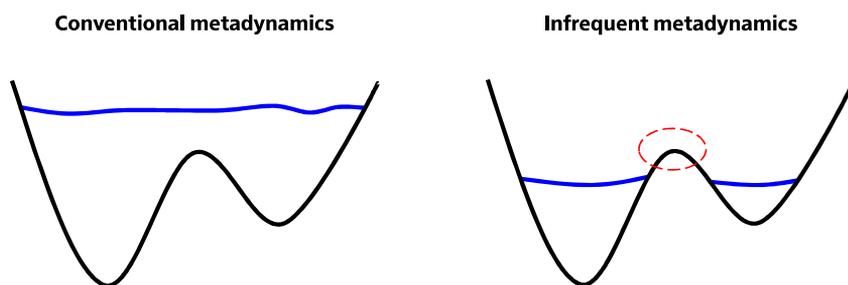
the system are enhanced if  $\gamma$  is increased.  $\gamma$  can also be interpreted as a scaling factor of the sampling temperature, so that sampling is enhanced as if carried out at higher temperature  $\gamma T = T + \Delta T$ , in which  $\Delta T$  is the bias temperature. A practical realization of WT-metaD relies on this concept of bias temperature. Instead of constructing the bias of (1.10) using hills of constant height  $w_i = w$ , it is determined upon deposition as a function of the instantaneous bias, according to  $w_i = w e^{-\Delta V(s(t_i))/k_B \Delta T}$ . Standard metadynamics does not fully converge to the actual FES, but fluctuates around the correct value as a function of the deposition parameters.<sup>43</sup> In contrast, it can be proved that the smoothly decaying hills in WT-metaD add up to a converged bias potential, and an equally converged FES can be recovered though (1.12).<sup>44</sup> Note that for  $\gamma \rightarrow \infty$ , original metadynamics is recovered, whereas setting  $\gamma = 1$  represents unbiased MD.

Continued development has expanded the application domain of metadynamics to fields as diverse as chemistry, biology, materials science, and statistical mechanics.<sup>45–48</sup>

#### 1.3.4 Dynamics from metadynamics

Conceptually, adding bias potentials to enhance the escape rate from metastable states is rather similar to hyperdynamics, but because metadynamics aims to fully fill the FES, it does not fulfill the hyperdynamics requirement that  $\Delta V = 0$  at dividing surfaces between states. Tiwary & Parrinello recognized that this restriction can be more or less enforced if the bias deposition rate is sufficiently low.<sup>49</sup> After all, the time spent by the system at the dividing surfaces is very short, so the probability that bias is deposited in these regions can be made arbitrarily small (Figure 1.4). This approach, dubbed the *infrequent metadynamics* method, allows to compute transition rates between metastable states by combining the history-dependent metadynamics bias potential (1.10) with the hyperdynamics-style time scale reweighing of (1.3). Spurious bias deposition in the transition state regions is not strictly avoided—to verify the quality of the calculated rates, an *a posteriori* check must be performed. If the obtained distribution of escape times is that of a Poisson process, it can be assumed that the rescaled dynamics has not been corrupted.<sup>50</sup>

In contrast to hyperdynamics, that focuses on generating trajectories through the full configuration space, infrequent metadynamics is still a metadynamics method, confined in collective variable space. The objective of the method is



**FIGURE 1.4:** Contrasting standard direct metadynamics and its infrequent variant.

to generate many transitions between a small number of metastable states that are distinguishable by the chosen set of CVs  $s$ , and it has been successfully applied to challenging processes such as protein-ligand unbinding in explicit solvent.<sup>51</sup> Furthermore, chemical reaction rates could be obtained from *ab initio* force fields, an application that however also exposes a weakness of the infrequent metadynamics approach: the requirement to use very slow bias deposition rates, with intervals larger than 1 ps, is difficult to combine with expensive quantum chemical calculations that limit the overall time scale of the simulation to only a few tens of ps.<sup>52</sup>

### 1.3.5 Choosing the right CVs

The accuracy and efficiency of any metadynamics simulation is largely determined by the appropriateness of the chosen CVs. The key issue is that the CVs should not only be able to distinguish between all states and pathways of interest, but also have to contain all slow modes in the system. Because the bias potential is meant to accelerate sampling along the slow modes, it should in some way be a function of them—otherwise the barriers associated with these modes are hidden to the algorithm.<sup>46</sup> If this is not the case, hysteretic behavior can be observed, and accurate sampling of the underlying FES is not possible.

CV selection requires the user to have a good *a priori* idea of the relevant mechanisms and their associated slow modes. In many cases, finding a suitable set of CVs is therefore the most challenging aspect of applying metadynamics. One practical way to sidestep this complication is given by *path collective variables*: if some initial guess of the mechanism can be made—in the form of a number of configurations that describe the transformation of interest—a reference path can be interpolated.<sup>53</sup> Biasing is performed on two collective vari-

ables, one of which is a projection of the system’s configuration on the reference path, while the other measures how much the system is straying from it. As such, path CVs are able to naturally incorporate states and pathways that are not covered by the initial guess of the mechanism.

It can be argued that all employed CVs for biasing are supposed to have an actual physical meaning, because only then meaningful insight into the dynamics of the system can be obtained. Therefore, one preferably should try to derive CVs that represent specific motions of the system, rather than relying on brute force approaches such as path CVs that might allow for correct biasing, but have no direct correlation with the actual physics taking place. Two approaches have been recently presented that attempt to identify slow modes in the system and construct appropriate CVs based on physical descriptors of the dynamics. Both methods, *spectral gap optimization of order parameters* (SGOOP)<sup>54</sup> and the *variational conformational dynamics approach to metadynamics* (VAC-metaD),<sup>55</sup> entail performing an initial WT-metaD simulation with a suboptimal CV that nevertheless suffices to (crudely) sample the transition of interest. Then, the biased trajectory is post-processed and improved CVs are obtained as a linear combination of a large set of possibly relevant, simpler CVs.

In SGOOP, a single CV is used for biasing: a linear combination of  $d$  CVs

$$f = c_1 s_1 + c_2 s_2 + \dots + c_d s_d, \quad (1.13)$$

in which the initial coefficients  $\{c_i\}$  are trial values. Through post-processing of a biased trajectory on this trial CV, improved  $\{c_i\}$  are obtained that maximize the spectral gap of  $f$ , i.e., produce the CV that is the best representation of the slowest modes in the system.

The initial CVs used in VAC-metaD can be just a subset of a larger number of candidate CVs. First, a biased trajectory is obtained with a few simple CVs, after which the observed conformational dynamics of the system is expanded as a linear combination of the full set of candidate CVs. More precisely, the full CV set is a basis set in which the approximate eigenvectors of the dynamical propagator of the system can be expanded; the eigenvectors corresponding to the largest eigenvalues are approximate representations of the slowest modes, and will be optimal CVs.

Besides leading to improved sampling, this class of methods also opens new avenues to improved physical insight. Using SGOOP, Tiwary & Berne analyzed

host–ligand unbinding in water to assess the relative importance of the water–solute interaction relative to the (strong) host–ligand interactions—and found that an optimal reaction coordinate has to contain both, in contrast to what is typically assumed.<sup>56</sup> With `VAC-metaD`, Piccini *et al.* demonstrated that even for ostensibly simple chemical reactions, slow modes do not only correspond to bond breaking or formation: conformational degrees of freedom are also part of the reaction coordinate, and the relative contributions of these motions are temperature-dependent.<sup>57</sup> The prospect of being able to simultaneously improve sampling *and* gain insight into mechanistic aspects of microscopic processes is very exciting, and will likely have a large impact on the field of molecular simulation. Very recently, new methods were presented to generate CVs by deep neural network processing of simulation trajectories,<sup>58</sup> or by analysis of fluctuations in the metastable state, i.e., without having to sample any transition or even guess a reaction path *a priori*.<sup>59</sup>

## 1.4 Reconciling the two worlds

It is clear that the two families of long time scale atomistic simulation techniques are based on shared theoretical concepts, and have largely similar objectives. Yet, in many ways both approaches also seem to live on two entirely different planes of existence.

In the world of accelerated MD, highly sophisticated approaches have been devised to generate nonequilibrium trajectories as efficiently as possible. In most cases, though, these algorithms are either tied to the solid state, require systems in which events have mechanisms involving few atoms, or assume that interatomic interactions can be described by fairly cheap force fields.

The application domain of collective variable-based enhanced sampling methods is enormous and spans the physical, chemical, and biological sciences. CV spaces of many forms can be efficiently explored. Within the single framework of the metadynamics formalism one can reconstruct the FES, disentangle transition mechanisms, and compute reaction rates of activated processes. The only missing family member is a generic approach to discover reaction pathways, and generate trajectories through configuration space of entire reaction systems. Of course, the foundational concept of the CV space is a deliberate *reduction* of the degrees of freedom, which is antithetical to letting the simulation

off the leash.

In this work, we attempt to combine the strengths of accelerated MD, and enhanced sampling methods. Specifically, we wish to develop a hyperdynamics method that has some of the flexibility of metadynamics. Primarily, the method should use an adaptive bias that does not need to be determined *a priori*, and the ability to be applied to a much wider range of systems without having to modify the formalism. Such an algorithm could also be seen as a metadynamics variant—one that extends the family with the ability to discover new pathways.



## CHAPTER 2

# The CVHD method

### 2.1 Objectives

We wish to develop an accelerated MD method that retains the key strength of hyperdynamics (large boosts), but cures its main weakness (need for system-specific optimization of the bias potential function). Therefore, the method should possess the following properties.

A CV-based algorithm. As demonstrated by the tremendous flexibility of methods such as metadynamics, CVs form an excellent basis for a truly generic biasing strategy. A clear separation between the core biasing algorithm and any system-specific details would be desirable.

CVs that do not impose any specific mechanism or pathway on the system.

Most standard CVs are designed to describe one specific kind of motion. One can either use detailed CVs to specifically describe the dynamics of small set of atoms (such as coordination number), or variables that capture the averaged behavior of a large group of atoms (e.g., the mean squared displacement (MSD)). We, however, wish to bias all atoms in a large system, but still have the flexibility to follow any of the accessible detailed chemical pathways. No *a priori* definition of the pathway or mechanism should be required.

The ability to explore the full configuration space. Standard metadynamics requires that the employed set of CVs is able to discriminate between

---

The results presented in this chapter were published in:

K. M. Bal and E. C. Neyts. Merging metadynamics into hyperdynamics: accelerated molecular simulations reaching time scales from microseconds to seconds. *J. Chem. Theory Comput.* **11**, 4545 (2015).

all possible states that are encountered in the simulation. It is our intention to *discover* new pathways and predict product states, which is not possible with a small set of traditional CVs, which assume that these are all already known to some degree.

An always-appropriate bias potential. Because the simulation will end up in many different states of various kinds of systems, a suitably flexible bias potential is required. In particular, the pre-defined shape and strength of the bias function has limited other hyperdynamics implementations to a specific subset of processes in a single system. Our method should be able to have some sort of self-learning ability to tailor its bias to the needs of the system at each stage of the simulation, without user intervention.

To meet these goals, we will have to draw from both the accelerated MD world, as from the field of enhanced sampling methods. The resulting algorithm is dubbed the *collective variable-driven hyperdynamics* (CVHD) method.

## 2.2 Description of the algorithm

### 2.2.1 The CV

Interestingly, a class of CVs that appears to satisfy our demands already exists. Based on graph theory, the social permutation invariant (SPRINT) coordinates of Pietrucci & Andreoni define a unique topology for any possible configuration of all  $N$  atoms in a system.<sup>60</sup> The eigenvector corresponding to the highest eigenvalue of the  $N \times N$  contact matrix gives a set of  $N$  atomic CVs that suffice to identify any possible bonding configuration. We deem SPRINT coordinates unsuitable for our purposes because for large systems, the number of CVs to bias becomes very large, while the performance of metadynamics deteriorates exponentially with the number of CVs. Indeed, SPRINT coordinates have only been applied to chemical transformations in fairly small systems ( $N \approx 10$ ) with very aggressive biasing.<sup>61</sup> Furthermore, it has been claimed that even for small molecules, SPRINT CVs cannot always properly distinguish all possible isomers.<sup>62</sup>

In the CVHD approach, we attempt to devise a class of CVs that avoid the pitfalls of SPRINT coordinates, but have a similar or even greater flexibility. The CV we use here is inspired conceptually by the Bond Boost method,

which is known to be able to handle large systems, and implementation-wise by a functional form introduced by Tiwary & van de Walle in their SISYPHUS method.<sup>63</sup> We assume there exists a (large, possibly redundant) set of *local CVs*  $\mathbf{s} = (s_1, s_2, \dots, s_n)$  that can distinguish between all available escape paths out of a reference state  $\mathbf{s}^{\text{ref}}$ . For each local CV  $s_i$ , a *local distortion*  $\chi_i = \chi(s_i)$  can now be defined, which is a continuous function that can return values between 0 and 1.  $\chi$  must be designed in such a way that if a CV  $s_i$  is directly involved in a transition, and takes a value  $s_i^\ddagger$  at the corresponding dividing surface,  $\chi(s_i^\ddagger) = 1$ . If  $s_i$  is far enough from  $s_i^\ddagger$  and closer to its equilibrium value,  $\chi_i < 1$ . Whenever any local CV  $s_i$  is involved in a transition somewhere in the system (and hence  $\chi_i = 1$ ), the system as a whole is about to cross a dividing surface, and the global CV describing the full system must reflect this. For this purpose, we calculate the *global distortion*  $\chi_t$  as:

$$\chi_t = \left( \sum_i^n \chi_i^p \right)^{1/p}, \quad (2.1)$$

in which  $p > 1$ . While SISYPHUS does not need continuous and vanishing derivatives at both “edges” of the CV, this is required in a hyperdynamics implementation like CVHD. Therefore, the actual CV  $\eta$  is here calculated from  $\chi_t$  according to the Tersoff-style<sup>64</sup> cutoff function

$$\eta = \begin{cases} \frac{1}{2} (1 - \cos(\pi\chi_t^2)) & \text{if } \chi_t \leq 1 \\ 1 & \text{if } \chi_t > 1 \end{cases}. \quad (2.2)$$

In short, the exponent  $p$  is used in the calculation of  $\chi_t$  to ensure that large distortions make a larger contribution to  $\eta$  than small ones. As a result, most of the bias energy flows into CVs that are about to contribute directly to a transition (modulated by the magnitude of  $p$ ), similar to what the Bond Boost method does with the breaking of bonds. Hence, the global CVs  $\chi_t$  or  $\eta$  can be used to selectively boost large changes in small parts of the included configuration space, insensitive to smaller fluctuations in the other parts. No biasing forces are applied to the system when  $\chi_t > 1$ : if  $\chi(s)$  is properly designed, this means that no bias is added in a transition state region. Also, the specific form of  $\chi_t$  does not only allow to describe transitions involving single degrees of freedom ( $\chi_t = 1$  because a single  $\chi_i = 1$ ) but also the simultaneous distortion of multiple local CVs (e.g.,  $N$  CVs all having only  $\chi_i = (1/N)^{1/p}$  such that  $\chi_t = 1$ ). The precise value

of  $p$  is not a critical parameter determining the accuracy of the CVHD method. Rather, it only controls to what extent the bias energy is distributed across the system: large values of  $p$  will suppress the influence of smaller distortions on  $\chi_t$  and will thus lead to a more localized bias potential. To optimize the method,  $p$  can be varied in function of how “concerted” events are expected to be, but its effect was found to be rather small for values between 4 and 12.

As presented here, a CV like  $\eta$  will only be of use as long as the system remains in the state described by  $\mathbf{s}^{\text{ref}}$ . Once the system escapes from this basin, and ends up in a different state,  $\eta$  will have a constant value and no physical meaning anymore. In order to be able to describe multiple consecutive events in a single biased MD run, without user intervention, a criterion is added to “reset” the procedure: if  $\eta = 1$  over a waiting time  $t_w$ , the system is assumed to have undergone a transition, and thermalized in its new state. Then, a new set of  $\mathbf{s}^{\text{ref}}$  is chosen, based on the current state, and the accelerated MD procedure is resumed. Hence, by slightly bending the rules, a single CV can be used to (1) give the system full freedom to escape out of a basin and (2) repeat this feat for each subsequent state.

A practical example of a local CV  $s_i$  is the stretch of a bond, which can be used to study reactive events involving bond breaking. Here, it is assumed that for every bond pair  $i$  with length  $r_i$ , there are distances  $r_i^{\text{min}}$  and  $r_i^{\text{max}}$ , which mark the begin and end point of possible reactive events. If  $r_i < r_i^{\text{min}}$ , the bond is not likely to dissociate soon, and not biased, whereas if  $r_i = r_i^{\text{max}}$ , the bond is about to dissociate and the system is close to a dividing surface. When the simulation starts, a list of bonds is created from all atomic pairs that are shorter apart than a cutoff  $r_i^{\text{cut}}$ . Then, local distortions can be calculated as

$$\chi_i = \begin{cases} 0 & \text{if } r_i < r_i^{\text{min}} \\ \frac{r_i - r_i^{\text{min}}}{r_i^{\text{max}} - r_i^{\text{min}}} & \text{if } r_i^{\text{min}} \leq r_i \leq r_i^{\text{max}} \\ 1 & \text{if } r_i > r_i^{\text{max}} \end{cases} . \quad (2.3)$$

Given the relative simplicity of this implementation of  $\eta$ , which only has a few parameters, we expect that it is rather widely applicable. As it only requires the slow to-be-boosted process to involve bond breaking, the CV can be readily applied to a wide variety of processes. Finally, as explained above, the basic functional form (2.3) could also be applied to local CVs other than bond lengths

to study a wide variety of other processes. Indeed, as long as the slow events of interest involve a significant distortion in a small subset of  $\mathbf{s}$ , the CVHD algorithm will be efficient. In the given example, we use bond distances to construct the global distortion function  $\chi_t$ , but the basis set of local CVs could also be something else, as we will show in the following section. The formalism as outlined in this section is sufficiently general and flexible to be used as a starting point for the development of such new CVs, without having to modify the full boosting algorithm.

Some time after the initial derivation of the CVHD CV, it was recognized that an even more general and flexible formulation of  $\eta$  exists. Instead of directly encoding transition state safety into the definition of  $\chi$ , this role can be transferred to  $\eta$ . For example, again using the example of bonds lengths as local CVs, one could write:

$$\chi_i = \begin{cases} 0 & \text{if } r_i < r_i^{\text{ref}} \\ \frac{r_i - r_i^{\text{ref}}}{r_i^{\text{ref}}} & \text{if } r_i \geq r_i^{\text{ref}} \end{cases} \quad (2.4)$$

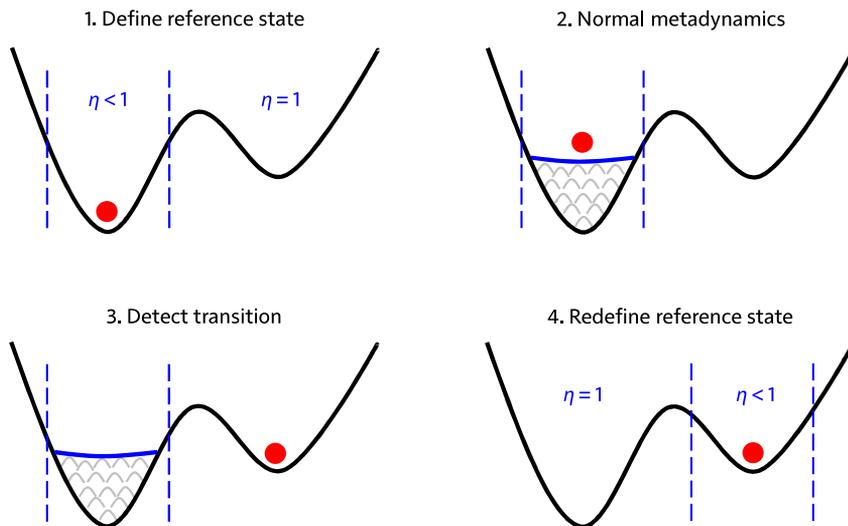
and

$$\eta = \begin{cases} \frac{1}{2} \left( 1 - \cos \pi \left( \frac{\chi_t}{\chi^{\text{cut}}} \right)^2 \right) & \text{if } \chi_t \leq \chi^{\text{cut}} \\ 1 & \text{if } \chi_t > \chi^{\text{cut}} \end{cases}. \quad (2.5)$$

Such a formulation has the advantage that  $\chi_t$  behaves more like a standard CV—one that measures the distance from some reference state. Ensuring that transition states remain uncorrupted is then relegated to the proper choice of  $\chi^{\text{cut}}$ .

### 2.2.2 The bias

If  $\eta$  is now used as a CV in a metadynamics simulation, a bias potential  $\Delta V(\eta)$  will be a slowly “grown” function of the form (1.10). The nature of the metadynamics method ensures that this bias potential matches the underlying FES of the studied processes and should thus guarantee a safe and efficient bias. An important difference between the CVHD method and metadynamics, however, is that the bias deposited in a state  $A$  is deleted once a new state is reached: even when state  $A$  is visited again later, the bias potential must be generated again. This is a consequence of the fact that the CV  $\eta$  cannot distinguish all possible states in the system, like in conventional (infrequent) metadynamics, but is only required to identify the system’s position in the *current* state. Note that this also



**FIGURE 2.1:** Graphical representation of the key steps occurring in a CVHD simulation of a single transition. At any time, the CV can only properly resolve the current reference state. Whenever a transition to a new state is detected, a new CV will be defined and the simulation forgets the old state.

explains the use of a criterion based on  $t_m$  to determine whether a new state has been reached, rather than having this explicitly reflected by the CV value. The CVHD method thus gains flexibility in the size of the configuration space that can be explored, but at the same time loses some efficiency. All in all, CVHD is metadynamics paired with a dynamical redefinition of its CV, as summarized in Figure 2.1.

The use of CVs is not limited to metadynamics, and could also be the basis of a traditional “static” hyperdynamics potential.  $\eta$  can satisfy the important constraint that no bias can be added in the transition state region, provided that a proper form of  $\chi$  is selected (for instance,  $r_i^{\max}$  in equation (2.3) is made sufficiently small). This means that a simple linear function of  $\eta$  could be an appropriate bias potential:

$$\Delta V(\eta) = \Delta V^{\max}(1 - \eta), \quad (2.6)$$

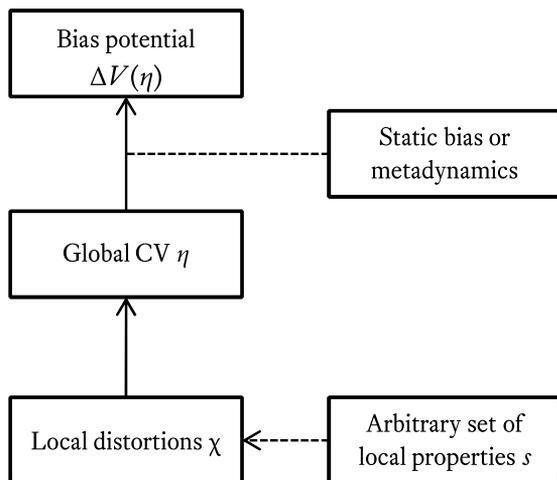
in which  $\Delta V^{\max}$  is the maximal bias strength. Just as is the case for traditional hyperdynamics implementations,  $\Delta V^{\max}$  must be chosen appropriately: large enough to obtain a substantial boost, but not larger than the barriers of interest.

Note that this static approach requires some *a priori* knowledge of the possible events in the system, in contrast to the dynamic application (see below). If bond lengths are used as local property, CVHD with a static bias essentially reduces to the Bond Boost method (equation (1.5)).

The expression for the bias potential itself should preferably not be specifically tailored for a particular system or process. Rather, in the spirit of other CV-based methods, all the complexity of the investigated process should be abstracted by the proper choice of a CV. As a result, the simple linear function of equation (2.6) is sufficient. As such, the CVHD family of methods is highly flexible. Both the underlying local property (as implemented by its associated distortion function) as the bias function itself can be changed to fit the needs of the studied process. The only characteristic that is shared by all (sub)methods is the CV  $\eta$ .

An important advantage of a metadynamics-based accelerated MD protocol (*dynamically biased* CVHD) is that it does not require the definition of an analytical bias potential function. Rather, a suitable bias potential is constructed on-the-fly. This means that as long the system dynamics can be represented by a combination of local contributions such as (2.1), the algorithm can adapt to all kinds of processes with unknown activation barriers. As such, the algorithm can be interpreted as a self-learning implementation of hyperdynamics. However, if the studied system is reasonably well-characterized, a predefined static CV-based bias (*statically biased* CVHD) could be more suitable. Such a bias can be more efficient as it will immediately start at its full strength and does not need to be built up during the simulation. Therefore, it can be expected that the performance of the dynamic bias method will not be as good as for the static bias in the case of relatively fast successive events. Furthermore, a static bias eliminates the additional simulation parameters introduced by metadynamics: the Gaussian hill width, height and deposition rate. Both methods do, however, share the same fundamental structure summarized in Figure 2.2.

It must be stressed that although the dynamically biased CVHD method uses a metadynamics protocol to construct the bias potential it is, strictly speaking, not a metadynamics method. Indeed, the deposited bias is not stored during the whole simulation, as the CV  $\eta$  does not span the whole configuration space, but only the current state. Once a new state is reached, all deposited bias is deleted and bias deposition is initiated again. Furthermore, bias is never deposited in



**FIGURE 2.2:** Schematic depiction of the structure of the CVHD algorithms.

transition state regions. Therefore, the CVHD method cannot be used to calculate free energy profiles, which require extensive complete sampling of a (limited part of) the configuration space and for which “traditional” metadynamics is required. On the other hand, the many technical aspects shared by CVHD and metadynamics make the former very easy to implement in common computer programs that already support metadynamics. After all, as represented in Figure 2.1, a dynamically biased CVHD simulation is a sequence of standard metadynamics simulations, but each with a different CV.

Calculation of the hypertime, i.e., the physical time recovered from the boost factor, is done the hyperdynamics way, through equation (1.3). Accurate hypertime calculations require good sampling of the regions with large boost, which limits the imposed strength of a static bias potential. In the case of a dynamic bias, however, the deposition of a large bias in a certain region of the CV space implies this region is frequently visited, because the number of hills deposited in a specific region depends on the time the system spends there. This means that the dynamic bias method can apply strong bias potentials while still maintaining a high accuracy of the calculated hypertime.

### 2.2.3 Critical aspects

The CVHD method as conceived above satisfies all requirements we set to meet. Before testing the method on real-life systems, we here summarize the main

critical aspects and intrinsic limitations of the CVHD method in its current form.

First, every valid accelerated MD method employing a bias potential has to ensure that this bias *vanishes in the transition state regions*. In CVHD, the key to achieving this is the parametrization  $\chi(s)$  or choice of  $\chi^{\text{cut}}$ . Our static bias approach will vanish at the transition state if for every local CV  $s_i$ ,  $\chi(s_i^\ddagger) = 1$  (or  $\chi^{\text{cut}} < \chi(s_i^\ddagger)$ ). In the case of eq. (2.3), this means that every  $r_i^{\text{max}}$  of a bond should be smaller than the corresponding  $r_i$  at any transition state  $r_i^\ddagger$ . Parametrizing  $\chi(s)$  or  $\chi^{\text{cut}}$  to achieve this behavior is not a trivial task, as it requires some knowledge of the transition states the system will encounter: a first estimation of “safe” parameter choices can be obtained from some (presumed) relevant transition states that are already known, and verified against other transitions that are discovered during an initial simulation.

When using a dynamic bias, it is possible to impose the same constraint on  $\eta$  as in the static case, which will cause all bias force to vanish before any  $s_i^\ddagger$  is reached. A convenient consequence is that, by construction, no bias can be deposited in the transition state region. In contrast to infrequent metadynamics, CVHD is therefore not restricted to very slow bias deposition and will have larger boost factors. It must be noted that the metadynamics algorithm will by default keep depositing bias at any time, which will become problematic when the system starts spending a large part of its time at the boundaries of the well ( $\eta \sim 1$ ) once the bias is at its full strength in the minima, a situation that becomes more common with increasing system size. If this happens, a large bias could be deposited at the boundary of the CV, which is unphysical and will negatively impact the accuracy of the hypertime calculation and the dynamics in general. A simple solution to avoid this bias pile-up is to restrict the metadynamics algorithm from depositing any bias at large  $\eta$  values (for example, at  $\eta > 0.9$ ). Additional control of the magnitude of deposited bias can be achieved with WT-metaD.

A second phenomenon that can cause excess dynamic bias deposition is the presence of *hidden CVs*. When using CVHD-type CVs, one has to ensure that all relevant degrees of freedom are included in the global CV. This is both an inherent strength and a limitation of the CVHD method. On one hand, if it is possible to describe the full “slow” dynamics by a set of simple localized degrees of freedom—like how chemical reactions involving bond breaking are fully charac-

terized by changes in bond lengths—no hidden CVs will be present. In any case, the bias deposition rate in dynamically biased simulations may not be too high, in order to allow the system to equilibrate along the “fast” degrees of freedom not included in the CV. On the other hand, the requirement that the system dynamics can be decomposed into contributions by a small number of highly localized interactions renders the CVHD method impractical to study systems in which this is not the case: especially complex biological processes involving various types of non-bonded interactions are very difficult to study this way. Fortunately, the latter use case can be covered by infrequent metadynamics, employing a well-tailored set of CVs, whereas the CVHD method is more suitable for, e.g., the prediction of reaction product compositions or the properties of grown materials.

Third, an important disadvantage of CV-based methods (and many conventional hyperdynamics implementations as well) is their *poor scaling with the system size*: in large systems, events will occur more frequently, leading to additional overhead. Representing all to-be-accelerated dynamics by a single CV does not allow for a parallel treatment of events, leading to a deterioration of the boost in such large systems. Possibly, our method could be improved by incorporating aspects of a recently proposed “local” variant of the Bond Boost algorithm.<sup>65</sup> Similarly, existing solutions to deal with the small-barrier problem,<sup>24</sup> to improve the calculation of the hypertime,<sup>66</sup> or to construct appropriate CVs on-the-fly<sup>67</sup> could also be beneficial for the methods presented here.

#### 2.2.4 Related methods

A few other self-learning or adaptive hyperdynamics methods also exist. The recently proposed HD-MD method has a bias depending on the system’s instantaneous potential energy (like the aMD method) and uses short MD runs within each local energy basin to parametrize a  $\Delta V$  on-the-fly.<sup>68</sup> These parameters are chosen based on the desired boost, which can therefore be made equal in every new basin (this is not necessarily the case in CVHD). The fact that the bias potential is only a direct function of the system’s potential energy makes the method conceptually very simple. However, because the method is based on aMD and has no explicit handling of transition states, the bias potential may be nonzero at dividing surfaces.

Also within the context of the aMD method, a simpler algorithm, dubbed the

adaptive hyperdynamics method (AHD) method, was proposed.<sup>69</sup> In the AHD method, the threshold energy  $E$  (below which biasing is allowed) is adjusted in time intervals  $\tau$  by comparing the current potential energy of the system to the minimal energy in the previous interval, and changing the threshold accordingly. Although very simple, the algorithm causes the biasing force to be discontinuous due to the stepwise modification of  $\Delta V$ , which is not the case when metadynamics is used to deposit bias.

Finally, within the framework of the Bond Boost method, a self-learning hyperdynamics method was derived by Perez & Voter.<sup>70</sup> This algorithm calculates an on-the-fly estimate of a pairwise potential of mean force (PMF), which is then used to iteratively improve the bias strength. The method was found to be very efficient in finding exactly the optimal boost for a given process, but lacks the generality and simplicity inherent to using generic CVs and a metadynamics-like self-learning bias.

## 2.3 Applications

### 2.3.1 General methodology

All simulations were carried out using the LAMMPS package<sup>71</sup> and a modified version of the Colvars module.<sup>32</sup> To control the system temperature, ensure its homogeneity, and allow for swift decorrelation, a Langevin-type thermostat<sup>72</sup> with relaxation time of 1 ps was employed. The equations of motion were integrated with a MD time step of 1 fs, except when using the ReaxFF potential, which required a 0.1 fs time step. Boost and hypertime were calculated by updating a running average of eq. (1.3) at every step; we did not compensate for the overhead induced by the bias calculation as it was found to be insignificant compared to the evaluation of the interatomic potential.

### 2.3.2 Bond-based CV: diffusion on Cu(001)

As an illustration of the bond-based CV in equation (2.3), we apply both the static and the dynamic CVHD methods to diffusion on the Cu(001) surface. Specifically, we apply the methods to the diffusion of adatoms, dimers and vacancies, which can all diffuse through simple hopping mechanisms. Copper adatom and dimer diffusion, however, can also occur by a two-atom exchange.<sup>73</sup> Thus, we can assess the performance of the bias methods for a set of competitive mechanisms with different characteristics in terms of the number of atoms

and bonds involved, and minimal  $r_i^\ddagger$ . The same processes have also been previously studied with the Bond Boost method,<sup>23</sup> which allows us to directly compare the performance of our generic CV-based methods to a dedicated hyperdynamics implementation.

The studied system consisted of a six-layer slab, each layer containing 50 atoms. The Cu–Cu interactions were described using a standard EAM potential.<sup>74,75</sup> The two bottom layers were kept fixed and, depending on the studied mechanism, an adatom or dimer was placed on top of the slab, or a vacancy was created by removing an atom from the top layer.

Climbing Image Nudged Elastic Band (CI-NEB)<sup>11</sup> calculations were used to obtain information on the minimal  $r_i^\ddagger$  values of breaking bonds associated with every mechanism. Of all the mechanisms considered, these are the smallest for adatom hopping, where the two partially broken bonds have a length of 3.3 Å at the transition state. Therefore, a global  $r_i^{\max}$  value of 3.3 Å was used for each bond in all simulations. The  $r_i^{\min}$  value for every bond was chosen to be the average bond length as obtained from an equilibration run of  $t_w$  (after an initial waiting time  $t_w$  before detecting a transition), and  $r_i^{\text{cut}}$  was a global constant of 3.0 Å. We furthermore set  $p = 8$  and  $t_w = 5$  ps. In the dynamic bias simulations, Gaussian hills with a height of  $w = 0.005$  eV and width of  $\sigma = 0.025$  were added with an interval of 1 ps. WT-metadD with a bias temperature  $\Delta T = 2000$  K was used to deposit the dynamic bias. The hill height was kept sufficiently low ( $< k_B T$ ) to minimize temporary heating effects—and incorrect sampling—along the CV. (Taking  $10w \approx k_B T$  is a rule of thumb we will also implicitly adopt for later applications.) Analogously, we set  $\Delta V^{\max} = 0.3$  eV in the static bias simulations, as a compromise between large boosts and accurate hypertime sampling.

Reaction rates  $k_i$  for all event types  $i$  were quantified by counting the observed number of diffusion events  $n_i$  and dividing this by the calculated hypertime,  $k_i = n_i/t_{\text{hyper}}$ . At temperatures between 150 and 600 K, rates were obtained and averaged over multiple runs of  $2 \times 10^8$  steps ( $1 \times 10^8$  for vacancies). Then, the calculated rates were fitted to the Arrhenius relation  $k(T) = Ae^{-\beta E_A}$ . The fitted prefactors  $A$  and activation energies  $E_A$  can be readily compared to those reported by Miron and Fichthorn in their Bond Boost study of the same system.<sup>23</sup>

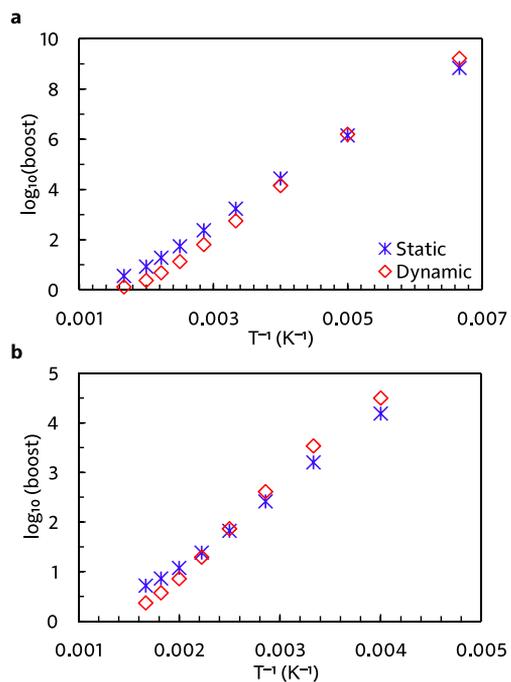
It can be seen from Table 2.1 that both the static and dynamic biasing methods yield the same values for  $A$  and  $E_A$ , both in excellent agreement with the

**TABLE 2.1:** Prefactors ( $A$ , in THz) and activation energies ( $E_A$ , in eV) for elementary diffusion processes on the Cu(001) surface, as obtained from dynamic bias (DB) and static bias (SB) CVHD simulations (see text for details). For comparison, the Bond Boost (BB) estimates and static barriers computed with the CI-NEB method are also included. DB and SB error bars reflect the 90% confidence interval.

	$A^{\text{DB}}$	$A^{\text{SB}}$	$A^{\text{BB}}$	
vacancy hop	$83e^{\pm 0.1}$	$54e^{\pm 0.5}$	$54e^{\pm 0.5}$	
adatom hop	$54e^{\pm 0.3}$	$33e^{\pm 0.1}$	$40e^{\pm 0.5}$	
adatom exchange	$430e^{\pm 1.0}$	$130e^{\pm 1.2}$	$270e^{\pm 0.6}$	
dimer hop	$34e^{\pm 0.2}$	$21e^{\pm 0.2}$	$30e^{\pm 0.7}$	
dimer exchange	$137e^{\pm 1.3}$	$213e^{\pm 1.2}$	$190e^{\pm 0.8}$	
	$E_A^{\text{DB}}$	$E_A^{\text{SB}}$	$E_A^{\text{BB}}$	$E_A^{\text{CI-NEB}}$
vacancy hop	$0.44 \pm 0.01$	$0.43 \pm 0.02$	$0.44 \pm 0.03$	0.44
adatom hop	$0.53 \pm 0.01$	$0.51 \pm 0.01$	$0.52 \pm 0.03$	0.51
adatom exchange	$0.76 \pm 0.04$	$0.71 \pm 0.05$	$0.73 \pm 0.04$	0.71
dimer hop	$0.51 \pm 0.01$	$0.49 \pm 0.02$	$0.47 \pm 0.03$	0.49
dimer exchange	$0.74 \pm 0.06$	$0.76 \pm 0.05$	$0.71 \pm 0.06$	0.70

Bond Boost result and the static barriers calculated by the CI-NEB method. The performance of both CV-based methods, as expressed by the achieved boost is, as Figure 2.3 shows, quite comparable, yielding a boost up to  $3 \times 10^4$  for adatom diffusion at a temperature of 250 K and up to  $10^9$  for vacancy diffusion at 150 K. The methods show the same basic behavior typical for hyperdynamics methods, with a boost that declines with increasing temperature, due to the inverse temperature dependence of  $\beta$ .<sup>23</sup> Furthermore, it can be seen that the relative efficiency of the static bias method compared to its dynamic counterpart improves with increasing temperature. This is because waiting times between events are shorter at higher temperatures which, as a result, puts the dynamic method at disadvantage. This difference becomes irrelevant at lower temperatures where waiting times, even with full-strength static bias, are much longer than the time needed by the metadynamics protocol to construct the dynamic bias.

Not all processes occur at similar rates at every temperature. As Figure 2.3a shows, vacancy diffusion was studied at temperatures as low as 150 K, where



**FIGURE 2.3:** Calculated boost factor as function of the temperature, of both the static as the dynamic cvhd method, for (a) vacancy and (b) adatom diffusion on Cu(001). (As explained in the text, dimer diffusion behaves essentially the same as adatom diffusion.)

we were able to observe about 60 events over a total MD time of 0.5  $\mu$ s: this is because the associated hypertime reached about 500 s. All other processes, however, are much slower at this temperature. For example, according to the kinetic parameters in Table 2.1, adatom hopping will be about 2000 times slower, with an average waiting time in the order of  $10^4$  s, explaining why we were able to observe the latter process with an appreciable frequency only from 250 K and higher, as depicted in Figure 2(b). The higher barriers and waiting times of the adatom diffusion processes, compared to vacancy diffusion, also increase the relative efficiency of the dynamic bias method, as expected. Similarly, exchange processes could only be observed starting from 400 K, and show the same boost characteristics as their respective hopping counterparts. In general, the bond-based CV as used in the CVHD algorithm performs about as well as the Bond Boost method, albeit being part of a much more generic framework. Furthermore, because all relevant processes were already well-characterized, use of the dynamic biasing method has no added advantage since an optimized static bias could be applied at lower cost.

### 2.3.3 Dihedral-based CV: folding of a helix

To demonstrate the flexibility of the CVHD framework, we study a process governed by a different local property or CV: the folding of extended chain to a helix. This process is very different from the Cu(001) diffusion example in two ways. First, the activated processes underpinning the system evolution are not bond breaking, but rotation around bonds, changing the dihedral angles. Second, whereas the Cu(001) system remained in equilibrium, and every state was associated with the same handful of possible escape pathways, a folding chain may visit a much larger number of different states, all of which may have wildly different and unpredictable kinetic and thermodynamic stabilities. The dihedral angle is also a four-atom local property, setting it apart from the pairwise properties that are commonly used in hyperdynamics methods such as the Bond Boost method.

We use a simple model system, consisting of a chain of 50 connected beads, interacting through a number of standard potential energy functions

$$V = V_{\text{bond}} + V_{\text{angle}} + V_{\text{non-bond}} + V_{\text{dihedral}}. \quad (2.7)$$

The bond and angle terms are simple harmonic potentials of the bond length  $r_i$

and valence angle  $\theta_i$ , respectively:

$$V_{\text{bond}} = \sum_i^{\text{bonds}} K_r (r_i - r_0)^2, \quad (2.8)$$

and

$$V_{\text{angle}} = \sum_i^{\text{angles}} K_\theta (\theta_i - \theta_0)^2. \quad (2.9)$$

We choose parameters that are in the order of magnitude of those used for C–C bonds. There is no specific real system we wish to model, but as a global target, the beads are chosen to be rather close to carbon atoms in behavior and properties. This way, we can ensure that all values and system properties are in the range of those of real molecular systems, rather than first calibrating a true coarse-grained potential. Based on these considerations we use  $r_0 = 1.52 \text{ \AA}$ ,  $\theta_0 = 109.5^\circ$ ,  $K_r = 5 \text{ eV/\AA}^2$ , and  $K_\theta = 2 \text{ eV/rad}^2$ .

Non-bonded interactions are taken into account for all pairs other than 1–2, 1–3 and 1–4 interactions using a standard Lennard-Jones potential:

$$V_{\text{non-bond}} = \sum_i^{\text{pairs}} 4\varepsilon \left[ \left( \frac{\sigma}{r_i} \right)^{12} - \left( \frac{\sigma}{r_i} \right)^6 \right], \quad (2.10)$$

whose parameters are  $\sigma = 4 \text{ \AA}$  and  $\varepsilon = 2.5 \text{ meV}$ . This interaction is mainly included to provide a repulsive core to all particles, in order to prevent inter-particle distances to become unphysically close. Also, care must be taken that the non-bonded attractions do not lead to similar energy contributions as the dihedral term, as this will make the simple potential energy surface more complicated, and anomalously stabilize more compact, less ordered states.

Finally, the dihedral interaction is given by<sup>76</sup>

$$V_{\text{dihedral}} = \sum_i^{\text{dihedrals}} A [1 - \cos(\phi_i)] + B [1 + \cos(3\phi_i)] + C \left[ 1 + \cos\left(\phi_i + \frac{\pi}{4}\right) \right], \quad (2.11)$$

which produces three minima, one corresponding to the trans ( $t$ ) state, and the two gauche states ( $g^+$  and  $g^-$ ). Of these,  $g^-$  is preferred. We found that setting  $A = B = C = 0.3 \text{ eV}$  gives  $t \rightarrow g^-$  transition barriers in the order of 8 kcal/mol,

which is similar to those in peptide backbones. With these settings, we can guarantee that a helical structure is the global energy minimum and crossing torsional barriers is an infrequent event but, at the same time, the full folding of helix is still possible to study in long unbiased MD simulations (within about 10  $\mu$ s).

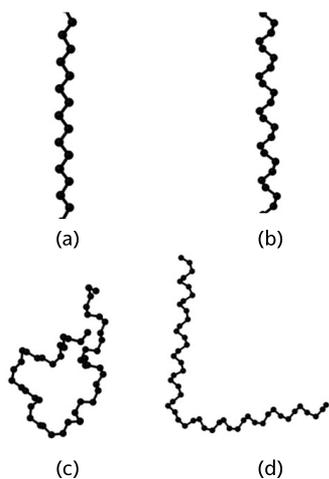
Similar to the bond-based system property of eq. (2.3), the local distortion function can be calculated from a dihedral angle  $\phi_i$  as

$$\chi_i = \begin{cases} \left| \frac{\phi_i - \phi_i^{\text{ref}}}{\Delta\phi_i} \right| & \text{if } \phi_i \in [\phi_i^{\text{ref}} \pm \Delta\phi_i] \\ 1 & \text{if otherwise} \end{cases} . \quad (2.12)$$

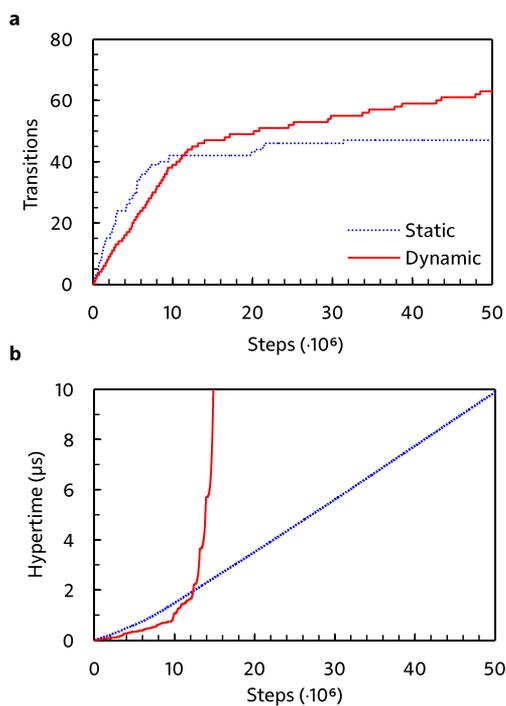
Here,  $\phi_i^{\text{ref}}$  is a reference dihedral angle, which is determined as the dihedral angle of the closest local minimum (here,  $\pm 60^\circ$  or  $\pm 180^\circ$ ) at the moment the property list is created.  $\Delta\phi_i$  is the maximal deviation from  $\phi_i^{\text{ref}}$  that keeps  $\phi_i$  far enough from transitions, which we set to be  $40^\circ$ . The choice of  $\Delta\phi_i$ , like all parameterizations of a local distortion  $\chi$ , requires some *a priori* knowledge of the system; in our case, we use a well-defined model potential, but in more complicated systems one should always verify that for all observed events the requirement that  $\eta = 1$  at the transition state is satisfied. The other parameters do not depend on the local property used, and we set  $p = 8$  and  $t_w = 5$  ps. In the dynamic bias simulations, Gaussian hills with a height of 0.005 eV and width of 0.05 were added with an interval of 1 ps and a bias temperature of  $\Delta T = 1000$  K. In the case of the static bias, we use  $\Delta V^{\text{max}} = 0.15$  eV.

We start every simulation from a fully extended chain, with all dihedrals in the  $t$  conformation (see Figure 2.4a), and run unbiased MD, static and dynamic bias simulations at 300 K, while monitoring the number of dihedral rotations. Because of the way our bias potential is designed, it is highly unlikely any other transitions will occur before there are no more  $t \rightarrow g^-$  events possible. Therefore, 47 transitions were always found to be sufficient to reach a perfect helix Figure 2.4b.

Examples of the system evolution in terms of the number of transitions are depicted in Figure 2.5a for both static and dynamic bias. It can be seen that initially, using a static bias is initially more efficient than a dynamic bias. As discussed before, a well-tailored static bias has the advantage of starting at its full strength, whereas a dynamic bias takes time to be constructed. However,



**FIGURE 2.4:** Some possible states of the helix model system: (a) fully extended, (b) perfect helix, (c) half-folded after 23 folding events and (d) after 46 (or, rather,  $46 + 2n$ ) events.



**FIGURE 2.5:** (a) Number of transitions as a function of simulation time, and (b) hyper-time as a function of simulation time in the helix folding test system, for both static and dynamic bias.

we notice that after 42 transitions, the dynamically biased simulation surpasses the statically biased case. Indeed, because the studied process obeys first order kinetics in the number of  $t$  dihedrals, the waiting time between events increases the closer the system is to the fully folded state. The choice of  $\Delta V^{\max}$  was calibrated to the initial phase of the folding process, and is very well-suited for this first stage with relatively short waiting times, but becomes less efficient when the number of  $t$  dihedrals is low and the waiting time is increased. The dynamic biasing scheme, on the other hand, will keep strengthening the bias potential while waiting for a transition to occur. The dynamic CVHD method thus dynamically uses a larger bias in the case of long waiting times.

A different look at this subject is given by Figure 2.5b, which compares the hypertime reached by both methods as a function of the simulation time. Whereas this quantity linearly increases for the static bias, indicating a constant boost, it shows an exponential increase in the case of the dynamic bias. Because of this property, the dynamic biasing scheme resulted in a perfect helix about two times faster than the static scheme.

The achieved boost factors of both biasing methods are, as Figure 2.5b shows, quite different. On one hand, the static biasing scheme provides a constant boost of about 200. On the other hand, the dynamic scheme adapts itself to match the boost requirements of the current state in which the system resides – arriving at an accumulated boost of about 500 when completing the folding process. This also explains the different behavior of both methods after the perfect helix is formed: any transition from the  $g^-$  state has a barrier that is about twice as high as the one associated with the  $t \rightarrow g^-$  transition, which means that the statically biased simulation will not be able to escape from the helix state within a reasonable computational time. However, in the dynamically biased simulation, the bias will slowly be increased until a transition to a less favorable state (Figure 2.4d) can occur. Such a process only takes place after about 0.1 ms, much longer than the 5  $\mu$ s required to obtain the folded helix. The escape pathways from this unfavorable “kinked” helix have lower barriers than from the perfect helix. As a result, the construction of a suitably strong bias to return to the latter state requires fewer simulation steps than the reverse reaction, explaining the successive occurrence of short and long “steps” in the dynamic transitions curve in Figure 2.5a.

Although the general performance characteristics of both accelerated meth-

**TABLE 2.2:** Average folding and half-folding times ( $\mu\text{s}$ ) for the helix model system, as obtained from unbiased MD and CVHD with static and dynamic bias, respectively. All error bars are at the 90% confidence level.

	unbiased	static bias	dynamic bias
half folding	$0.45 \pm 0.08$	$0.38 \pm 0.05$	$0.39 \pm 0.04$
full folding	$6 \pm 2$	$7 \pm 2$	$6 \pm 1$

ods in the case of the folding model system have now been established, their accuracy must still be ascertained. Table 2.2 therefore collects the average times for full folding (47 transitions) and obtaining a half-folded structure, such as the one depicted in Figure 2.4c (23 transitions), as obtained by CVHD simulations but also long unbiased MD runs. For both processes, all considered methods are in excellent agreement with each other.

### 2.3.4 Methane dissociation on Ni(111)

Finally, as an example of the kind of complex dynamics that can be accessed with CVHD simulations, we consider the catalytic dissociation of methane on the Ni(111) surface. This process is important not only in methane reforming processes, but also in chemical vapor deposition growth of carbon nanostructures. The initial dissociative adsorption of  $\text{CH}_4$  has an activation energy in the order of 20 kcal/mol,<sup>77</sup> rendering direct MD simulations of this process difficult; to observe appreciable  $\text{CH}_4$  dissociation, previous simulation attempts were required to use elevated temperatures<sup>78–80</sup> (up to 1500 K), instead of experimental temperatures of 800–1000 K, or only focus on plasma-activated species<sup>81,82</sup> ( $\text{CH}_x$  radicals). From a technical point of view, this type of reaction is a useful additional test case for the CVHD method, being both an example of a system with a phase boundary, and of heterogeneous catalysis in general.

The methane dissociation process is modeled starting from a single  $\text{CH}_4$  molecule above a six-layer nickel slab (64 atoms per layer), with the two bottom layers held fixed. A reflective wall is used at a  $z$ -height of 20 Å, leading to a gas phase volume of about  $19.9 \times 17.4 \times 8 \text{ \AA}^3$ . The interatomic interactions are described by the ReaxFF potential<sup>83</sup> as implemented in LAMMPS,<sup>84</sup> using the Ni/C/H parameter set of Mueller *et al.*<sup>85</sup> and the QEq method<sup>86</sup> to calculate atomic charges.

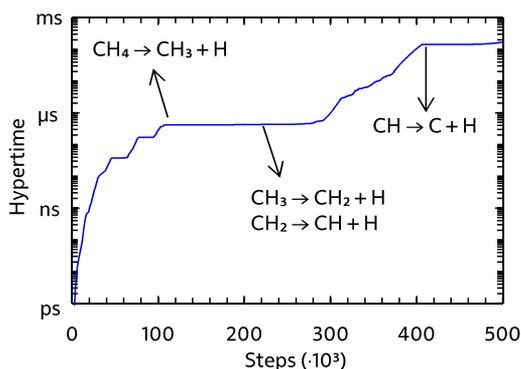
The simulations are carried out at 800 K, applying the bond-based CV of

**TABLE 2.3:** Average reaction time for all elementary reaction steps of the full methane dehydrogenation process  $\text{CH}_4(\text{g}) \rightarrow \text{C}(\text{ad}) + 4 \text{H}(\text{ad})$ , on Ni(111) at 800 K, as obtained from ReaxFF dynamically biased CVHD simulations. Reaction times are given as a 90% confidence interval.

reaction	time
$\text{CH}_4(\text{g}) \rightarrow \text{CH}_3(\text{ad}) + \text{H}(\text{ad})$	4–9 $\mu\text{s}$
$\text{CH}_3(\text{ad}) \rightarrow \text{CH}_2(\text{ad}) + \text{H}(\text{ad})$	0.09–0.22 $\mu\text{s}$
$\text{CH}_2(\text{ad}) \rightarrow \text{CH}(\text{ad}) + \text{H}(\text{ad})$	37–91 ps
$\text{CH}(\text{ad}) \rightarrow \text{C}(\text{ad}) + \text{H}(\text{ad})$	0.3–0.8 ms

eq. (2.3) with  $p = 6$  and  $t_w = 0.1$  ps to C–H bonds, and dynamic biasing with a deposition stride of 10 fs, a hill width of 0.025, a hill height of 0.25 kcal/mol and a bias temperature of 4000 K. Compared to the previously discussed processes, metal-catalyzed methane decomposition poses two additional challenges. First, the general problem of thermostating gas phase species is that it is a poor model of energy exchange in such a system: in reality, this only occurs at discrete moments in time during collisions. Also, a Langevin thermostat distorts the diffusion path of gas-phase particles. However, thermostating the methane molecule is necessary to dissipate the excess energy introduced by the dynamic bias procedure, and to avoid unphysical heating of the molecule. As a compromise, we apply a separate Langevin thermostat to the  $\text{CH}_4$  molecule that only acts on the vibrational and rotational degrees of freedom, and leaves its translational motion untouched. The second problem is specific to the successive dehydrogenation pathway of methane, in which the separate steps have very different bond lengths at the transition state ( $r^\ddagger$ ) for the dissociating bonds, ranging from 1.55 Å for  $\text{CH} \rightarrow \text{C} + \text{H}$  to 1.80 Å for  $\text{CH}_3 \rightarrow \text{CH}_2 + \text{H}$  on the Ni(111) surface.<sup>87</sup> Safe  $r^{\text{max}}$  values for the former were found to perform very poorly when attempting to boost the latter process. Therefore, we used a global  $r^{\text{max}}$  value of 1.8 Å and  $r^{\text{min}} = 0.9$  Å. Although the safety of such a setting is not completely guaranteed, we found that little to no bias was effectively deposited in the transition state regions of the “unsafe” cases.

We carried out 15 independent simulations of  $10^6$  steps, corresponding to a MD time of 100 ps each, and were always able to observe the full methane decomposition process. For every elementary step in the reaction, we calculated the average reaction time, summarized in Table 2.3. These results demonstrate



**FIGURE 2.6:** Hypertime evolution in a dynamically biased CVHD simulation of methane decomposition on Ni(111), at 800 K. The observed elementary steps are shown at the time step at which they occurred.

the usefulness and power of a dynamic biasing method. Indeed, methane decomposition at 800 K is a process that consists of rather fast steps such as the dissociation of a C–H bond of adsorbed  $\text{CH}_2$  (which takes about 50 ps), to the very slow decomposition of adsorbed CH, which requires more than 0.1 ms. Studying this reaction sequence with a static bias would therefore not be achievable; the vast time scale spread of the various elementary processes is illustrated in Figure 2.6. Boosts of  $2 \times 10^6$  are achieved.

The very long reaction time of the various reactions means that we cannot verify the accuracy of the values in Table 2.3 by direct comparison to MD simulations: even at a temperature of 1000 K we did not observe any reaction within 100 ps. It is, however, possible to compare the relative rates of the elementary steps to estimations based on differences between their respective activation energies.<sup>87</sup> The barrier for C–H dissociation in adsorbed  $\text{CH}_3$  is 10 kcal/mol higher than for adsorbed  $\text{CH}_2$ , meaning that the latter is about 1000 times faster than the former, in agreement with our findings. Similarly, the dissociation of adsorbed CH has a barrier that is 14 kcal/mol higher than the dissociation step involving adsorbed  $\text{CH}_3$ , leading to a rate that differs by an order of magnitude of  $10^4$ , again in agreement with the results in Table 2.3. Finally, according to kinetic theory, the initial  $\text{CH}_4$  pressure is about 40 bar, with a flux to the surface of  $0.3 \text{ ps}^{-1}$ . Considering a dissociation barrier of 19 kcal/mol,<sup>87</sup> we can make a crude estimation of the average reaction time to be 0.5  $\mu\text{s}$ , which is also in line with our observations.

## 2.4 Conclusions

We have developed a theoretical framework, the *collective variable-driven hyperdynamics* (CVHD) method, which is an implementation of the hyperdynamics method that includes some of the strengths of metadynamics. The CVHD method is intended to be used as an accelerated molecular dynamics method, in which the waiting time between infrequent events is shortened by adding a bias potential to the energy minima in the system, without requiring a priori knowledge of the pathways and states that will be encountered. A self-learning implementation of hyperdynamics is, essentially, realized as a sequence of discrete metadynamics simulations. At each point in time, a single global CV  $\eta$  measures the deviation of the system from its current reference state, while the metadynamics biasing algorithm is used to dynamically build up a suitable bias potential for every new potential energy basin the system encounters, and eventually escapes from. This *dynamically biased* CVHD method does not require an *a priori* knowledge of the activation barriers the system can encounter during its long time scale evolution. Furthermore, unlike standard metadynamics CVs, no specific details about possible pathways or mechanisms have to be supplied for the construction of  $\eta$ . Rather, CVHD only requires the identification of all possibly relevant degrees of freedom, from which a fresh definition of  $\eta$  is automatically generated after each event. CVHD is therefore easily adaptable to wildly different processes: in this chapter, we have demonstrated its applicability to solid state diffusion, heterogeneous catalysis, and chain folding.

If the studied process is already well-characterized and all relevant activation barriers are known, the *statically biased* CVHD method is the optimal choice: it is easy to construct well-optimized one-size-fits-all static bias, where the on-the-fly construction of a dynamic bias will only cause additional overhead. On the other hand, in systems undergoing a more complex evolution, using a dynamical metadynamics-based bias may be the more optimal choice, as it is generally not possible to construct a single static bias that is both safe and efficient for every process encountered. This ability of the dynamically biased CVHD method to adapt its bias to the specific requirements of the system at any time is an important advantage of the method. Irrespective of their relative efficiency, however, both biasing methods give rise to a correct sequence of state-to-state transitions and are therefore formally correct hyperdynamics methods.

Although the CVHD method is inherently flexible in the kind of local CVs it can use to calculate its global CV, its performance does not seem to suffer from this genericity. For example, in the case of low-temperature diffusion on the Cu(001) with the bond length local property, accelerations as large as  $10^9$  can be obtained, corresponding to physical times up to several seconds. In general, the CVHD method is about as efficient as the Bond Boost implementation of hyperdynamics, but has the added advantage of being more general. The local distortion functions developed so far already span a large range of processes and systems, and additional ones can be incorporated to further extend the scope of the method. We therefore believe that the CVHD method will be a valuable tool in the study of slow or activated processes in a wide range of scientific fields including growth, conformational sampling, and catalysis. More generally, CVHD demonstrates how well-known simulation techniques can be combined and applied in surprising new ways; it will give access to simulations that used to be impossible.

# Advanced applications of CVHD

### 3.1 Pyrolysis and combustion as a case study

A detailed understanding of pyrolysis and combustion is of great technological and industrial importance. A fundamental insight in (bio)fuel decomposition chemistry is essential to improve the selectivity of cracking and reforming processes and increase the efficiency of combustion engines and minimize their production of pollutants. For example, low temperature combustion (LTC) strategies can significantly decrease production of particulate matter (PM) and nitrogen oxides ( $\text{NO}_x$ ) in engines, but additional insights in their operation are required for further optimization.<sup>88,89</sup> To screen and improve possible operating conditions, kinetic modeling can be used to explain and guide experimental investigations.<sup>90–93</sup> It is, however, extremely challenging to create sufficiently complete and accurate kinetic models due to the wealth of possible intermediates and pathways that can all contribute significantly to the overall process, which generally limits predictive power of these models. Therefore, such a class of processes provides an excellent test case for the *predictive* capabilities of CVHD. In principle, CVHD only requires the definition of an initial state of the system (i.e., the fuel or fuel–oxygen mixture, density, and temperature), and all relevant pathways, intermediates and products under these conditions will follow naturally from these choices. Hence, CVHD simulations can not only be used to predict the outcome of a complex chemical process, but the thus obtained fundamental knowledge can also be used to extend and improve existing kinetic models.

---

The results presented in this chapter were published in:

K. M. Bal and E. C. Neyts. Direct observation of realistic-temperature fuel combustion mechanisms in atomistic simulations. *Chem. Sci.* **7**, 5280 (2016).

Interesting for our purposes is also the availability of a perfectly cromulent ReaxFF potential, which has been successfully applied to various pyrolysis and combustion reactions, and gives us the ability to conduct simulations of large-scale processes over fairly long MD time scales.<sup>94-102</sup> CVHD simulations also have the potential to significantly improve on the state-of-the-art: previous MD studies invariably used very high (>2000 K) simulated temperatures to be able to observe any appreciable pyrolysis or combustion chemistry within the short MD time scale. Of course, it is difficult to correlate insights from high-temperature simulation with industrially relevant processes at lower temperatures, such as alkane cracking at ~1000 K or low-temperature diesel engines, and a clear motivation to extend the simulation time scale ought to be present in the combustion community. It is a problem that also touches on the broader question raised earlier in this thesis: with ReaxFF, much effort has been directed towards the development of a rather complex and accurate description of the interatomic forces, which is then essentially wasted on simulations that have little relevance to the actual processes of interest. Hence, the quality of the model's Hamiltonian should be matched by a proper correspondence with experimental conditions, which means that the simulation time scale must be dramatically extended.

CVHD is the only candidate method that can meet this goal. As mentioned before, methods that require saddle-point searching, such as TAD or adaptive KMC, have difficulties handling liquid- or gas-like systems. The PRD method imposes almost no constraints on the simulations and has been applied to the thermal decomposition of *n*-hexadecane<sup>103</sup> and 1-hexene.<sup>104</sup> In the latter case, pyrolysis could be simulated at 1350 K over a simulated time of ~1  $\mu$ s by using up to 180 replicas. A further extension of the PRD time scale to the millisecond-to-second range necessary for capturing processes at temperatures of 1000 K or lower is, however, impractical. Indeed, because the boost factor of PRD is proportional to the number of processors, simulating this kind of process would put unrealistic demands on available computational resources.

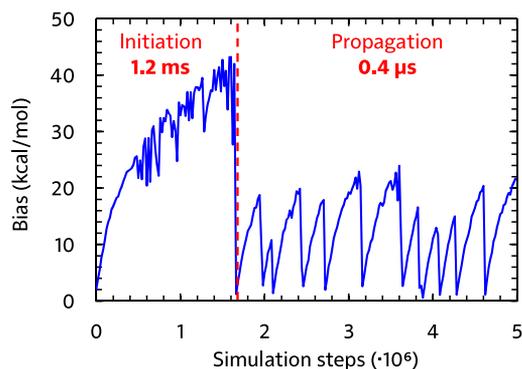
Arbitrarily long time scales can be reached with hyperdynamics, but traditional implementations of the method such as Bond Boost would also not be suitable: there is a large separation of reaction barriers (and associated reaction time scales) that can be encountered during the process, ranging from ~30 kcal/mol for alkyl radical  $\beta$ -scissions to ~80 kcal/mol for initiation reac-

tions of alkane pyrolysis. This has a major impact on the applicability of hyperdynamics, since a simple “static” bias potential can only be designed to work well for a small range of possible barriers; a bias that achieves a good acceleration of  $\beta$ -scissions will still fall short in bringing the initiation reaction within reach. In some specific cases, a conventional hyperdynamics scheme can be sufficient: Cheng *et al.* exploited the very fast radical chemistry in hydrogen combustion by only applying a predefined Bond Boost bias potential to radical initiation.<sup>105</sup> However, the much longer lifetimes of hydrocarbon radicals<sup>104</sup> and the employed ReaxFF-specific concepts render this approach not generally applicable. The self-learning bias of CVHD, however, should be uniquely suited to deal with this type of problem.

In this work, we apply the CVHD method to the initial phase of *n*-dodecane pyrolysis and combustion to, for the first time, uncover detailed atomic-level fuel decomposition pathways under realistic conditions. These simulations are the first direct atomistic simulations of fuel pyrolysis and combustion chemistry under realistic conditions and provide an additional validation of contemporary mechanistic insights.

## 3.2 Methodology

All simulations were carried out with LAMMPS<sup>71</sup> and the Colvars module,<sup>32</sup> using the ReaxFF potential<sup>83</sup> with the Chenoweth *et al.* parameter set<sup>94</sup> and QEq charge equilibration,<sup>86</sup> as implemented in LAMMPS.<sup>84</sup> The equations of motion were integrated with a time step of 0.1 fs, and the system was initially equilibrated at the target temperature with a Langevin-type thermostat.<sup>72</sup> Further sampling in the NVT ensemble was achieved through application of a Nosé-Hoover chain<sup>106</sup> with a relaxation time of 0.1 ps, whereas for isotropic NPT simulations, the Martyna-Tobias-Klein (MTK) equations of motion<sup>107</sup> were integrated through the scheme of Tuckerman *et al.*,<sup>108</sup> using a relaxation time of 1 ps. For pyrolysis simulations, the system consisted of 24 alkane molecules in a  $50 \times 50 \times 50 \text{ \AA}^3$  periodic box, corresponding to a density of about  $0.05 \text{ g/cm}^3$ . As local degrees of freedom we used C-C and C-H bond lengths to calculate the bond distortion according to equation (2.3). The  $r_i^{\min}$  and  $r_i^{\max}$  parameters were respectively 1.55 and 2.20 Å for C-C bonds, and 1.05 and 1.65 Å for C-H bonds. The  $r_i^{\max}$  values were specifically chosen to be smaller than the lengths of the



**FIGURE 3.1:** Applied maximal bias potential during the initial steps of a 1000 K CVHD pyrolysis simulation. The time scales of the two displayed distinct regimes are also shown.

breaking C–C and C–H bonds in the transition states of radical  $\beta$ -scissions and intramolecular hydrogen atom transfers, respectively, to ensure these states remain unbiased. This choice of CV means that only events involving bond breaking are accelerated, and conformational changes are unbiased; following a similar reasoning as in previous work, low-barrier conformational dynamics can be considered to have reached equilibrium well within the time spent while waiting for a reaction.<sup>103</sup> Gaussian hills of width  $\sigma = 0.025$  and height  $w = 0.25$  kcal/mol were added every 0.2 ps; the waiting time to detect events was  $t_w = 1$  ps. CVHD simulations were carried out between 1000 and 1800 K; for comparison, unbiased MD simulations were conducted at a temperature of 2500 K.

Constant density combustion simulations were carried out for a  $40 \times 40 \times 40 \text{ \AA}^3$  box containing 5 *n*-dodecane and 100 oxygen molecules, corresponding to a fuel-lean mixture with a density of about  $0.1 \text{ g/cm}^3$ . The CVHD parameters are the same as those of the pyrolysis simulations, with all interactions involving oxygen atoms being described by the corresponding values for carbon. Biased simulations were carried out between 700 and 1800 K, and conventional MD was again performed at 2500 K. The average pressures in these simulations range from  $\sim 200$  bar at 700 K to almost 500 bar at 1800 K.

In order to capture the pressure dependence of the oxidation process over the range of pressures relevant to practical combustion applications, we also carried out a set of constant pressure simulations at 1000 K and pressures between 10 and 500 bar. CVHD simulations of gas-phase systems suffer from a

**TABLE 3.1:** Lowest temperatures achieved in the CVHD simulations of *n*-dodecane pyrolysis and combustion, and corresponding physical times and boost factors.

	pyrolysis	combustion
Lowest temperature	1000 K	700 K
Longest simulated time	57 ms	39 s
Largest boost	$6.3 \times 10^6$	$1.3 \times 10^9$

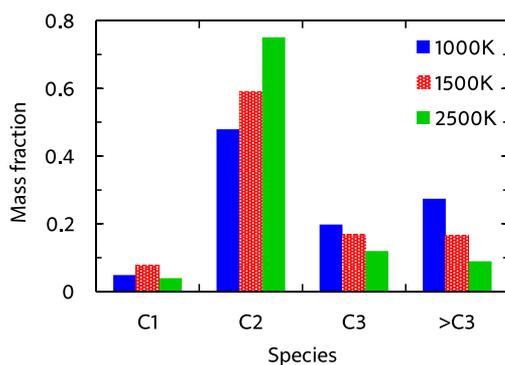
particular complication: lower pressures also mean lower collision frequencies in the system. Therefore, to prevent excessive buildup of bias between possible reactive collisions, and an overestimation of the time scale, the Gaussian deposition stride must be lowered accordingly. While 0.2 ps suffices for the high-density NVT simulations, we found that the 10 bar simulation requires a deposition interval of 0.5 ps, a value we used in all NPT simulations. When applying CVHD to other gas-phase systems, care must again be taken to choose an appropriate deposition stride.

Unless noted otherwise, all comparisons between simulations at different temperatures, such as of product compositions and time scales, are made at a fixed conversion level. For pyrolysis, analysis was performed at 50% fuel conversion. Combustion simulations were carried out until 20% of the O<sub>2</sub> molecules were consumed. For every condition, two independent trajectories were calculated to obtain reliable statistics. Error intervals, if reported, reflect the 90% confidence level.

## 3.3 Results

### 3.3.1 Accessible time scale

The dynamic self-learning nature of the CVHD method is illustrated in Figure 3.1, which shows the evolution of the applied bias potential in the first stages of a pyrolysis simulation: the bias strength is slowly increased until an event is detected, and the biasing procedure is restarted. It can also be seen that the initiation reaction, which is a C–C bond fission, is the slowest event that requires the largest bias potential, whereas subsequent radical isomerizations and  $\beta$ -scissions have lower barriers. Thus, the bias strength is automatically tuned to be optimal for the current stage of the simulation. As summarized in Table 3.1, application of CVHD allows us to observe alkane pyrolysis and combus-



**FIGURE 3.2:** Products of the *n*-dodecane pyrolysis simulations at different temperatures. The mass fraction is that of carbon only, and reflects how carbon is distributed over the various species.

tion at temperatures as low as 1000 and 700 K, respectively; the largest boost factor in our simulations is  $8 \times 10^6$  larger than that of the longest pyrolysis PRD simulation.<sup>104</sup> The longest simulated physical time is therefore almost 40 s.

### 3.3.2 Pyrolysis

In general, the alkane decomposition chemistry observed in the CVHD simulations is similar to previous high-temperature (>2000 K) MD simulations of alkane pyrolysis.<sup>96,97</sup> Most reactions of large alkyl radicals are either isomerization by intramolecular H-transfer, or decomposition to 1-alkenes through  $\beta$ -scission (the *Rice-Kossiakoff mechanism*). At high temperatures, the entropically favored decomposition reactions are the dominant process: ethylene is by far the dominant reaction product, in agreement with previous high-temperature MD simulations. In contrast, low-barrier isomerization occurs much more frequently at low temperatures, forming more stable secondary radicals which give rise to the formation of larger 1-alkenes after eventually undergoing  $\beta$ -scission. Therefore, lower pyrolysis temperatures yield larger product molecules, as shown in Figure 3.2. In contrast to the 2500 K simulation, where the C2 fraction is dominant and higher fractions are negligible, heavier molecules (C3 and higher) comprise about 50% of the products at 1000 K. Similarly, we observe that low-temperature propagation reactions involving H-abstraction by small radicals such as H, CH<sub>3</sub> and C<sub>2</sub>H<sub>5</sub> constitute the main consumption channel of unreacted alkanes, but at high temperatures unimolecular initiation through

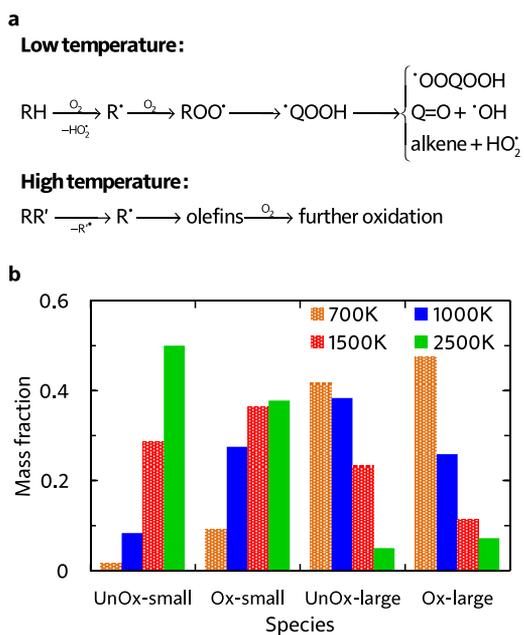
bond fission gains importance.

The relative stability of C–C and C–H bonds is also found to be temperature-dependent. Because a C–H bond is about 25 kcal/mol stronger than a C–C bond, unimolecular initiation at low temperatures only occurs through C–C dissociation; at high temperatures, considerable C–H dissociation is also observed, resulting in highly reactive free H atoms, in agreement with earlier high-temperature simulations of *n*-heptane pyrolysis.<sup>97</sup> A constant supply of free H radicals has a large impact on the overall reactivity of the system and the propagation rate, again illustrating the temperature-dependence of the pyrolysis mechanism. At low temperatures, C–H dissociation is only observed in radicals: ReaxFF predicts that C–H bonds vicinal to a radical site are about 50 kcal/mol weaker than those in alkanes (dissociation energies of ~50 and ~100 kcal/mol, respectively), significantly facilitating their dissociation. Especially the ethyl radical, which has a C–H bond dissociation energy of 45 kcal/mol, frequently decomposes into  $C_2H_4 + H$ .

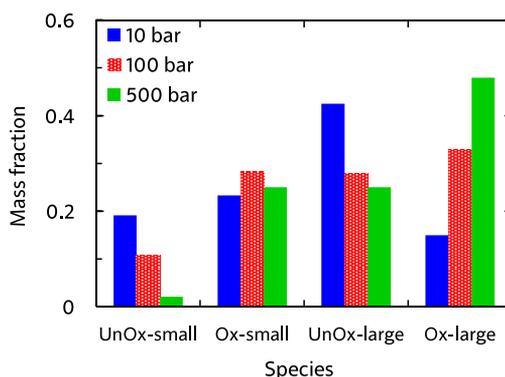
### 3.3.3 Oxidation

More complicated mechanisms are observed in the oxidation simulations, of which there are two distinct limiting cases, summarized in Figure 3.3a. In the low-temperature mechanism, the oxidation process is always initiated with hydrogen abstraction by an oxygen molecule. The subsequently formed alkyl radical combines with another oxygen molecule to form a peroxy radical  $ROO\cdot$  and further isomerization leads to a hydroperoxyalkyl radical  $\cdot QOOH$ , which can react through a variety of pathways. Additional H abstractions by  $O_2$  or reactive oxygen species from alkenes, radicals and carbonyl-containing compounds lead to the formation of compounds such as (conjugated) alkenes, ketenes and keto-hydroperoxides. At high temperatures, on the other hand, initial steps are essentially a pyrolysis process initiated by unimolecular C–C bond fission and subsequent  $\beta$ -scissions, forming primarily  $C_2H_4$  which is further oxidized in a later stage.

At intermediate oxidation temperatures, both mechanisms are at play: below 1500 K, alkanes are initiated by H abstraction but then easily break down into olefins, whereas from 1000 K and lower, C–C bond fission only rarely occurs in the initial oxidation stages. These temperature-dependent mechanisms are reflected by the product distributions of Figure 3.3b. High temperatures primar-



**FIGURE 3.3:** CVHD simulations of *n*-dodecane oxidation. (a) Limiting temperature-dependent initial mechanisms, and (b) products of the constant-density simulations at different temperatures. *Unox* species do not contain oxygen, whereas *Ox* do; *large* products are C3 or heavier.

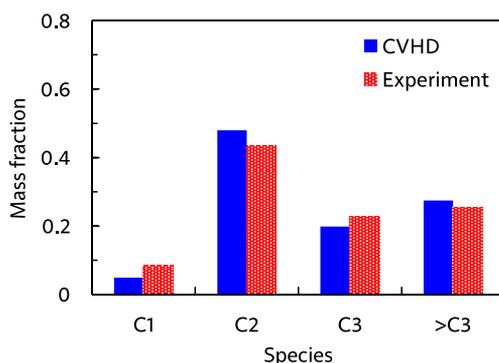


**FIGURE 3.4:** Pressure-dependence of *n*-dodecane oxidation at 1000 K. Presentation of the data is the same as in Figure 3.3.

ily produce  $C_2H_4$  and its oxidation products, whereas lowering the temperature suppresses dissociation events. In agreement with the findings of the pyrolysis simulations, alkyl radical  $\beta$ -scissions become less likely at lower temperatures, but the formed 1-alkenes are larger due to the relatively increased isomerization rate so that the mass fraction of produced hydrocarbons remains almost constant.

The temperature also has an impact on the formation of hydrogen peroxide and water. A first hydrogen atom transfer to  $O_2$  forms a hydroperoxyl radical,  $HO_2$ , which can subsequently either abstract another hydrogen atom and form  $H_2O_2$ , or transfer its hydrogen atom to another radical. The further reactivity of  $H_2O_2$  is strongly temperature-dependent, as it is found to be stable at low temperatures, whereas at high temperature, dissociation in two OH radicals occurs within a short amount of time. These highly reactive OH radicals can then carry out an additional hydrogen abstraction to form  $H_2O$ . Therefore, at low temperatures, the kinetically stable  $H_2O_2$  tends to accumulate whereas high temperatures favor the formation of OH and water. Indeed, at 700 K, the  $H_2O_2$  fraction accounts for ~17% of the non- $O_2$  oxygen atoms to be compared with ~14% in the  $H_2O$  fraction. At 1000 K, this ratio is already 10/20 and from 1200 K onwards, the  $H_2O_2$  fraction is negligible while the  $H_2O$  fraction contains about 30% of all reacted  $O_2$ . These observations are in agreement with conceptual models of low-temperature diesel engines.<sup>109</sup>

Combustion chemistry is also affected by pressure, and CVHD simulations can be used to investigate this effect. If the pressure-dependent reaction rate is



**FIGURE 3.5:** Comparison of experimental and simulated product distributions of *n*-dodecane pyrolysis at 1000 K.

proportional to  $p^n$  and, at constant temperature and assuming ideal gas behavior, the average reaction time  $\langle t \rangle \sim p/\text{rate}$ , the overall reaction order  $n$  can be determined by fitting  $\ln \langle t \rangle = m \ln p + \ln \langle t \rangle_{p=1}$ , in which  $m = 1 - n$ . This way, we obtained  $n = 2.07 \pm 0.07$ , indicating that the rate-determining step of the oxidation is of second order, most likely involving hydrogen abstraction. Average oxidation time scales ranged from 0.6 ms at 500 bar, to 45 ms at 10 bar. The pressure effect on the relative importance of uni- and bimolecular processes is also reflected by the product distribution, as depicted in Figure 3.4. Although this effect is less pronounced than the influence temperature has on the oxidation process, it can be seen that pyrolytic mechanisms are favored at low pressures, but suppressed in denser systems.

### 3.3.4 Comparison with experiments and unbiased MD

Our CVHD simulations also compare well with experimental results and existing kinetic models. The product distribution of the 1000 K pyrolysis process can be compared with a recent experimental study at the same temperature and is, as depicted in Figure 3.5, in good agreement with our results.<sup>110</sup> Moreover, the half-life of *n*-dodecane was found to be in the order of 20–40 ms, which compares well with the results in Table 3.1. The temperature-dependent oxidation mechanisms observed in CVHD simulations are also consistent with generally accepted models<sup>90,91</sup> and experiments.<sup>110</sup>

There exist some discrepancies between our simulations and oxidation experiments. While experimentally, an early pyrolytic stage is already observed at

**TABLE 3.2:** Kinetic parameters of *n*-dodecane pyrolysis and combustion as obtained from fitting apparent first order Arrhenius and Eyring equations.

	pyrolysis	combustion
Temperature range (K)	1000–1800	700–1800
$E_A$ (kcal mol <sup>-1</sup> )	70 ± 5	46 ± 1
$A$ (s <sup>-1</sup> )	5 × 10 <sup>15</sup> to 2 × 10 <sup>17</sup>	7 × 10 <sup>11</sup> to 3 × 10 <sup>12</sup>
$\Delta^\ddagger H$ (kcal mol <sup>-1</sup> )	68 ± 5	44 ± 1
$\Delta^\ddagger S$ (cal mol <sup>-1</sup> K <sup>-1</sup> )	12 ± 4	-8 ± 1

1050 K, our simulations suggest that this requires higher temperatures above 1200 K. This can be attributed to the high pressures in our constant density simulations, which will favor bimolecular over unimolecular reactions and thus a relative decrease of  $\beta$ -scissions over alkyl radical reactions with oxygen-containing species. Indeed, as shown earlier, lowering the pressure in our 1000 K CVHD simulation suppresses bimolecular reactions and gives rise to an early pyrolytic stage at lower temperatures than suggested by high-pressure simulations. Moreover, deficiencies of the force field can also reduce agreement with the experiment.

Finally, apparent first order rate constants for pyrolysis and combustion were computed from the C<sub>12</sub>H<sub>26</sub> and O<sub>2</sub> consumption rates, respectively. Two sets of kinetic parameters, collected in Table 3.2, were derived from these rate constants. By fitting a linearized Arrhenius equation

$$\ln k = \ln A - \frac{E_A}{k_B} \frac{1}{T} \quad \text{with} \quad k = \frac{1}{\langle t \rangle}, \quad (3.1)$$

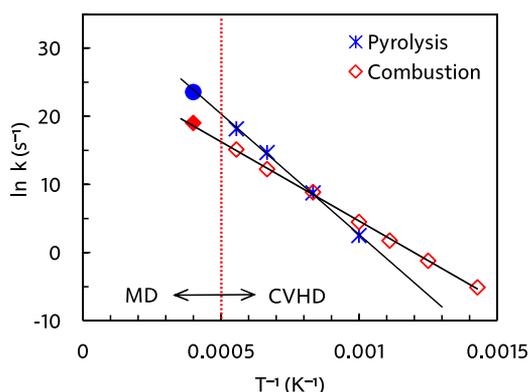
prefactors  $A$  and activation energies  $E_A$  were obtained, and activation enthalpies  $\Delta^\ddagger H$  and entropies  $\Delta^\ddagger S$  were calculated from the Eyring equation

$$k = \frac{k_B T}{h} e^{-\frac{\Delta^\ddagger G}{k_B T}}, \quad (3.2)$$

which can be linearized in  $1/T$  as

$$\ln \frac{hk}{k_B T} = \frac{\Delta^\ddagger S}{k_B} - \frac{\Delta^\ddagger H}{k_B} \frac{1}{T}. \quad (3.3)$$

The pyrolysis parameters are consistent with other ReaxFF pyrolysis studies of *n*-dodecane, in which values of  $E_A$  between 56 and 66 kcal/mol and  $A$



**FIGURE 3.6:** Arrhenius plots of the apparent first order rate constants of *n*-dodecane pyrolysis and combustion as obtained from CVHD simulations. Filled symbols at 2500 K are unbiased MD simulations that were not included in the fit.

from  $10^{15}$  to  $10^{16}$   $s^{-1}$  are found,<sup>96</sup> and with a unimolecular C–C dissociation as rate-determining step, as indicated by the positive entropy of activation. For combustion, the activation energy matches that of a hydrogen abstraction by  $O_2$ . Indeed, experimental barriers of hydrogen atom transfers from alkanes to  $O_2$  lie between 44 and 51 kcal/mol<sup>111</sup> and ReaxFF predicts a barrier of  $\sim 50$  kcal/mol for  $O_2$ -mediated hydrogen abstraction from methane.<sup>94</sup> The negative  $\Delta^\ddagger S$  value for combustion is also in line with a bimolecular mechanism. This means that hydrogen atom transfers to  $O_2$  are rate-determining at all temperatures, regardless of the different temperature-dependent initial reaction steps. Furthermore, as can be seen from the Arrhenius plots in Figure 3.6, the CVHD values are also consistent with unbiased MD simulations: extrapolation of the CVHD results to higher temperatures agree with the MD results, therefore further validating the application of CVHD to pyrolysis and combustion.

### 3.4 Conclusions

We have applied the CVHD method to pyrolysis and combustion of the *n*-dodecane model fuel. Owing to the unprecedented long time scale of our simulations, we were able to conduct the first explicit verification of temperature- and pressure-dependent pyrolysis and combustion mechanisms through direct atomistic simulations. Reaction pathways uncovered by CVHD simulations agree well with experiments and kinetic models and prove CVHD's ability to

extend and supplement chemical kinetic models. Moreover, these results show that a flexible accelerated molecular dynamics method such as CVHD can give access to the long-timescale dynamics of complex chemical processes, and how it can further extend the interpretive and predictive power of atomistic simulations by bridging the gap between theory and experiment.

### 3.5 Further applications of CVHD

The CVHD method has already been applied in research not directly related to this PhD project, applications that are briefly summarized here.

**Hydrogen etching of graphite.** In a collaboration with researchers from the Dutch Institute for Fundamental Energy Research (DIFFER), the interaction of graphite with a hydrogen plasma was investigated.<sup>112,113</sup> The time scale limitation of MD simulations has meant that previous MD studies of hydrogen atom etching of materials had to impose artificially high fluxes in order to observe any relevant mechanisms. By applying CVHD in between individual hydrogen impacts, Aussems *et al.* could tune the inter-impact time and, hence, the overall ion flux—the longest inter-impact times ( $> 1 \mu\text{s}$ ) corresponded to the conditions on experimentally feasible set-ups. The CVHD simulations revealed that at low fluxes, different C–C bond breaking mechanisms are at play, also significantly affecting the erosion yield. Smaller hydrocarbon species are ejected from the surface due to the switch from ion-induced etching to chemical erosion. Similarly, hydrogen atoms were found to have flux-dependent C–H interactions, going from sputtering and Eley-Rideal surface recombination at high fluxes, towards thermal processes as Langmuir-Hinshelwood recombination and H desorption upon reduction of the flux.

**Methanol oxidation on  $\text{V}_2\text{O}_5$ .** In previous MD studies of this process, the  $\text{V}_2\text{O}_5$  catalyst was isothermally kept at 650 K, while the methanol gas phase molecules were given a high temperature of 2000 K to be able to observe any chemistry within the MD time frame. Neyts & Bal applied CVHD with specific aim to study the effect of a plasma-supplied electric field at a more realistic temperature of 650 K for the *entire* system.<sup>114</sup> For the specific setup employed in the study, it had to be concluded that the

electric field had little impact.

**More pyrolysis.** Ashraf *et al.* performed ReaxFF-based MD simulations of the supercritical pyrolysis of fuel mixture (JP10/toluene and *n*-dodecane/toluene).<sup>115</sup> To justify the high temperatures (>2000 K) used in these simulations, however, some CVHD simulations at 1200 K were also performed. At least the qualitative differences in reactivity of the individual fuels were found to be preserved at this lower temperature, although the authors acknowledged that the overall chemistry would most likely be very different.

**Polymer degradation.** Arash *et al.* used CVHD to investigate the impact of water on the thermal degradation of polyamide 6,6 (PA 6,6).<sup>116</sup> Realistic temperatures as low as 1000 K were investigated, and kinetic parameters related to the degradation process could be extracted, consistent with experimental observations.

**Rate calculations** Some of CVHD's properties can be spliced back into its parent enhanced sampling methods. Fu *et al.* calculated gas-phase reaction rates with infrequent metadynamics to assess various types of CVs.<sup>117</sup> They found that simulations with the CVHD CV  $\eta$ , using bond lengths or SPRINT coordinates as local CVs, were more efficient than simulations that directly biased these local CVs, by an order of magnitude. These authors identified CVHD's dimensionality reduction as the key factor behind the method's efficiency, making it the best choice for high-demanding simulations based on *ab initio* force fields.

# A different approach to biasing

## 4.1 Towards even more challenging applications

The previous chapters have documented the remarkable successes of the CVHD method across a diverse range of problems. Nevertheless, a few areas of possible improvement can still be identified, and these must be addressed to give the method an even wider reach.

- Metadynamics-style dynamic bias deposition is in many cases quite inefficient.  $10^6$  time steps until transition (as observed for alkane pyrolysis initiation) are trivial to reach for modestly-sized classical systems, but utterly intractable once simulations become more expensive.
- Folding all dynamics into a single CV is great for efficiency, but also very restricting. In particular if some of the constituent local CVs are much stiffer than others, unbalanced or inefficient biasing could be observed. For example, if both bonded and non-bonded pairs are combined, the large fluctuations from the latter drown out the stiff, small-amplitude vibrations of the former.
- A mechanism to prevent transition state corruption is explicitly encoded into the CVHD CV, but still relies on a proper parametrization. Restricting the bias to small distortions is a safe choice, but will also result in limited biasing and poor performance.

It is clear that the application of CVHD to more complex phenomena described by more accurate *ab initio* force fields demands an extension of the method. By leveraging CVHD's inherent flexibility, some of the method's key weaknesses could perhaps be alleviated. The option to use different biasing

strategies has been available from its inception, and alternatives to the practical but sometimes unsatisfactory metadynamics bias can be easily integrated if available. Here, we will develop a variant of the CVHD method that uses the recently developed *variationally enhanced sampling* (VES) method<sup>118</sup> to generate its bias.

## 4.2 A variational approach to CVHD

### 4.2.1 Variationally enhanced sampling

In metadynamics, a bias potential is added to the system to enhance exploration of the CV space. Instead of obeying the true probability distribution  $P(\mathbf{s}) \sim e^{-\beta F(\mathbf{s})}$ , the free energy landscape is flattened, defined by a new distribution  $p(\mathbf{s})$ . Conventional metadynamics simulations completely fill the underlying FES and hence impose a uniform distribution  $p(\mathbf{s}) = c$ , whereas WT-metaD samples  $p(\mathbf{s}) \sim [P(\mathbf{s})]^{1/\gamma}$ . It would be interesting to design a bias potential  $\Delta V(\mathbf{s})$  to specifically sample a completely arbitrary *target distribution*  $p(\mathbf{s})$ , i.e., to limit sampling to specific regions of CV space. Such a bias potential should hence obey the relation

$$\Delta V(\mathbf{s}) = -F(\mathbf{s}) - \frac{1}{\beta} \ln p(\mathbf{s}), \quad (4.1)$$

up to an irrelevant additive constant. The extent to which a trial bias is able to produce the desired target distribution can be written as the functional

$$\Omega[\Delta V] = \frac{1}{\beta} \ln \frac{\int e^{-\beta[F(\mathbf{s}) + \Delta V(\mathbf{s})]} d\mathbf{s}}{\int e^{-\beta F(\mathbf{s})} d\mathbf{s}} + \int p(\mathbf{s}) \Delta V(\mathbf{s}) d\mathbf{s}, \quad (4.2)$$

which is minimized by the optimal bias of eq. (4.1). In other words, it is possible to use the *variational principle* to iteratively obtain a suitable bias.<sup>118</sup>

A practical choice of  $\Delta V$  is a linear expansion in some basis  $f_k(\mathbf{s})$ ,

$$\Delta V(\mathbf{s}) = \sum_k \alpha_k f_k(\mathbf{s}), \quad (4.3)$$

using the expansion coefficients  $\{\alpha_k\}$  as variational parameters. Common choices of basis functions are either polynomials of the Legendre or Chebyshev types (for nonperiodic CVs) or plane waves (if periodicity is required). In the

most general implementation of VES, the basis set consists of one-parameter functions—the actual bias expansion in a multidimensional CV space uses all possible products of the single-CV basis functions. Of course, nothing in the VES formalism enforces the use of a global linear expansion bias, and different forms have already been proposed for specific purposes.<sup>119,120</sup>

As  $\Omega[\Delta V]$  is based on expectation values, an iterative optimization of the bias demands some sort of dynamic sampling. To be precise, the optimization of  $\Delta V$  is based on the gradient and Hessian of  $\Delta V$  with respect to the expansion coefficients, which also entail the calculation of several ensemble averages on the biased potential. Hence, a VES simulation consists of periodic updates of the bias following short ( $\sim 1$  ps) sampling runs to estimate the required averages.<sup>118</sup>

The considerable freedom offered by the target distribution-based approach makes VES a potentially very powerful method. In particular high-dimensional CV spaces can be much more efficiently explored if sampling is limited to only the small subspace of interest.<sup>120</sup> Furthermore,  $p(\mathbf{s})$  need not be a predefined analytical function. Just like the actual bias potential, a target distribution can also be periodically updated—which is how VES implements the well-tempered target distribution.<sup>121</sup>

#### 4.2.2 Integrating VES in CVHD

Infrequent metadynamics allows for rate calculations and served as one of the inspirations for CVHD. A VES variant for similar purposes also exists, but is implemented by exploiting its variational nature. In this method, dubbed *variationally optimized free energy flooding*, the FES is filled up to a predefined cutoff value  $F_c$ .<sup>122</sup> As long as the free energy barrier is located above  $F_c$ , transition states remain uncorrupted, and rates can be extracted in a manner similar to infrequent metadynamics.

The flooding bias  $\Delta V$  is constructed as

$$\Delta V(\mathbf{s}) = v(\mathbf{s}) S(-v(\mathbf{s}) - F_c), \quad (4.4)$$

in which  $v(\mathbf{s})$  is the VES bias that is optimized to minimize  $\Omega[\Delta V]$  to satisfy eq. (4.1).  $S(x)$  is a switching function that turns off the bias when it becomes too large, and is usually taken to be of the Fermi type:

$$S(x) = \frac{1}{1 + e^{\lambda x}}. \quad (4.5)$$

Finally, the target distribution is dynamically computed based on an iteratively improving guess of  $F(\mathbf{s})$ , i.e.,  $F^*(\mathbf{s})$ ,

$$p(\mathbf{s}) = \frac{S(F^*(\mathbf{s}) - F_c)}{\int S(F^*(\mathbf{s}') - F_c) d\mathbf{s}'}. \quad (4.6)$$

By using this particular form of  $p(\mathbf{s})$ , one essentially enforces uniform sampling below  $F_c$ , and no biasing above. In practice, the flooding bias is first optimized through the iterative VES procedure, and then used as a static bias in a subsequent rate calculation.<sup>122,123</sup>

All in all, many aspects of CVHD are also included in the variational flooding formalism. VES-based methods can efficiently handle higher-dimensional CV spaces<sup>118</sup> and transition state protection can be natively—and explicitly—enforced.<sup>122</sup> Moreover, the basis set expansion (4.3) provides a more efficient route to a suitable bias than slow deposition of small Gaussian kernels. With the aim of developing a generic hyperdynamics method to study arbitrary chemical reactions, we can join aspects of CVHD and variational flooding as follows. We still use a CV of the form

$$\chi_t = \left( \sum_i^n \chi_i^p \right)^{1/p}, \quad (4.7)$$

and take as local CV  $s$  a commonly used<sup>30-32</sup> switching function of the inter-atom distance  $r_i$

$$s_i(r_i) = \frac{1 - (r_i/d)^n}{1 - (r_i/d)^m}. \quad (4.8)$$

The choice of  $d$  will depend on the nature of the atom pair, and should be very well-calibrated if different types of contacts have to be combined into a single  $p$  norm-based CV (4.7). The actual distortion  $\chi_i$  can then be simply taken as  $\chi_i = |s_i - s_i^{\text{ref}}|$ , where any  $s_i^{\text{ref}}$  is either 0 (i.e., non-bonded reference) or 1 (bonded reference).

To lift the stringent requirements on the choice of  $d$ , we propose to split up radically different interaction types—bonded or non-bonded pairs, light or heavy elements—into their own global CVs. Then, the FES is flooded up to a pre-chosen  $F_c$ . No safe distortion cutoffs have to be selected, and all interaction types will be biased equally in a natural fashion. In contrast to the standard flooding approach, time scale recovery is carried out on the dynamic bias, because each transition requires its resetting and re-optimization. In other words,

instead of first optimizing a suitable flooding potential and then using that as a static bias for rate calculations, we simply flood the FES and keep optimizing  $\Delta V$  until a transition is observed.

Only one user-selected choice remains, that is, a proper choice of  $F_c$ . For this, we define an initial cutoff  $F_{c,\text{init}}$ , that is gradually (every few ps) incremented. If  $F_{c,\text{init}}$  and its increments are sufficiently small the generated bias will always remain safe. In a way, they represent the target accuracy of the simulation— $F_{c,\text{init}}$  can be chosen to be a bit higher than strictly possible, sacrificing some accuracy for small barrier events of lesser importance.

### 4.3 Application

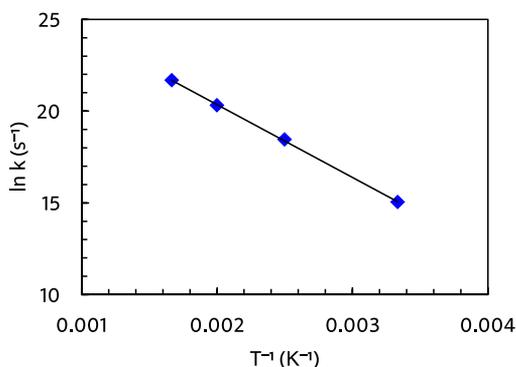
As a prototypical test, we considered the symmetric gas phase  $S_N2$  reaction of methyl chloride, i.e.,  $\text{ClCH}_3 + \text{Cl}^- \rightleftharpoons \text{Cl}^- + \text{CH}_3\text{Cl}$ : a common benchmark reaction that requires the simultaneous biasing of bond breaking and forming, and hence an excellent test case for our variational CVHD method. To assess its accuracy, repeated reactions were sampled in a temperature range of 300–600 K to compare its performance to infrequent metadynamics simulations of the same process.<sup>52,117</sup> As in these previous studies, energies and forces were described at the semi-empirical PM6 level,<sup>124</sup> as implemented in the Quickstep module<sup>125</sup> of the CP2K 4.1 package.<sup>126</sup> A single methyl chloride molecule and chloride ion were placed at the center of a  $20 \times 20 \times 20 \text{ \AA}^3$  nonperiodic box in which electrostatics were treated by a wavelet-based solver.<sup>127</sup>

The equations of motion were integrated with a time step of 0.5 fs. At all temperatures, the system was first equilibrated for 1 ps with a stochastic global thermostat,<sup>128</sup> followed by five independent NVT production runs of 1 ns each. To prevent the ion and molecule from moving too far apart, both C–Cl distances were held below 3 Å by harmonic restraints. Two CVs were used to describe bonded and non-bonded C–Cl interactions, respectively, which were both described by a switching function of the type (4.8) with  $d = 2.5 \text{ \AA}$  and  $m = 2n = 12$ . The bias was expanded as a linear combination of Chebyshev polynomials up to 12th order and updated every 1 ps with a step size of 0.05 kcal/mol.  $F_{c,\text{init}}$  was 5 kcal/mol, with  $F_c$  being incremented by 1 kcal/mol every 10 ps until transition. Transitions were detected when any of the CV values exceeded 0.9.

The average reaction times at each temperature were computed from a fit to

**TABLE 4.1:** Rate calculations for the symmetric gas phase  $S_N2$  reaction of methyl chloride from variational CVHD. Average reaction times (in seconds) are reported at different temperatures (in K). The total number of observed events,  $p$  value from the Kolmogorov–Smirnov analysis, and average boost factor are also reported.

temperature	$\langle t \rangle$	$n_{\text{event}}$	$p_{\text{KS}}$	boost
300	$2.94 \times 10^{-7}$	58	0.15	$3.42 \times 10^3$
400	$9.69 \times 10^{-9}$	70	0.19	$1.36 \times 10^2$
500	$1.51 \times 10^{-9}$	86	0.29	$2.60 \times 10^1$
600	$3.86 \times 10^{-10}$	113	0.21	$8.72 \times 10^0$



**FIGURE 4.1:** Arrhenius plot for the symmetric gas phase  $S_N2$  reaction of methyl chloride.

an exponential escape time distribution as used in the Kolmogorov–Smirnov test developed for infrequent metadynamics,<sup>50</sup> and collected in Table 4.1. Despite not being designed for continuous biasing, the variational flooding algorithm appears to be able to guarantee a safe bias even while unconverged, as evidenced by the fact that  $p_{\text{KS}} > 0.05$  in all cases, meaning that the escape time distribution is Poissonian (and uncorrupted).<sup>50</sup> The performance of variational CVHD appears to be quite competitive, too: while infrequent metadynamics with two bond distance CVs can provide a boost factor of 60 at 300 K for this particular system,<sup>117</sup> we obtain a value of  $3.42 \times 10^3$ . If multiple CVs are required for biasing, variational CVHD can hence be considered to be more efficient than its metadynamics-based variant. The boost also deteriorates with increasing temperature, as expected.

The average reaction times at different temperatures can be used to construct

an Arrhenius plot, as depicted in Figure 4.1. From a linear fit of  $\ln k = \ln\langle t \rangle^{-1}$  versus  $T^{-1}$  an activation energy of  $E_A = 7.9 \pm 0.5$  kcal/mol and pre-exponential factor  $A = 1.93e^{\pm 1.1}$  THz can be obtained (error bars reflect the 95% confidence level). These values agree well with estimates from infrequent metadynamics (on the same PM6 PES), although  $E_A$  falls somewhat on the high end of the known range ( $\sim 6.5$ – $8.0$  kcal/mol).<sup>117</sup> Combined with the fairly low values of  $p_{KS}$ , that usually exceed 0.5 in the competing infrequent metadynamics simulations, some minor transition state corruption cannot be ruled out. However, the average transition times obtained from our simulations are of excellent quality: other estimates at 300 K fall in the range of 183 to 312 ns, consistent with our value of 294 ns.

## 4.4 Conclusions

The inherent flexibility of the CVHD method makes it easy to incorporate novel biasing strategies. A variational version of CVHD can be achieved by replacing the metadynamics-style bias with a VES flooding potential. By exploiting the variational principle, several user-supplied parameters can be replaced or made less critical, and an overall more flexible method is obtained.



## **PART II**

Untangling the experiment through  
modeling: the case of surface  
charges in plasma catalysis



# Computational models of charged surfaces

## 5.1 The plasma-catalytic enigma

### 5.1.1 The virtues of plasma catalysis

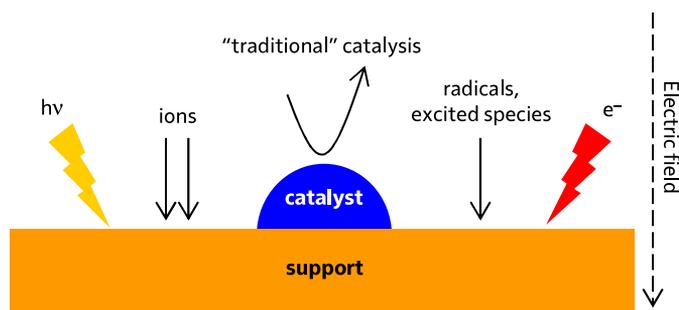
A plasma or gas discharge is a partially ionized gas, consisting of free electrons, ions, and neutrals (molecules, radicals and excited species), which can all interact with each other, giving rise to a “rich” and reactive plasma environment.<sup>129</sup> In plasma catalysis, a heterogeneous solid catalyst is integrated in a plasma to obtain a technique that combines the reactivity of the two, potentially resulting in a better efficiency or selectivity than would be accessible with the individual approaches. Plasma catalysis is a rapidly growing research domain, with as hugely important envisaged application the efficient conversion of greenhouse gases into value added chemicals. The key strength of plasma catalysis can be summarized as follows:

- Plasmas are in a far-from-equilibrium state at relatively low temperatures, typically in the range 300 K-1000 K. The combination of reactivity, far-from equilibrium state and low temperature operation enables plasma-based processes to be operated at much milder conditions than traditional thermal methods. It also means that a plasma-exposed catalyst is not necessarily subjected to the scaling laws<sup>130</sup> that limit the performance of traditional thermal catalysis, and thermodynamically unfavorable conversions can be made to take place.

---

The results presented in this chapter were published in:

K. M. Bal and E. C. Neyts. Modelling molecular adsorption on charged or polarized surfaces: a critical flaw in common approaches. *Phys. Chem. Chem. Phys.* **20**, 8456 (2018).



**FIGURE 5.1:** Possible plasma-catalyst cross-interactions, as seen from the perspective of the catalyst.

- In many cases, synergistic effects have been claimed, i.e., the conversion, yield, energy efficiency or selectivity is observed to be greater than the sum of pure plasma processing of the gas and pure thermal catalysis.<sup>131–134</sup>
- The plasma is driven by electrical power, making plasma technology much more flexible and easier to switch on or off than its thermal counterparts. This property fits well with the current trend towards electrification of our society, where fossil fuel-based thermal energy is being replaced by electricity from renewable sources.<sup>135</sup> The fluctuating supply of wind and solar energy demands significantly higher flexibility from chemical industry.

The mechanisms underpinning this apparent efficiency are not fully understood: they must be unraveled to achieve a better understanding of the process and optimize its performance. The key characteristic—and key difficulty—that sets plasma catalysis apart from pure plasma-technological or catalytic approaches is the presence of a strong cross-interaction between the plasma and the catalyst surface, mutually changing each other’s properties (see Figure 5.1).

### 5.1.2 Constructing models of plasma-catalyst interactions

From the perspective of the catalyst, the impact of a plasma can be considered to be perturbation of the catalytic chemistry which, on itself, is already a very complex process with a massive number of chemical and physical degrees of freedom. To untangle all these influencing factors, a “bottom-up” approach based on theoretical atomistic calculations is ideally suited to study the role

of the chemical building blocks that make up the overall catalytic process.<sup>136</sup> For traditional catalytic approaches to CO<sub>2</sub> activation, for example, this kind of incrementally improved understanding of increasingly complex technologies has already been extensively demonstrated in the literature, mostly based on density functional theory (DFT) calculations.

Theoretical studies of traditional catalytic processes are strictly structure-activity focused. That is, the performance of catalyst in a specific application is assumed to be dictated only by its composition and morphology, so that calculations are applied to evaluate and screen the properties of current materials and new candidates in a systematic manner. This has been rather extensively achieved for the CO<sub>2</sub> conversion processes, through extensive characterization of the energetic and kinetic parameters of a variety of simple catalyst models, such as flat transition metal surfaces<sup>137-144</sup> (both pure metals and alloys) and oxide single crystals.<sup>145-153</sup> However, realistic industrial applications (as well as plasma catalysis<sup>154</sup>) typically do not employ “pure” materials, but rather *supported* metal particles deposited on some other material. Naturally, the activity of such catalysts depends on a significantly larger set of interactions and cross-interactions between reactants, intermediates, products, metal catalyst, and support.

Models can again provide valuable insights, but with increasing model complexity computational work becomes scarcer, and just a few material combinations have been already studied.<sup>155-158</sup> Yet, these studies have led to the conclusion that the catalyst-support interface plays a significant role in the catalytic activity of the metal, and that reaction mechanisms on a supported cluster can be quite different from those on a pure metal catalyst.<sup>159</sup> The as such obtained insight from incrementally more complex models highlights the power of computational approaches to increase our understanding of catalytic processes.

In principle, plasma catalysis can be treated as another layer of complexity that is added to the computational model of the catalyst, one that goes beyond standard scaling laws and structure-activity relations. Plasma effects can be introduced separately, so as to disentangle the effect of the various mechanisms through which the plasma can interact with the catalytic process. This way, the effect of phenomena such as plasma-generated radicals, excited molecules, ions, photons, electrons and electric fields can be studied in isolation to assess their individual impact and relative importance.<sup>160</sup> In fact, most experimental

studies of plasma catalysis currently primarily treat the process as a black box, and focus on improving conversion or energy efficiency. Detailed experimental insights in the atomic-level aspects of the process are sorely lacking, which gives modeling a unique opportunity to dramatically advance the understanding of the plasma-catalytic process.

From the comparatively extensive simulation literature on plasma-enhanced material growth and surface modification, it is already known that plasma-supplied radicals are an important source of interesting chemistry. Therefore it is perhaps not surprising that the few computational investigations of potential plasma-catalytic effects have also introduced the effect of a plasma radical flux on the catalytic process. Using DFT calculations<sup>161,162</sup> and molecular dynamics simulations,<sup>81</sup> it was confirmed that gas phase plasma activation of inert gasses such as CH<sub>4</sub> leads to improved chemisorption, which lowers the temperature threshold for plasma catalysis. Moreover, a high surface coverage of plasma-generated radicals can significantly modify the activity of the catalyst towards CO<sub>2</sub> activation.<sup>163,164</sup> In other words, these limited initial studies already explicitly show that the performance of a plasma-catalytic process is not purely determined by traditional scaling laws or structure-activity relationships. The presence of a plasma could therefore allow for conversions not accessible by standard catalytic approaches.

### 5.1.3 Understanding the effect of charging

Of the many other possible plasma-surface interactions, perhaps the most intriguing is the ability of a plasma to modify the electronic structure of the catalyst through charging. All surfaces exposed to a gas discharge accumulate a negative charge due to the influx of plasma-supplied electrons, which is much larger than the influx of ions. Although physical models<sup>165</sup> and experiments<sup>166-169</sup> suggest that these surface charges can be quite substantial and long-lived, little to nothing is known about their effect on the chemical properties of the catalysts. Nevertheless, this effect can be expected to be important because catalytic bond breaking and formation processes are governed by the flow of electrons to and from the surface. Again, a plasma effect might be able to tune catalyst properties beyond the standard correlations explored in catalyst development.

Because charging is a fully reversible process that does not modify the cata-

lyst's physical structure, a recent set of plasma-catalytic experiments is particularly intriguing: the plasma-catalyst synergistic effect is found to also be fully reversible, i.e., no permanent plasma-induced chemical or physical modification of the catalyst is observed,<sup>170</sup> which suggests that surface charging could indeed play a role in this process. Since no direct experimental work in this direction has been carried out—and a controlled set-up to isolate the charge effect on the surface chemistry is difficult to achieve—computational approaches must be applied to gauge its impact on the plasma-catalytic process. Computational approaches to charged catalysts in general are, however, rather challenging, and have mostly tackled catalyst charging in an indirect fashion. For example, charged electrocatalysts were modeled as slabs in contact with a partially ionized aqueous phase,<sup>171,172</sup> and charge effects in photocatalysts could be approximated through the introduction of defects or dopants.<sup>173,174</sup>

Only a few computational studies have attempted to directly model the effect of a surface charge. Here, the focus has mostly been on reversible gas adsorption: through addition or removal of electrons, the (relative) affinity of a nanomaterial for gas molecules can be changed, implying that charging can effectively be used as a switch in gas capture, storage, and separation technologies.<sup>175</sup> DFT calculations consistently predict that extensive charging greatly modifies the selective adsorption of gasses such as CO<sub>2</sub> or H<sub>2</sub> on nanomaterials, i.e., hexagonal boron nitride (*h*-BN)<sup>176</sup>, graphitic carbon nitrides<sup>177–182</sup>, MXenes,<sup>183</sup> doped graphene or carbon nanotubes,<sup>184–187</sup> borophene,<sup>188,189</sup> and nanocomposites of these materials.<sup>190,191</sup> Similar switchable properties have also been demonstrated for electric fields in reversible gas adsorption<sup>192–195</sup> and processes relevant to catalysis.<sup>196–198</sup> Hence, this type of studies predicts that electrically switchable interfacial properties will play a major role in the development of new “green” technologies, where reversible gas capture and activation methods are required for carbon capture and sequestration, and the transport and storage of hydrogen gas. Consequently, charge effects in plasma catalysis can also be anticipated to be significant.

In order to make accurate predictions of charging-related material properties, the employed computational approaches must be adequate. Typically, models of adsorption and reactions on solid surfaces use small semi-infinite slabs, with periodic boundary conditions approximating the macroscopic surface. However, properties of systems with excess charges or large polarization are difficult

to properly converge.<sup>199</sup> Furthermore, traditional computational approaches to charged periodic simulation cells have been found to be inadequate for inhomogeneous systems.<sup>200</sup> Indeed, these issues are well-known for charged defects in solids. A proper model of the bulk solid requires the application of periodic boundaries, but also introduces artificial long-ranged Coulomb interactions between the periodic images of the localized charges. Moreover, the inclusion of an implicit neutralizing background charge introduces unphysical contributions to the total energy, rendering a straightforward calculation of the defect formation energy impossible. For this situation, a number of techniques has been devised that allow to eliminate any spurious electrostatic interaction, i.e., to treat the defect as isolated, embedded in a semi-infinite neutral matrix.<sup>201–203</sup>

Calculation of adsorption on charged surfaces poses a very different challenge. Indeed, because all periodic cells of interest—containing the pristine surface slab and adsorption complex, respectively—carry the same net charge, the energetic term originating from an implicit background charge is eliminated. Furthermore, the surface is assumed to be homogeneously charged, meaning that long-range interactions along periodic boundaries are a physically justifiable aspect of the model. However, *full* periodic boundary conditions (as usually applied) introduce an unphysical interaction of the material with itself, normal to its surface. In other words, the calculation cell dimensions along the surface directions have a physical meaning as they are a function of the material lattice parameters, whereas the edge length along the surface normal represents an arbitrary degree of freedom. To avoid spurious interactions from surface dipoles or charges, correction schemes exist, but the contemporary literature on electrochemically switchable adsorbents and catalysts does not adopt any specific corrective measures. Rather, sizable vacuum regions—as is also common practice in the study of nonpolar or neutral surfaces—have been claimed to be sufficient without explicit confirmation of the adequacy of this approach.<sup>176–197</sup> In other words, the comparatively new field of charge- and field-tunable surface chemistry has not yet adopted a rigorous, consistent and properly validated standard computational procedure.

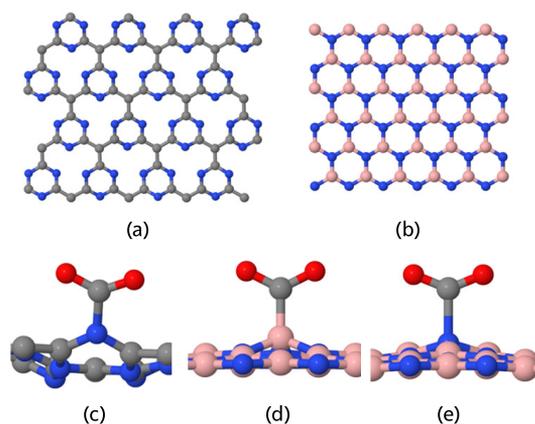
The challenging problem of plasma-catalyst interactions should only be tackled if the proposed methodology is properly validated on comparatively simple problems. Therefore, we first revisit three prominent applications of switchable CO<sub>2</sub> capture from the literature as representative case studies for the modeling

challenges posed by charged or polarized surfaces. Specifically, we consider charge-enhanced adsorption on  $g$ -C<sub>4</sub>N<sub>3</sub> (a half metal) and  $h$ -BN (a wide-gap semiconductor) as well as electric field-tuned adsorption, again on  $h$ -BN. We attempt to reproduce previously reported results for these systems, and investigate to what extent theoretical predictions are dependent on the employed computational model.

## 5.2 Methodology

All DFT calculations were carried out with the Quickstep module<sup>125</sup> in the CP2K 4.1 package<sup>126</sup> within the Gaussian and plane wave (GPW) framework,<sup>204</sup> employing Goedecker-Teter-Hutter (GTH) pseudopotentials<sup>205,206</sup> for the core-valence interactions and a polarized double- $\zeta$  (m-DZVP) basis set<sup>207</sup> to expand the Kohn–Sham valence orbitals. An auxiliary plane wave basis set defined by a cutoff of 1000 Ry was used to expand the electron density. Exchange and correlation were treated with the PBE functional,<sup>208</sup> supplemented by Grimme’s D3 dispersion correction<sup>209</sup> in its Becke-Johnson damping form.<sup>210</sup> Atomic partial charges were calculated by the self-consistent Hirshfeld-I scheme.<sup>211</sup> All energies are reported without thermal or zero-point energy corrections, and adsorption energies on the nanosheet are defined as  $E_{\text{ads}} = E_{\text{mol+sheet}} - E_{\text{mol}} - E_{\text{sheet}}$  so that a negative  $E_{\text{ads}}$  corresponds to a stable adsorption configuration. Two different periodicities were considered: fully periodic calculations employed the default CP2K Poisson solver for electrostatics, whereas the Martyna-Tuckerman (MT) solver<sup>212</sup> allowed to impose two periodic (along the surface,  $XY$  plane) and one nonperiodic direction (along the surface normal,  $Z$  direction). Unless noted otherwise, geometries were optimized in cells with a nonperiodic  $Z$  length of 25 Å and used in single point energy calculations in all other cells.

The MT solver is a reciprocal space-based method to calculate electrostatic energies that also allows treating both isolated and periodically replicated systems within the same formalism. It can therefore be applied to systems with mixed boundary conditions. The MT solver has one specific limitation that must be kept in mind: the cell length along any nonperiodic axis must be at least two times the size of the charge distribution (including ions and electrons) of the system along this direction. For all of our systems, this means that very



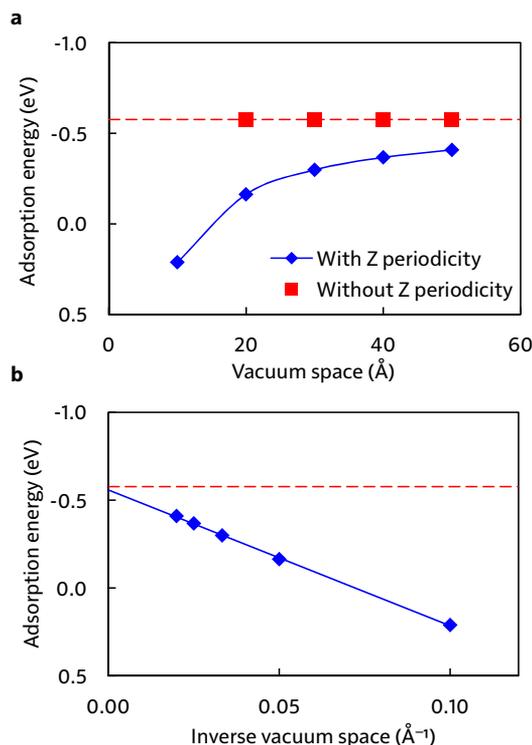
**FIGURE 5.2:** Investigated  $\text{CO}_2$  adsorption configurations on 2D nanomaterials. The two considered adsorbents are (a) a  $4 \times 4$   $g\text{-C}_4\text{N}_3$  sheet and (b)  $6 \times 6$   $h\text{-BN}$ . The molecule is adsorbed on (c)  $4e^-$  charged  $g\text{-C}_4\text{N}_3$ , (d)  $2e^-$  charged  $h\text{-BN}$ , and (e)  $h\text{-BN}$  subjected to an electric field of 0.04 a.u., pointing downward. All configurations have been previously described in the literature. Atom color codes: gray, carbon; blue, nitrogen; pink, boron; red, oxygen.

small vacuum spaces ( $10 \text{ \AA}$  or lower) cannot be used.

Because all studied systems have already been characterized by DFT calculations, adsorption structures and relevant charge or field conditions can initially be directly taken from the literature. For charged  $g\text{-C}_4\text{N}_3$ , a charge state equivalent to  $1e^-$  in a  $2 \times 2$  supercell, as used in the literature,<sup>177</sup> is considered; to compensate for the lack of  $k$  point sampling beyond the  $\Gamma$  point in our calculations we use a larger  $4 \times 4$  supercell (Figure 5.2a), with a  $4e^-$  charge. An  $h\text{-BN}$  sheet was modeled as a  $6 \times 6$  supercell (Figure 5.2b), either carrying a  $2e^-$  charge<sup>176</sup> or subjected to an electric field of 0.04 a.u. (about  $2 \text{ V \AA}^{-1}$ ).<sup>194</sup> In all cases, the considered charge levels or field strength have been reported to lead to a shift of the  $\text{CO}_2$  adsorption mode from physisorption to chemisorption. The relevant adsorption structures for these systems are depicted in Figure 5.2c–e.

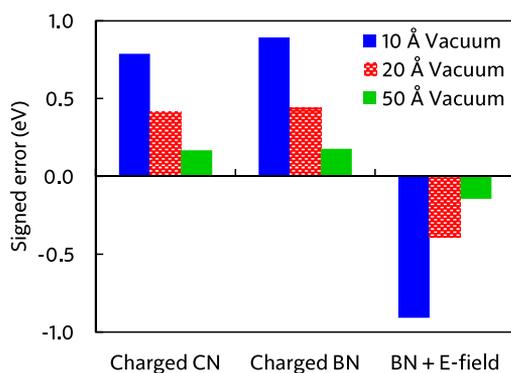
### 5.3 Results

The only “free” computational parameter with no true physical analogue is the vacuum space between the repeated periodic images of the adsorbent material,



**FIGURE 5.3:** Convergence dependence of CO<sub>2</sub> adsorption energies on charged *g*-C<sub>4</sub>N<sub>3</sub> on the vacuum size. (a) Dependence of the adsorption energy on the vacuum space for both fully as partially periodic cells. (b) Extrapolation to infinite vacuum. Details can be found in the text.

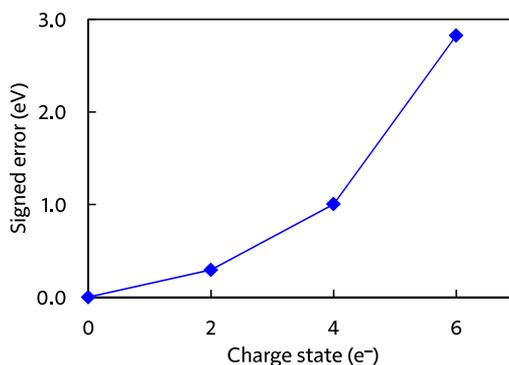
i.e., the *Z* length of the simulation cell. Therefore, we first perform adsorption energy calculations using different sizes of the vacuum region, for both the pristine sheet and the adsorption complex, to assess its impact. In Figure 5.3a, these energies are plotted for CO<sub>2</sub> adsorption on charged *g*-C<sub>4</sub>N<sub>3</sub>, demonstrating the huge impact of the vacuum region. Depending on the imposed intersheet spacing, the interaction of CO<sub>2</sub> with the sheet is either predicted to be very favorable (with a strongly negative adsorption energy), or not binding at all, which means that not only quantitative predictions, but even qualitative trends are sensitive to the vacuum size. Furthermore, no clear convergence is observed, even when using a very large cell length of 50 Å, which is significantly larger than the commonly reported vacuum sizes of 15–25 Å. No convergence studies of this type have been previously reported in the literature.



**FIGURE 5.4:** Absolute signed errors for adsorption energies in all systems for various vacuum sizes. As a reference, converged energies obtained from partially periodic calculations are used.

One way to avoid the problem of periodic images interacting with each other is to use specialized Poisson solvers for electrostatics, that allow to mix periodic and nonperiodic boundaries. We repeat the adsorption energy calculations in different cell sizes, but now apply the MT solver to impose nonperiodic boundaries along the  $Z$  direction. This setup gives rise to adsorption energies independent of the cell dimensions, as expected (Figure 5.3a). If the adsorption energy obtained in a partially nonperiodic cell is assumed to reflect the properly converged solution, it can be inferred that calculations in fully periodic cells consistently underestimate this property, even when large vacuum sizes are used. This can be understood as follows: significant charge transfer toward the adsorbed molecule takes place upon adsorption, and the region carrying a net charge becomes “thicker” compared to the pristine “clean” nanosheet. The repulsive interaction between periodic images of clean sheets can be expected to be proportional to  $L^{-1}$ , with  $L$  the vacuum spacing along the  $Z$  direction. The effective separation between sheets carrying adsorbed molecules, then, is reduced. For large  $L$ , the as such introduced error will be simply proportional to  $L^{-1}$  and extrapolating  $L \rightarrow \infty$  should give adsorption energies free of any error. Indeed, as depicted in Figure 5.3b, extrapolated adsorption energies from fully periodic cells are identical to the solution obtained in a cell with partial periodicity, also confirming the appropriateness of the latter.

The same treatment can also be applied to the other systems—some key results are summarized in Figure 5.4. Similar to  $g\text{-C}_4\text{N}_3$ , the effect of charge-



**FIGURE 5.5:** Absolute signed errors for adsorption energies on *h*-BN in different charge states, in a fully periodic box with a vacuum of 30 Å. As a reference, energies obtained from partially periodic calculations are used.

enhanced adsorption on *h*-BN is significantly underestimated in fully periodic cells: the commonly applied vacuum space of 20 Å gives rise to a significant error of 0.44 eV, and even an impractically large cell edge length of 50 Å is still off by 0.18 eV. Furthermore, the error increases with the magnitude of the surface charge, as depicted in Figure 5.5. Evidently, such a large charge-dependent error renders computational studies of charge response properties<sup>213</sup> useless.

A somewhat different picture arises for electric field-modulated adsorption. Here, fully periodic cells tend to overstabilize adsorbed CO<sub>2</sub> (Figure 5.4). This can be explained by the difference in interaction across periodic images, which is not a charge-charge repulsion, but an attractive dipole-dipole attraction. Electric field-enhanced adsorption of CO<sub>2</sub> relies on a large polarization effect from the electric field to push electrons from the nanosheet into the molecule, which causes the adsorption complex to have a larger induced dipole, and hence stronger interactions across periodic boundaries than the pristine nanosheet, where electrons have nowhere else to go. Nevertheless, energies from the fully periodic cells can also be extrapolated to the partially periodic result for this system.

While energetics are strongly affected by the treatment of periodicity, electronic structures of the considered systems are fairly insensitive to this. For example, the charge transfer to CO<sub>2</sub> (as an indication of the relevant electronic structure) on charged *h*-BN is 1.187e<sup>-</sup> in a converged calculation, and 1.175e<sup>-</sup> in a fully periodic system with a modest vacuum size of 20 Å. Similar observations

for this transition can be made for *h*-BN subjected to an electric field ( $0.523e^-$  to  $0.554e^-$ ) and charged *g*- $C_4N_3$  ( $0.416e^-$  to  $0.383e^-$ ). These highly similar electron densities also imply a similar ground state structure, which indeed appears to be the case: reoptimizing the adsorption complex on charged *g*- $C_4N_3$  in a periodic cell with 25 Å vacuum yields a structure with an N-CO<sub>2</sub> bond of 1.53 Å, as compared to the 1.54 Å in the cell with partial periodicity. On *h*-BN we find somewhat larger deviations, with the N-CO<sub>2</sub> bond in electric field-enhanced adsorption changing from 1.71 Å to 1.82 Å upon reducing periodicity.

Unfortunately, we have been unable to cross-check our results with some of the relevant previous studies<sup>177,194</sup> because these are inherently unreproducible. Rather than reporting exact box dimensions, employed vacuum spacings are consistently reported as being “larger than” a certain value, which is a very troubling carelessness that appears to permeate the literature (see, e.g., any publication by Smith and coworkers). Because results can be “tuned” by changing the vacuum spacing, such vague language is not acceptable: it makes independent validation and comparison across materials, molecules, and computational codes impossible. Thankfully, we *could* find well-defined computational results for charged *h*-BN, for which a vacuum distance of 15 Å was used.<sup>176</sup> For that particular setup, we find a CO<sub>2</sub> adsorption energy of  $-2.45$  eV, in good agreement with the previously reported value of  $-2.32$  eV. Compared to fully converged cell sizes that produce an adsorption energy of  $-3.04$  eV, these fully periodic results are in error by  $\sim 0.6$  eV, or 25%.

## 5.4 Conclusions

We have shown that conventional computational approaches to charge- and electric field-controlled gas capture, storage and conversion processes can give rise to significant artifacts due to incomplete convergence of interlayer interactions across periodic boundaries. These errors appear to affect the bulk of the literature on the subject.<sup>176–197,213</sup> In future work, fully converged adsorption energies can be obtained through one of two possible strategies: extrapolation of fully periodic calculations to infinite vacuum, or adoption of specialized Poisson solvers that also allow for a nonperiodic direction along the surface normal. The former approach can be easily realized in any of the standard periodic DFT codes, whereas the latter will also be able to self-consistently yield converged

minimum energy configurations and is hence preferred. We stress that in general, calculations of fully periodic systems with large vacuum can reasonably capture charge-modulated adsorption phenomena, but quantitative predictions will be off.

More generally, these results show the difficulty of constructing physically sound models. Even if one consciously limits the scope of the model—only pure charging effects, no embedding of the catalyst in a solvent or plasma, no defects, only 2D materials, etc.—plenty of pitfalls remain. Within the stylized nature of the model, with little direct correspondence to real-life problems, it is even more important to at least ensure *internal* consistency, and remain critical of conventional wisdom. Only when we get the simplest of models right, we can build towards bigger and better things.

Armed with these new insights, we can now move to the construction of meaningful models of plasma-charged catalysts.



# Catalytic effects in plasma catalysis

## 6.1 Surface charging and plasma catalysis

A plasma catalysis reactor is commonly created by packing a plasma reactor with catalyst beads, which are usually made of an oxide support material on which transition metal particles are deposited. At an atomistic level, the catalytic process is dependent on the morphology and chemical composition of the catalyst. While a negative surface charge is known to change the chemical characteristics of a material, it is unclear to what extent this effect is influenced by the nature of the material, which raises a number of questions. Is the effect of plasma charging the same for all catalysts? Which catalytic processes are affected by charging, and how? Is it possible to control this effect? Is the magnitude of these effects relevant for typical experimental conditions? Can charging help explain the differences between thermal and plasma catalysis?

In a first attempt to answer these questions, we construct a number of model systems of common plasma-catalytic experiments, i.e., CO<sub>2</sub> splitting by different oxide-supported transition metal catalysts. In particular, atomically dispersed Ti, Ni and Cu on a  $\gamma$ -Al<sub>2</sub>O<sub>3</sub> (110) surface are considered in order to (1) characterize the structure of single atom catalysts on Al<sub>2</sub>O<sub>3</sub> and (2) investigate the CO<sub>2</sub> reduction ability of these catalysts and the dependence of their chemical properties on the nature of the metal. Besides being a very promising class of materials,<sup>214,215</sup> single atom catalysts also allow us to “purify” the model from the structural complexity of larger supported clusters, models of which

---

The results presented in this chapter were published in:

K. M. Bal, S. Huygh, A. Bogaerts and E. C. Neyts. Effect of plasma-induced surface charging on catalytic processes: application to CO<sub>2</sub> activation. *Plasma Sources Sci. Technol.* **27**, 024001 (2018).

have many more degrees of freedom and therefore require somewhat arbitrary choices of cluster size, structure and orientation.<sup>155-158</sup> For this reason, using a model based on single metal atoms allows for a fairer and clearer comparison of different catalytic transition metals, although preliminary tests indicate that our general conclusions are unaffected by cluster size.

## 6.2 Methodology

### 6.2.1 Model systems

The same GPW/PBE-D3/m-DZVP level of theory as in Chapter 5 was used, but with a slightly higher plane wave cutoff of 1200 Ry for improved accuracy. Calculations were carried out on a slab of the  $\gamma$ -Al<sub>2</sub>O<sub>3</sub> structure proposed by Digne *et al.*<sup>216</sup> The (110) surface was modeled as a 2 × 2 supercell containing 240 atoms, corresponding to six layers of which the bottom two were kept fixed at their bulk positions. The simulation cell dimensions were 16.1606 × 16.8106 × 40 Å<sup>3</sup>. Partial periodicity was enforced with the Martyna-Tuckerman Poisson solver;<sup>212</sup> calculations involving isolated atoms or molecules were also carried out in these cell sizes. The surface exposes both coordinatively unsaturated Al and O atoms. Tri- (Al<sub>III</sub>) or tetracoordinate Al (Al<sub>IV</sub>) atoms provide Lewis-acidic sites, whereas di- (O<sub>2</sub>) and tricoordinate (O<sub>3</sub>) surface atoms are Lewis basic.<sup>217</sup> Although the (110) surface termination is the most common, it is not stable in its “dry” form, which is why a hydrated variant was also considered in this work (structure s1a from the literature<sup>217</sup>) containing 4 adsorbed water molecules, corresponding to a density of about 3 OH nm<sup>-2</sup>. This surface is the most stable adsorption configuration of a single adsorbed water molecule per unit cell, which is dissociated into an OH group adsorbed on the Al<sub>III</sub> site and a proton bonded with an O<sub>2</sub> atom. Comparison of the two surfaces allows assessing the impact of adsorbed water on the properties of the Al<sub>2</sub>O<sub>3</sub> support.

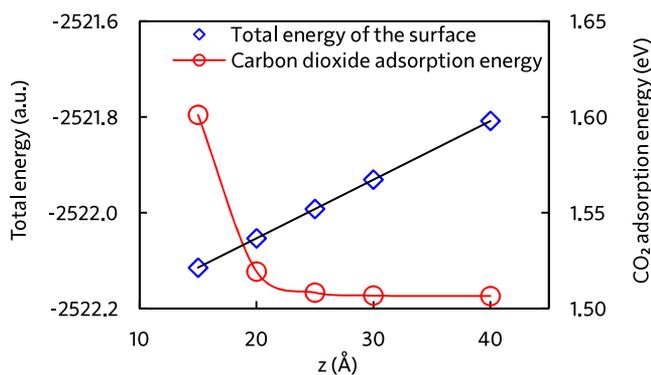
Unless noted otherwise, the abovementioned PBE-D3 based methodology was employed for all calculations, but a small subset of structures was re-optimized using different exchange-correlation functionals in order to assess the reproducibility of our results and their dependence on the chosen approximations. A more detailed description of all cross-checks is given in Section 6.5. We find that the sensitivity of our results on the choice of the density functional approximation is very small, and has therefore no impact on the general conclu-

sions presented here. When MD simulations were carried out, a reduced plane wave cutoff of 400 or 600 Ry and box  $Z$  length of 25 Å was used, with full periodic boundaries. The equations of motion of the Nosé-Hoover chain<sup>106</sup> were integrated with a 0.5 fs time step. Before production runs, each system was equilibrated for 1 ps at the desired temperature. CVHD biasing was applied with the PLUMED plugin,<sup>31</sup> in the form of equations (2.4) and (2.5). Bond distortions were biased up to a maximal value of 0.5 (50% bond elongation compared to equilibrium) through addition of a Gaussian of height 0.01 eV and width 0.05 every 10 fs, with a bias factor of  $\gamma = 20$ . The boost factors that were obtained range from  $\sim 100$  at 800 K, to over  $3 \times 10^6$  at 400 K.

### 6.2.2 Modeling correct charge distributions

A naive approach to model a plasma-charged catalyst surface would be a straightforward introduction of the charge, i.e., adding an additional electron to a surface slab model to generate a negative surface charge. Although this a reasonable way to treat homogeneous systems (such as a solvated ion) in which the background charge essentially approximates the effect of a uniform distribution of counterions, it breaks down for systems with an inhomogeneous countercharge distribution.<sup>200</sup> In particular, such a charge distribution will be a poor approximation of a surface exposed to a plasma, in which there is a clear charge separation between the negatively charged surface and the plasma sheath that contains positively charged ions. We believe that an explicit counterion added to the gas phase at a sufficient distance above the surface, with periodic boundaries parallel to the surface, will essentially act as a charged plate of opposite charge and generate an electric double layer. This way, an electric field perpendicular to the surface of a magnitude appropriate to the surface charge density will be naturally obtained as a byproduct of the procedure. Hence, such an approach yields a realistic model of a charged catalyst surface exposed to a plasma because (1) the charge distributions match those of the true system, i.e., a negatively charged surface exposed to a gas phase carrying positive countercharges and (2) an electric field, perpendicular to the surface follows self-consistently from these charge distributions.

The practical realization of this approach in a standard DFT code (such as CP2K) is as follows. The negative surface charges in this work require a positive countercharge which, in the simplest case, can be a proton. It is, however,



**FIGURE 6.1:** Convergence of computed CO<sub>2</sub> adsorption energies (in the d-IVb configuration) on a negatively charged slab with respect to the position of the neutralizing countercharge. For the total energies, a straight line is fitted.

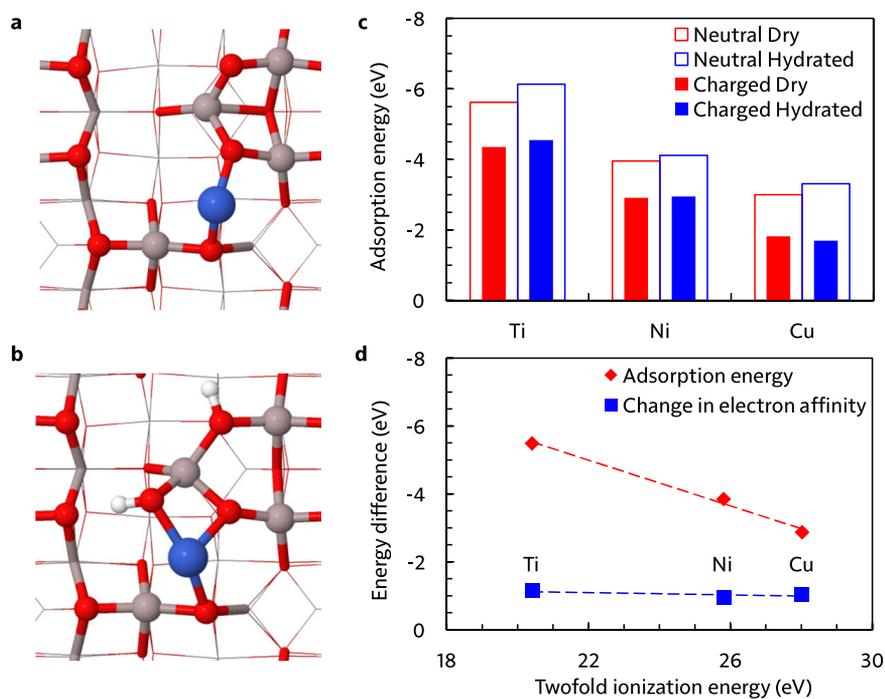
not always straightforwardly possible to introduce gas phase ions of specified charge into the simulation box. Indeed, if this approach were attempted in a plane wave DFT code, charge transfer could occur to the point charge due to use of a non-localized basis set, making it impossible to control the charge of the slab; after all, the ground state solution of such a surface+free atom system, given full variational freedom of the electron density, is perhaps not the required charge-separated state. This can be compensated by using a DFT code that expands the Kohn–Sham orbitals in an *atom-centered* (localized) basis: if no basis functions are added on the counterion, no electronic density can spill over, its charge can be precisely controlled, and the desired surface charge can be enforced. The method is in principle readily usable in any DFT code that uses localized basis sets but has, to the best of our knowledge, not yet been described in the literature.

In the setup adopted in this work, a single additional electron is considered and a proton countercharge is placed at a  $Z$  position of 40 Å in a box of dimensions  $16.1606 \times 16.8106 \times 100 \text{ Å}^3$ . The accuracy of the countercharge approach hinges on the assumption that if the energetic contribution from the point charge (besides generating a perpendicular electric field) is the same for all systems, adsorption energies are not affected because its effect is cancelled when subtracting the energies of the slab + adsorbate and the clean slab. To verify this assumption, energies were computed for different  $Z$  positions of the countercharge, depicted in Figure 6.1. The interaction energy of the counter-

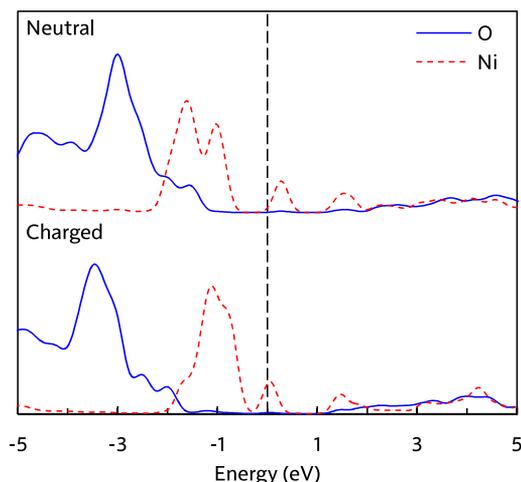
charge and the slab is linearly dependent on their mutual distance, which is the expected behavior for two interacting infinite charged plates. At sufficient separation, the effect of the adsorbed species on this interaction becomes negligible, and the computed adsorption energy converges. That is, the electrostatic interactions between the surface and the countercharge reduce to a simple plate-plate interaction (with corresponding electric field) in an averaged sense, rather than a point charge-adsorbed molecule interaction that can be observed if the point charge is too close. A  $Z$ -position higher than 30 Å (or a distance of  $\sim 20$  Å) suffices, and the value of 40 Å used in our production calculations is a very conservative choice.

The fact that the slab-countercharge interaction quickly converges at fairly short distances is also a further validation of the physical soundness of the model. Typical Debye lengths  $\lambda_D$  in nonthermal plasmas are in the order of  $\mu\text{m}$  to  $\text{mm}$ ,<sup>218</sup> so that the plasma sheath thickness is beyond the length scales of our simulation cell which, however, is shown here to give converged electrostatics. In other words, our model of the charge distribution is a rather good approximation of an electric double layer, provided that the characteristic length scales are large. It would therefore not work for charged slabs in contact with an electrolyte solution. For example, a room temperature aqueous solution with a 0.1 M ionic strength has  $\lambda_D \approx 1$  nm, which means that counterions are mostly present at distances with unconverged slab-ion interactions, and more fine-grained descriptions of the charge distribution must be adopted.

For the surface model used, a single excess electron corresponds to an electron density of  $3.68 \times 10^{17} \text{ m}^{-2}$  or a surface charge density of about  $-0.06 \text{ C m}^{-2}$ . Recent measurements on alumina exposed to a multi-filament atmospheric pressure dielectric barrier discharge (DBD) put the plasma-induced surface electron density in the order of  $10^{15}$ – $10^{17} \text{ m}^{-2}$ , close to values used here.<sup>168</sup> In view of these results, and assuming that the charge penetration depth is no more than 1 nm, the relatively small surface model employed in this work is in fact also a realistic approximation of the charged plasma-exposed alumina surface itself.



**FIGURE 6.2:** Transition metal adsorption on neutral and negatively charged alumina surfaces. (a) and (b) top view of the most favorable transition metal adsorption configuration on the dry and hydrated surfaces, respectively. Hydrogen: white, oxygen: red, aluminium: gray, and metal: blue. (c) Metal adsorption energies on the two surfaces, with and without extra charge. (d) Correlation of metal binding energies and the change of surface electron affinity  $\Delta\chi = E_{\text{ads}}(M, \text{neutral}) - E_{\text{ads}}(M, \text{charged})$  induced by metal binding with metal ionization energies.



**FIGURE 6.3:** Projected densities of states (PDOS) for Ni supported on the dry surface. Shown are the states of Ni and surface oxygens. Energies are centered on the Fermi level. It can be seen that mixing of metal and surface states is essentially nonexistent.

## 6.3 Results

### 6.3.1 Transition metal atom adsorption on the $\text{Al}_2\text{O}_3$ support

The  $\text{Al}_2\text{O}_3$  surface is known to provide strong anchoring sites for adsorbed metal atoms, and is therefore an excellent support material to stable single atom catalysts.<sup>219–221</sup> For the Ti, Ni, and Cu atoms, different adsorption sites were probed on both the dry and hydrated surface. As discussed in Section 6.5, additional coordination by adsorbed water has an impact on the adsorption characteristics and relative energetics of the surface sites. However, for all metal–surface combinations, the adsorption configuration in which the metal atom is coordinated by two  $\text{O}_2$  atoms (Figure 6.2a) was found to be the most favorable, and is the only one considered in the following (all configurations and their energies are given in Section 6.5). The effect of surface hydration (and additional OH coordination, Figure 6.2b) on the metal adsorption energy is limited (< 10%), indicating that transition metal bonding at the surface does not depend strongly on the precise hydration degree or pattern. In all configurations and on all surfaces, Ti adsorbs much more strongly on the oxide surface than Ni or Cu, as depicted in Figure 6.2c.

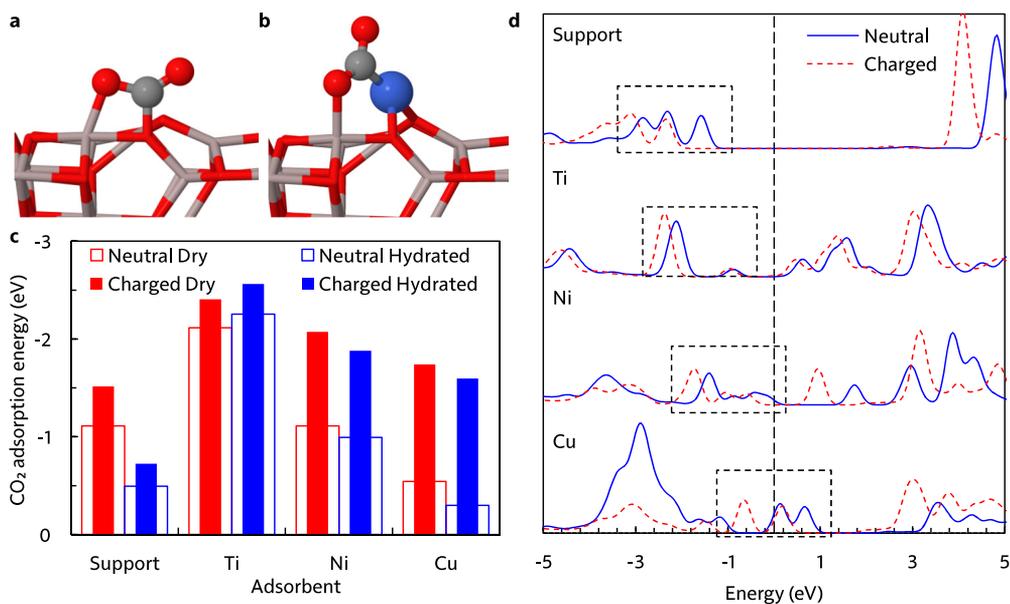
Transition metal adsorption on the negatively charged surface is not as favor-

able. The structures of all metal–support combinations were reoptimized with an additional electron, and absolute metal adsorption energies are about 1 eV smaller in all cases or, alternatively, the electron affinity of the support consistently decreases by this quantity when a transition metal atom is adsorbed. In support of the latter phrasing we find two major indications that the metal–support interaction is mostly ionic in character, with the metal atom adsorbed in its  $M^{2+}$  state. First, only very limited mixing of the metal and support electronic states is observed in the projected density of states (PDOS, see Figure 6.3 showing Ni as example), which can be associated with a primarily ionic bond. Second, the adsorption energies of the metal atoms on the dry support correlate very well with their combined first and second ionization energies, i.e., the energetic cost of  $M \rightarrow M^{2+} + 2e^-$  in the gas phase (Figure 6.2d). Combined with the near-constant  $\sim 1$  eV metal-induced downward shift of the support’s electron affinity, it can be inferred that metal atom adsorption on  $Al_2O_3$  is a redox reaction wherein the support is reduced, which therefore becomes more resistant to further reduction through the absorption of (plasma-supplied) electrons. This reduction of the support upon metal adsorption is of the same magnitude independent of the metal, which is always oxidized to  $M^{2+}$  (in this particular configuration), meaning that the support’s electron affinity is also modified in the same constant fashion.

### 6.3.2 $CO_2$ adsorption

$CO_2$  can either chemisorb on the metal atom, or on the  $Al_2O_3$  support. In all cases, the adsorbed  $CO_2$  molecule adopts a bent carbonate-like structure, with the O–C–O angle deformed by over  $40^\circ$ , as shown in Figure 6.4a-b.

On the support, the preferential adsorption site is on an  $Al_{IV}-O_2$  Lewis pair, forming Al–O and O–C bonds (Figure 6.4a). Another configuration involving an  $Al_{III}-O_2-Al_{IV}$  site is 0.46 eV less favorable due to the higher Lewis acidity of the  $Al_{III}$  site. Indeed,  $CO_2$  is a Lewis acid and consequently its affinity with a surface site is proportional with the site’s basicity, which is why it is typically used as probe molecule to determine surface basicity. In line with this reasoning, the Lewis acidity of the most favorable  $Al_{IV}$  site increases upon hydroxylation of  $Al_{III}$ ,<sup>217</sup> correlating with the lower (by 0.62 eV)  $CO_2$  adsorption energy on the hydrated surface. A negative charge transfer, respectively  $-0.33e$  and  $-0.31e$  on the dry and the hydrated surface, further confirms the Lewis acidic behavior of



**FIGURE 6.4:** Effect of surface charging on CO<sub>2</sub> adsorption. (a) and (b) Most favorable adsorption configuration on the support and supported transition metal atom. (c) Adsorption energies on all sites, with and without extra charge. (d) PDOS of C in CO<sub>2</sub> adsorbed on all relevant sites on the dry support, centered on the Fermi level (or, rather, the energy of the highest occupied orbital). The relevant high-lying bonding orbitals are marked with dashed boxes.

the CO<sub>2</sub> molecule. CO<sub>2</sub> chemisorption on the  $\gamma$ -Al<sub>2</sub>O<sub>3</sub> (110) surface is generally quite similar to adsorption on many other oxides, with adsorption energies in the range of  $-0.5$  to  $-2.5$  eV, formation of a surface carbonate with Lewis basic surface oxygens, strongly bent bi- or tridentate adsorption configurations, and negative charge transfer to the molecule.<sup>137-144</sup> The fairly strong adsorption of CO<sub>2</sub> on the alumina support might also increase the retention time of the molecule near the surface—giving it more time to reach an active catalyst site—although it could also increase the competition between metal and support sites.

For all metal–surface combinations, the IVa adsorption configuration is the most stable, and is therefore used in the CO<sub>2</sub> adsorption calculations. In all cases, CO<sub>2</sub> is found to adsorb in a bridged structure on both the metal atom and the neighboring Al<sub>IV</sub> surface atom, highlighting the important effect of the support material on the chemical properties of the adsorbed transition metal (Figure 6.4b). Similar binding modes were observed for larger supported metal clusters, for which the metal–support interface was also the preferred CO<sub>2</sub> adsorption location.<sup>155,156</sup> Ni and Cu exclusively bind the CO<sub>2</sub> carbon atom, whereas the surface Al atom binds one of its oxygen atoms. Ti, on the other hand, forms an  $\eta^2$  complex with the molecule, coordinating both atoms of a C–O bond, while the Al surface atom coordinates the other C–O bond. The ability of the metal–support interface to provide Lewis acid–base pairs is an important property of oxide-supported metal catalysts that can significantly impact its reactivity, with the support material playing in active role beyond merely acting as support for the metal catalyst.

The supported metal atoms show a very diverse CO<sub>2</sub> binding behavior, with Ti having the strongest interaction of  $-2.12$  eV ( $-2.25$  eV on the hydrated surface), Ni half as strong with  $-1.11$  eV ( $-0.99$  eV), and Cu even weaker with only  $-0.54$  eV ( $-0.30$  eV), following trends that were established earlier for fcc (100) metal surfaces.<sup>140</sup> In fact, the van der Waals component contributes to about half of the Cu/CO<sub>2</sub> interaction (amounting to  $0.22$  eV and  $0.18$  eV on the dry and hydrated surface, respectively), pointing to only very limited chemical bonding, insomuch that adsorption on the alumina support is favored over adsorption on the Cu atom. On the dry surface, this is also true for Ni, although hydration greatly diminishes the support's CO<sub>2</sub> adsorption ability and favors adsorption on supported Ti or Ni (at least for the particular hydration pattern

employed here).

Introduction of an additional electron has a dramatic impact on the adsorption properties, significantly improving the binding characteristics of all CO<sub>2</sub> adsorption modes, as summarized in Figure 6.4c. The magnitude of the effect is the most striking in the case of Cu, which (on the hydrated surface) sees a four-fold increase of the binding energy upon charging, even becoming competitive to Ni. In general, surface charging appears to somewhat “level out” the differences between the metal catalysts, because the effect is much weaker for Ti, which already shows very strong binding with neutral charge.

From a Lewis acid–base theory perspective, negatively charging the surface will naturally increase its basicity and hence improve the binding with the acidic CO<sub>2</sub> molecule. To explain the differences between the adsorption modes, their electronic structure must however be analyzed. In particular, examination of the bonding states in the PDOS, depicted in Figure 6.4d, and their position relative to the Fermi level is here useful. The comparatively minor surface charging effect on adsorption at the dry support can be attributed by the fact that the highest bonding state, formed by overlap of CO<sub>2</sub> antibonding  $\pi^*$  orbitals with surface *p* or *d* states, is fairly low-lying, centered around  $-2.55$  eV (relative to the Fermi level) and shifting to  $-3.31$  eV upon charging; similar observations can be made for CO<sub>2</sub> adsorption on supported Ti ( $-2.13$  eV dropping to  $-2.38$  eV). In contrast, the bonding M–CO<sub>2</sub> states of the neutral dry Ni and Cu-based catalysts lie partially above the Fermi level, especially explaining the very limited Cu–CO<sub>2</sub> bonding: the lower the energy of the metal *d* states, the more difficult they overlap with the high-lying CO<sub>2</sub> antibonding  $\pi^*$  orbitals, resulting in a higher energy (i.e., less stabilization) of the bonding states. Surface charging can therefore have a much larger impact in these cases, lowering the bonding states from  $-1.00$  to  $-1.43$  eV (Ni), and from  $0.46$  to  $-0.28$  eV (Cu), relative to the energy of the highest occupied orbital. The relative lowering of the bonding states upon charging is also reflected by the charge of the adsorbed CO<sub>2</sub> molecule: increased occupation of these orbitals, which are partially localized on the molecule, leads to a larger electron density; for example, the charge of CO<sub>2</sub> adsorbed on Cu on the hydrated surface changes by  $-0.27e$  upon surface charging, compared to only  $-0.08e$  and  $-0.12e$  on Ti and Ni, respectively.

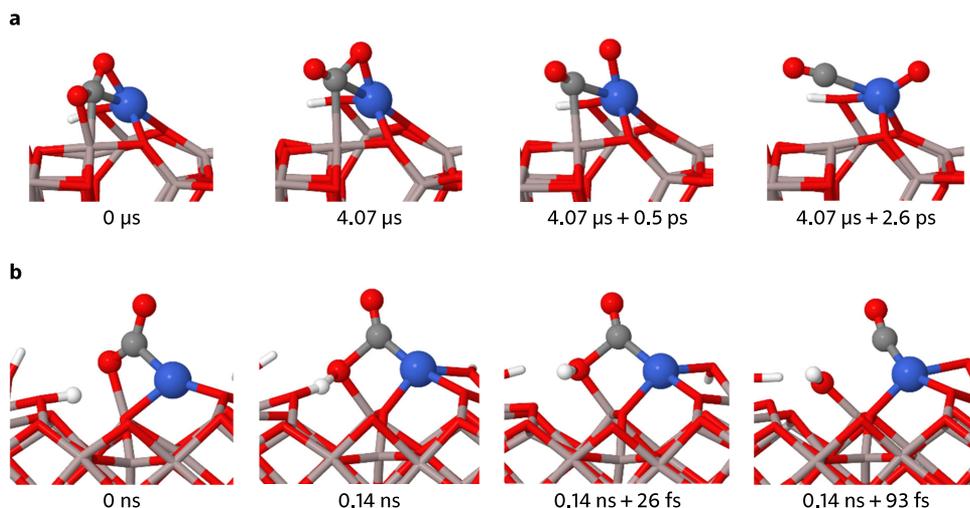
**TABLE 6.1:** Influence of surface charging on molecular adsorption energies (eV) at various sites on the  $\text{Al}_2\text{O}_3$  support (d: dry surface, h: hydrated surface).

molecule	site	neutral	charged
$\text{H}_2\text{O}$	d- $\text{Al}_{\text{III}}$	-2.42	-2.56
$\text{CH}_4$	d- $\text{Al}_{\text{IV}}$	-0.38	-0.46
	h- $\text{Al}_{\text{IV}}$	0.14	0.11
CO	d- $\text{Al}_{\text{III}}$	-1.38	-1.20
	d- $\text{Al}_{\text{IV}}$	-1.07	-1.02
	h- $\text{Al}_{\text{IV}}$	-1.29	-1.14

### 6.3.3 Adsorption of other molecules on the support

While the Lewis acidic  $\text{CO}_2$  shows improved adsorption behavior on a negatively charged substrate, this is not necessarily a good indicator for molecular adsorption in a general sense. Therefore, we calculated the adsorption energies of water, methane, and carbon monoxide on both the neutral and the charged surface, summarized in Table 6.1. For water, the hydration energy is considered, i.e., the reaction energy of forming the hydrated surface model from the dry surface. Similarly, for methane, dissociative adsorption into  $\text{CH}_3$  and H is the studied process. CO is commonly used as basic probe molecule to assess the Lewis acidity of a surface, and is also a major reaction product in  $\text{CO}_2$  reduction.

Generally, surface charging improves the adsorption behavior of  $\sigma$ -bonded species ( $\text{H}_2\text{O}$  and  $\text{CH}_4$ ), but to a much smaller degree (no more than 0.15 eV) due to absence of unoccupied states close to the Fermi level. CO shows the opposite behavior, consistent with its Lewis basic character, also again showing the relative hydration-induced decrease in basicity for this particular configuration. While CO binds on the surface by donating its lone electron pair on C, a negative surface charge can be donated back to CO by partially filling its antibonding  $\pi^*$  orbitals, which become more easily accessible because of the higher energy of the surface electronic states. For example, a CO molecule bound at an  $\text{Al}_{\text{III}}$  surface atom has a total charge of  $0.06e$  on the neutral, and  $0.02e$  on the charged surface. Hence, in all cases unoccupied states close to the Fermi level play a crucial role in determining the charge dependence on adsorption; the precise direction of the effect depends on their (anti)bonding nature. The implications of these contrasting respective bonding/antibonding interactions in adsorbed  $\text{CO}_2$  and CO will be further explored in the following section.



**FIGURE 6.5:** CVHD simulation of  $\text{CO}_2$  splitting by supported metal atoms. Reaction steps observed for (a) Ti at 400 K and (b) Ni at 800 K. Accelerated time is given below each frame. Given the time scales, it can be concluded that for Ni, proton transfer and C–O splitting are essentially concerted.

### 6.3.4 Impact on surface reactions

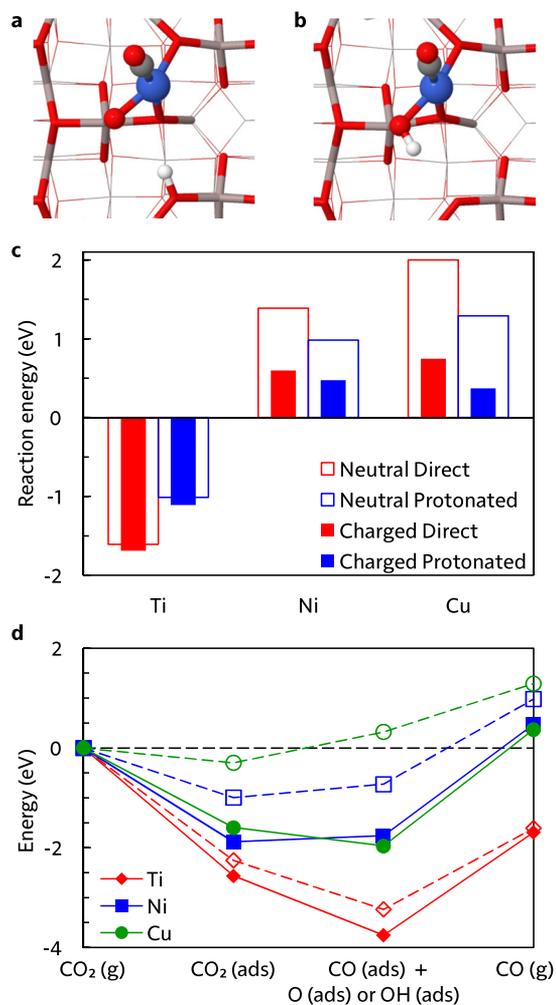
The uncatalyzed gas-phase splitting of  $\text{CO}_2$  (i.e.,  $\text{CO}_2 \rightarrow \text{CO} + \frac{1}{2}\text{O}_2$ ) is thermodynamically highly unfavorable ( $\Delta H = 2.9$  eV). On a suitable catalyst, the reaction  $\text{CO}_2(\text{g}) \rightarrow \text{CO}(\text{g}) + \text{O}(\text{ads})$  can be made to take place more easily, having a beneficial impact on the *overall* rate of any process that depends on  $\text{CO}_2$  splitting, including dry reforming. Although a structurally simple atomically dispersed catalyst can ostensibly only take part in a small number of reaction mechanisms (direct C–O splitting in this case), the chemical activity of the support material significantly increases the number of possible  $\text{CO}_2$  activation pathways. While it is not in the scope of this work to obtain a comprehensive picture of the complete catalytic processes of the considered systems, it is useful to have an initial picture of the most common reactions of adsorbed  $\text{CO}_2$ .

In CVHD simulations of  $\text{CO}_2$  adsorbed on the hydrated Ti-based catalyst, the *direct splitting* reaction,  $\text{CO}_2(\text{ads}) \rightarrow \text{CO}(\text{ads}) + \text{O}(\text{ads})$  (Figure 6.5 a), could be observed at a temperature as low as 400 K (which is indeed typically achieved in a DBD plasma) after a simulated time of 4.1  $\mu\text{s}$ . Ni is not found to be active at 400 K within the CVHD time scale (which does not, however, rule out the general possibility of a reaction), but does react at 800 K after 0.14 ns.

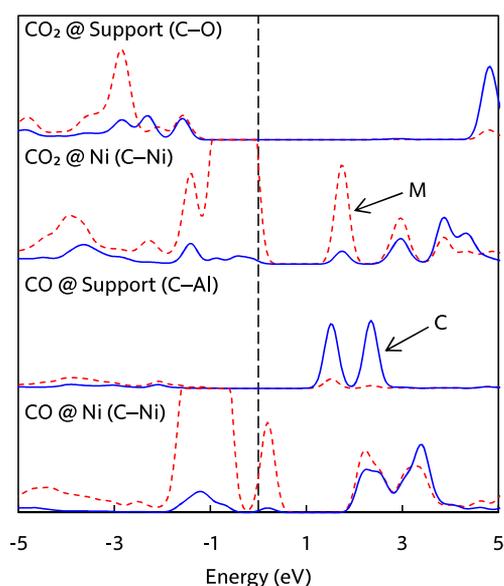
However, no direct splitting is observed in this case, but rather a *proton-mediated* mechanism in which a proton is first transferred to the CO<sub>2</sub> molecule from an OH group at the support, leading to instantaneous dissociation into CO and OH (Figure 6.5 b) in a near-concerted fashion.

Motivated by this apparent difference in the reactivity of Ti and Ni, the overall reaction energies of the two competing CO<sub>2</sub> activation pathways leading to CO (g) + O (ads) and CO (g) + OH (ads), respectively (in their most stable configuration as depicted in Figure 6.6a and b), were calculated for all metals on both the neutral and charged surface. It is indeed found that Ti is more active towards direct splitting, while Ni and Cu favor a proton-mediated mechanism (Figure 6.6c). Also in agreement with the simulations is the much more favorable reaction energy of the initial splitting step on Ti, which reacted at 400 K and exhibits a reaction energy of  $-0.98$  eV, as compared to Ni, which reacted at higher simulated temperatures and has a reaction energy of  $0.27$  eV. The much higher reactivity of Ti can be attributed to its higher intrinsic reductive abilities: while formally adsorbed in the Ti<sup>2+</sup> state, a further oxidation to Ti<sup>4+</sup> is possible, which can be achieved by reducing CO<sub>2</sub>.

When examining the effect of an excess electron on the overall splitting process the results largely echo those of CO<sub>2</sub> adsorption, with reactions on Ti relatively unaffected ( $\Delta\Delta E_{\text{ads}} = -0.08$  eV for direct splitting) and the process on Cu exhibiting a very strong influence ( $\Delta\Delta E_{\text{ads}} = -0.92$  eV for proton-induced splitting) by the additional negative surface charge (Figure 6.6c). Interestingly, when decomposing the energetic contributions of the separate process steps (depicted in Figure 6.6d), it can be seen that the initial CO<sub>2</sub> adsorption step is in fact the most affected by the charge, while the subsequent steps are not as dissimilar to their counterparts on the neutral surface. Larger effects are observed again for the desorption of CO, which is more strongly bound on the charged than the neutral surface, in contrast to what was found for the adsorption on the support. Indeed, CO<sub>2</sub> always adsorbs through hybridization of its antibonding  $\pi^*$  orbitals with high-lying surface states, giving rise to new bonding states around the Fermi level that can be further stabilized with an excess negative charge. CO, however, interacts with the support through a  $\sigma$ -bonded interaction involving its lone pair, leaving its mostly unmodified  $\pi^*$  orbitals available in an antibonding state that can be occupied by an excess electron, destabilizing the adsorption complex. In contrast, CO interaction with a metal involves a *d*-



**FIGURE 6.6:** Effect of an excess electron on the reaction energies of the CO<sub>2</sub> splitting process. (a) Product of the direct splitting reaction. (b) Product of proton-mediated splitting. (c) Overall reaction energies for the two studied mechanisms. (d) Most favorable pathways on all metals. Empty symbols and dashed lines: neutral surface, filled symbols and full lines: charged surface.



**FIGURE 6.7:** Contrasting adsorption behavior of  $\text{CO}_2$  and  $\text{CO}$ . PDOS of the C–M interaction, with M being the surface atom forming a bond with a C atom in the molecule. It can be seen that around the Fermi level,  $\text{CO}_2$   $\pi^*$  states are always hybridized with the surface states, giving rise to new bonding states. This is also true for  $\text{CO}$  bound on a metal atom (here Ni), but not on the support, where the  $\pi^*$  orbitals remain clearly recognizable as such, do not mix appreciably with the surface states, and remain antibonding. Blue lines: C, red dashed lines: M.

$\pi^*$  overlap that again produces a bonding interaction that will be strengthened by charging. These concepts are illustrated in Figure 6.7.

While it becomes more difficult to release CO from the metal catalyst upon the charging, the overall CO<sub>2</sub> splitting process is more favorable. Moreover, CO need not be the final product, but could react further to yield base chemicals, such as formaldehyde or methanol upon addition of a hydrogen source, just as well as the additional oxygen atom on the surface can take part in various oxidation processes. This kind of more detailed pathway studies should be investigated further.

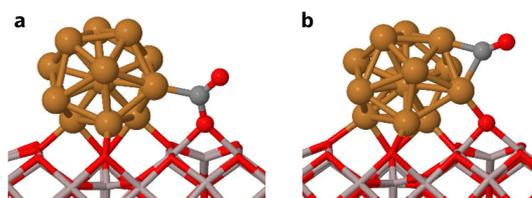
While we have primarily discussed the thermodynamic effects of the excess electron, the kinetics of the catalytic reaction are also of great importance. As a first assessment of the impact of the surface charge on reaction barriers, estimated transition states of the direct splitting reaction were determined. We find that the presence of a negative surface charge consistently lowers the energy of configurations with partially broken bonds such as transition states, lowering the estimated splitting barrier on all metals: from 1.15 eV to 0.75 eV on Ti, from 0.80 eV to 0.65 eV on Ni, and from 1.26 eV to 0.83 eV on Cu. Through the presence of an additional electron, partially unsaturated atoms in the transition state receive some additional stabilization, hence lowering the apparent reaction barrier and increasing the reaction rate. It must be mentioned that the calculated barriers do not necessarily reflect the lowest energy splitting pathway, but are chosen so as to provide a consistent set of benchmark configurations. For example, in our CVHD simulations we find that CO<sub>2</sub> splitting on Ti occurs from a rearranged state in which the molecule is bound exclusively on the metal, as opposed to the metal-support bridge we used as initial state here, as shown in Figure 6.5.

### 6.3.5 Larger clusters

Although we have primarily focused on supported single atoms, it is instructive to assess to what extent our extensive conclusions for these systems might be valid for larger supported clusters. In a general sense, introducing larger clusters defeats the purpose of using single atoms: the number of degrees of freedom increases with the number of metal atoms in the cluster, meaning that results could become more and more influenced by the choice of the configurations used in the calculations, in contrast to the rather limited set of structures

**TABLE 6.2:** Charge effect on CO<sub>2</sub> binding by a supported Cu<sub>13</sub> cluster. Energies (eV) of a Cu<sub>13</sub> cluster adsorption on the dry Al<sub>2</sub>O<sub>3</sub> surface, and of CO<sub>2</sub> adsorption and activation on this cluster.

	neutral	charged
Cu <sub>13</sub> adsorption on Al <sub>2</sub> O <sub>3</sub>	-6.13	-5.63
CO <sub>2</sub> adsorption on Cu <sub>13</sub>	-0.78	-1.16
CO <sub>2</sub> split on Cu <sub>13</sub>	-0.12	-0.15
CO desorption from Cu <sub>13</sub>	1.70	1.87
Overall CO <sub>2</sub> splitting	0.79	0.56

**FIGURE 6.8:** CO<sub>2</sub> splitting on a supported Cu<sub>13</sub> cluster. (a) CO<sub>2</sub> adsorption configuration, (b) configuration after splitting. The cluster is adsorbed on the dry Al<sub>2</sub>O<sub>3</sub> surface.

that must be considered in the case of single atom catalysts. Therefore, these initial calculations cannot offer the same fine-grained level of conclusions that has been reached for supported single atoms but, rather, reveal if our previous conclusions are not an artifact of the particular (pragmatic) choice of catalyst model.

We investigated the charging effect on the properties of a supported icosahedral Cu<sub>13</sub> cluster, bound on the surface by three O<sub>2</sub> surface atoms and one O<sub>3</sub> site. Upon absorption of an electron most of the additional charge is localized in the cluster, which is changed by  $-0.79e$ , and a destabilization of 0.5 eV is observed (Table 6.2). As the studied CO<sub>2</sub> binding mode we considered a bridged structure at the metal–support interface, in analogy with the structures obtained on the single atom catalysts (Figure 6.8). Again, as evidenced by Table 6.2, the trends observed for the single atom catalysts are retained, although in somewhat diminished form. Yet, the overall splitting reaction CO<sub>2</sub> (g) → CO (g) + O (ads) is 0.23 eV more favorable on the negatively charged surface, which is significant.

## 6.4 Conclusions

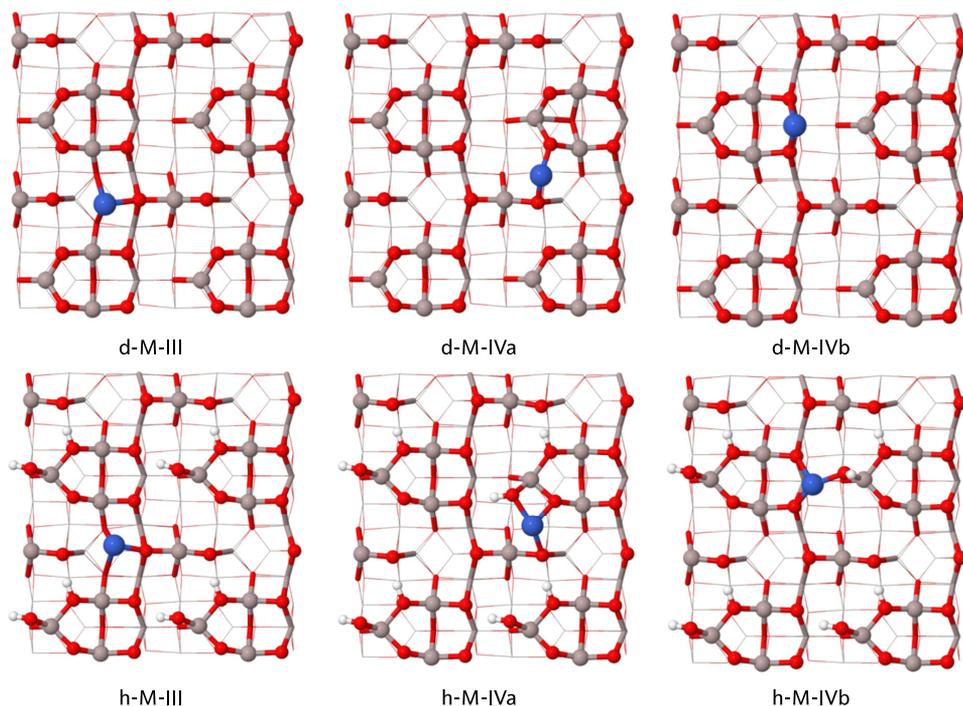
In the most general sense, electron deposition leads to a chemical reduction of the catalytic surface and, hence, increases its reductive capabilities. Specifically, this phenomenon has a very favorable effect on CO<sub>2</sub> activation, with respect to both adsorption strength and overall reaction energy of the splitting reaction. For the strongly oxidizable adsorbed Ti catalyst, this effect is not as pronounced as for Ni and Cu: while all metals formally adsorb in their M<sup>2+</sup> state, Ti can easily be further oxidized to Ti<sup>4+</sup>, allowing it to act as a strong reducing agent without having to be charged, as evidenced by its strong CO<sub>2</sub> activation abilities. The properties of the latter are also largely in line with the redox properties of TiO<sub>2</sub> surfaces, resulting from oxygen vacancy creation and annihilation and which allow for efficient reduction of CO<sub>2</sub>.

A less general interpretation of the phenomenon involves viewing the negatively charged catalyst as more Lewis basic, which is appropriate for the description of the bare Al<sub>2</sub>O<sub>3</sub> support, but is more difficult to apply once adsorbed transition metal clusters have to be considered, as evidenced by the different behavior of CO adsorbed on the support or the metal, respectively. An analysis of the electronic structure of the adsorption complex hence provides the most valuable and robust insight into its response to surface charging.

The major impact of surface charging on the catalytic performance of supported Ni and Cu—even inducing a reversal of their relative activity—demonstrates that conclusions drawn for “conventional” thermal catalysis not necessarily hold for processes involving charged catalysts in, e.g., a plasma. Indeed, the presence of a large surface charge might help explain often-observed but poorly understood synergistic effects in plasma catalysis.

It remains to be seen to what extent the large excess electron-induced effects observed for the systems and reactions of this study are applicable to other catalysts and processes. Different support materials (e.g., semiconductors rather than isolators), larger supported clusters, transition metal surfaces, and a more exhaustive set of redox processes should all be considered in order to assess the influence of a negative surface charge on catalysts in a more general sense. The methodology outlined in this work can provide the template for such a systematic undertaking. However, the results presented in this work already point to a phenomenon with potentially far-reaching consequences: by varying the dis-

charge parameters of the plasma and the degree of electron deposition on the plasma-facing catalyst, its Lewis acidity and redox properties can be modified as well. Thus, controlling the electron deposition on a catalyst opens another avenue towards activity and selectivity control of a plasma-catalytic process.



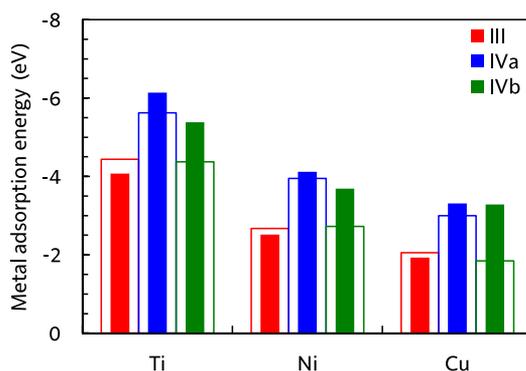
**FIGURE 6.9:** Metal adsorption configurations on the dry and hydrated surface. Configurations starting with “d-” denote the dry surface, while those with “h-” describe the hydrated surface. Hydrogen: white, oxygen: red, aluminium: gray, and metal: blue. The depicted configurations are for adsorbed Ni, but do not differ significantly for Ti and Cu.

## 6.5 Additional information

### 6.5.1 Transition metal adsorption

For the Ti, Ni, and Cu atoms, different adsorption sites were probed on both the dry and hydrated surface. These configurations are based on those of the Al atoms in the “next” surface layer; that is, the location of the Al atoms in a hypothetical additional atomic layer atop the actual surface layer in this work. The sites are named after the Al atom on which they were based meaning that, e.g., the d-III configuration is based on the position of an Al<sub>III</sub> atom in the hypothetical top layer. The possible metal atom adsorption configurations are depicted in Figure 6.9.

The most stable adsorption site on the dry surface for all three metals is d-IVa, at which two highly unsaturated (dicoordinate) O atoms can coordinate the



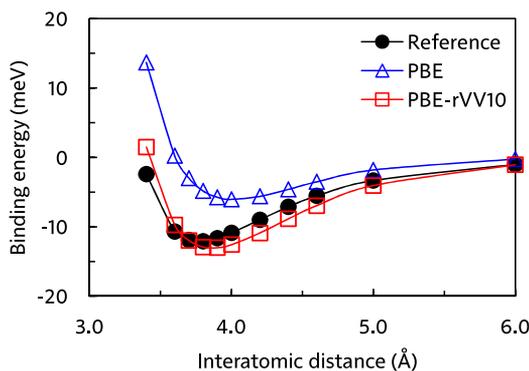
**FIGURE 6.10:** Relative energies of the metal adsorption configurations on the dry and hydrated surface. Empty bars reflect adsorption on the dry surface, filled bars represent the hydrated surface.

metal atom, while d-III and d-IVb have a similar stability (Figure 6.10, empty bars). Hydration has a drastic impact as it introduces OH groups on the surface that can provide additional coordination of the metal atom; for Cu, the h-IVa and h-IVb configurations are essentially degenerate (Figure 6.10, filled bars). Remarkably, however, the IVa configuration remains the most favorable in all considered cases. It is conceivable that at higher water coverages or with different hydration patterns, a wider array of adsorption configurations are possible; our calculations provide a first step towards a more complete understanding of transition metal adsorption on realistic alumina surfaces.

The deformation of the surface upon transition metal adsorption is quite limited in case of Ni and Cu, mainly amounting to a small ( $< 0.1 \text{ \AA}$ ) elongation of the Al–O bonds of the coordinating surface O atoms. Ti, however, can have a much larger impact on the support surface structure: in the d-Ti-IVa configuration, the  $\text{Al}_{\text{IVb}}\text{--O}_{2\text{b}}$  bond is extended from  $1.69 \text{ \AA}$  to  $2.41 \text{ \AA}$ , essentially converting  $\text{Al}_{\text{IVb}}$  into an  $\text{Al}_{\text{III}}$  site. Therefore, Ti can strongly affect the activity of the surface region close to it.

### 6.5.2 A reference van der Waals-corrected functional

The rvv10 nonlocal correlation functional,<sup>222</sup> which is a revision of the vv10 functional<sup>223</sup> that is better suited for plane wave calculations, consists of a standard rPW86PBE<sup>224</sup> generalized gradient approximation (GGA) exchange-correlation functional combined with a nonlocal correlation term due to Vydrov



**FIGURE 6.11:** Ar dimer binding curve. Energies from calculations employing uncorrected PBE and optimized PBE-rVV10 functionals are compared against an accurate reference.

and Van Voorhis. Combining a different base GGA functional with the nonlocal correlation term as van der Waals correction can be achieved by refitting the adjustable parameters in the nonlocal term.<sup>225</sup> Of these parameters,  $b$  controls the short range behavior, and  $C$  the long range. Hence, only  $b$  should be refitted when the base functional is changed, and the original value for  $C = 0.0093$  can be retained.

It is found that even for the original vV10 implementation, optimal  $b$  values depend on the fitting target. When fitting against a standard test set,  $b = 5.9$  is obtained, whereas a correct description of noncovalent interactions in liquid water<sup>226</sup> or layered solids requires  $b > 9$ .<sup>227</sup> We therefore decided to only fit PBE-rVV10 against a simple “fundamental” reference system, namely the argon dimer binding curve. This way, accuracy for specific systems is sacrificed in favor of a more clear and simple parameterization strategy with little empiricism.<sup>228</sup>

We use an accurate estimate of the attractive part of the Ar dimer binding curve as reference.<sup>229</sup> Calculations were carried out in a  $25 \times 25 \times 25 \text{ \AA}^3$  box using a 1200 Ry cutoff for the density. To minimize the basis set superposition error, the very large aug-QZV3P basis set<sup>125</sup> was used to expand the Kohn–Sham valence orbitals. All calculations were evaluated self-consistently.

The value of  $b = 9.5$  was found to be optimal (minimizing the RMS deviation from the reference), and the resulting PBE-rVV10 functional yields a good binding curve for the Ar dimer, depicted in Figure 6.11. Due to the shallowness

**TABLE 6.3:** Computational consistency checks of the effect of the exchange-correlation functional. Hydration (per H<sub>2</sub>O molecule) and CO<sub>2</sub> adsorption energies (in eV) for different adsorption configurations, as calculated by different density functional methods.

	PBE	PBE-D3	PBE-rVV10	revPBE-D3	TPSS-D3
H <sub>2</sub> O	-2.23	-2.42	-2.36	-2.36	-2.42
d-CO <sub>2</sub> -III	-0.37	-0.65	-0.64	-0.52	-0.66
d-CO <sub>2</sub> -IVb	-0.82	-1.11	-1.07	-1.02	-1.17
h-CO <sub>2</sub> -IVb	-0.20	-0.49	-0.44	-0.43	-0.56

of the curve, recovering the exact minimum is difficult with PBE-rVV10, but in overall terms, energetics are well-described.

### 6.5.3 Computational cross-checks

The stability of the various CO<sub>2</sub> adsorption sites was reinvestigated using a set of different approximations of the exchange-correlation energy. This small series of calculations should not be seen as a true “benchmark” of these approximations (due to the unavailability of a reference to measure against) but rather a “consistency check” to verify to what extent our conclusions depend on the computational choices that were made. Specifically, the treatment of dispersion interactions was verified by comparing the “plain” uncorrected PBE functional with its D3- and rVV10-corrected variants, whereas the role of the underlying functional was assessed by comparing three common approximations to the exchange-correlation energy: the widely used PBE GGA functional<sup>208</sup> (and used throughout the rest of this work), its revPBE variant<sup>230</sup> (which tends to perform better for thermochemistry<sup>231</sup> and surface science<sup>232</sup>), and the TPSS meta-GGA<sup>233</sup> (which represents a higher, more advanced, rung on “Jacob’s ladder” of density functional approximations<sup>234</sup>). These functionals were applied in their D3-corrected form.<sup>209,210</sup>

Inspection of Table 6.3 demonstrates that all methods are in close agreement with respect to the relative stability of the CO<sub>2</sub> adsorption sites and, for the dispersion corrected methods, even in terms of absolute adsorption energies. Furthermore, all methods give hydration energies in good agreement with the value of -2.34 eV/H<sub>2</sub>O calculated by Wischert *et al.*<sup>217</sup> It should be noted, however, that our plain PBE calculations do not recover the results of Pan *et al.*<sup>145</sup> of

**TABLE 6.4:** Computational consistency checks for transition metal adsorption. Comparison of D3 and rvv10 dispersion corrections for transition metal adsorption on the alumina surface. All calculations use PBE-D3 geometries. Adsorption energies are in eV.

	PBE-D3	PBE-rVV10
<i>M @ dry surface</i>		
d-Ti-III	-4.44	-4.24
d-Ti-IVa	-5.62	-5.49
d-Ti-IVb	-4.37	-4.26
d-Ni-III	-2.67	-2.53
d-Ni-IVa	-3.95	-3.85
d-Ni-IVb	-2.73	-2.61
d-Cu-III	-2.05	-1.90
d-Cu-IVa	-3.00	-2.87
d-Cu-IVb	-1.84	-1.71
<i>M @ hydrated surface</i>		
h-Ti-III	-4.07	-3.88
h-Ti-IVa	-6.14	-5.97
h-Ti-IVb	-5.38	-5.22
h-Ni-III	-2.52	-2.38
h-Ni-IVa	-4.12	-3.99
h-Ni-IVb	-3.69	-3.57
h-Cu-III	-1.93	-1.78
h-Cu-IVa	-3.31	-3.17
h-Cu-IVb	-3.28	-3.16

CO<sub>2</sub> adsorption on the  $\gamma$ -Al<sub>2</sub>O<sub>3</sub> (110) surface, as these authors reported a different relative ordering of d-III and d-IVb configurations, with adsorption energies of -0.43 and -0.27 eV, respectively, and a bidentate structure for d-III rather than a tridentate. The origin of this discrepancy is unclear to us, but our results are more in line with those of other oxides and the relative Lewis acidity of the sites.<sup>217</sup>

It can be seen that the effect of including dispersion corrections is quite large, amounting to about 0.3 eV for CO<sub>2</sub> adsorption. The necessity of including dispersion was previously demonstrated for CO<sub>2</sub> adsorption on TiO<sub>2</sub><sup>235</sup> and is confirmed here for Al<sub>2</sub>O<sub>3</sub>. It is reassuring to observe that the D3 and rvv10 meth-

**TABLE 6.5:** Computational consistency checks for CO<sub>2</sub> adsorption. Comparison of D3 and rvv10 dispersion corrections for CO<sub>2</sub> adsorption on the alumina surface. All calculations use PBE-D3 geometries. Adsorption energies are in eV.

	PBE-D3	PBE-rVV10
<i>CO<sub>2</sub> @ support</i>		
d-CO <sub>2</sub> -IVb	-1.11	-1.06
h-CO <sub>2</sub> -IVb	-0.49	-0.44
<i>CO<sub>2</sub> @ metal atom</i>		
d-Ti-CO <sub>2</sub>	-2.12	-2.09
d-Ni-CO <sub>2</sub>	-1.11	-1.09
d-Cu-CO <sub>2</sub>	-0.54	-0.52
h-Ti-CO <sub>2</sub>	-2.25	-2.23
h-Ni-CO <sub>2</sub>	-0.99	-1.01
h-Cu-CO <sub>2</sub>	-0.30	-0.27

ods, although based on different principles (atom pairwise and nonlocal density based, respectively), give very similar results, thus validating each other's applicability to this system. In all other calculations (of the neutral systems), we adopted the PBE-D3 method for all geometry optimizations but, because we deal with the rather challenging case of metal-containing systems, follow Hujó & Grimme's recommendation<sup>225</sup> and verify PBE-D3 results against single point cross-checks with PBE-rVV10. These tests are collected in Table 6.4 and Table 6.5, and show that while rvv10 consistently gives smaller adsorption energies for transition metal atoms (in the order of 0.1–0.2 eV), it is very close to D3 for adsorption of CO<sub>2</sub> on those metals and, most importantly, gives the same energy differences between adsorption sites.

#### 6.5.4 Should cell neutrality always be enforced?

In this chapter, we have enforced net neutrality of the simulation cell through the explicit treatment of the countercharge distribution. The prime motivation for this choice of model was its appropriateness for the problem at hand, where plasma sheath electrostatics could be easily approximated within an atomistic model. However, inclusion of countercharges also solves the somewhat more philosophical problem of the implicit neutralizing background charge that is present in charged simulation cells: to what extent is a calculation with this kind

of unphysical charge distribution really a good model of anything? In Chapter 5, we brushed over this question by pointing out that it is the best we can do in terms of building a model that *purely* captures charging effects, not affected by any external fields arising from interactions with a countercharge distribution. In that sense, it is in fact an excellent model for gaining fundamental insights because of its purity.

Yet, a very similar setup could also be realized in a neutral cell, and is worth investigating. An explicit countercharge has been used to simulate the plasma sheath by placing all charge on one side of the slab, clearly distinguishing one side as the surface, and one as the bulk end. Evidently, an additional electric field is created through the addition of the charge. However, if the countercharge were evenly distributed above and below the slab, the net field generated should be zero.

We test this assumption by revisiting the adsorption of  $\text{CO}_2$  on  $2e^-$ -charged  $h$ -BN of Chapter 5. Instead of imposing a net charge on the simulation cell, two counterions were added in the form of basis set-free protons. They were placed at  $Z$  positions of  $\pm 20 \text{ \AA}$  relative to the sheet, enclosed by a periodic box with a nonperiodic  $Z$  length of  $100 \text{ \AA}$ . The  $\text{CO}_2$  adsorption energy in the case of this *explicitly* compensated charge is  $-3.044 \text{ eV}$ , which is almost identical to calculations with *implicitly* compensated surface charges, for which we obtained  $-3.042 \text{ eV}$ . It can therefore be concluded that if one is interested in adsorption energetics at charged surfaces, the overall simulation cell charge is irrelevant. Rather, one should ask whether only the isolated charge effects are of interest, or if the model requires some additional electrostatic interactions. For instance, a plasma-style charge distribution—with all countercharge placed at one side,  $40 \text{ \AA}$  above the surface—yields a markedly different adsorption energy of  $-5.108 \text{ eV}$ . These conclusions mirror the results of Hub *et al.*, who found that not the magnitude of the cell charge is of relevance, but rather its distribution across the cell.<sup>200</sup>



# A brave new world for DFT?

## 7.1 The density functional zoo

In the preceding chapters, we have constructed computational models of the chemistry at charged surfaces. We have explicitly dealt with appropriate choices of the specific surface model, the charge distribution, and the properties of the simulation cell. What has not been addressed is the choice of the density functional approximation (DFA). A standard GGA functional (PBE) and van der Waals corrections (of the D3 type) were adopted as the workhorse method for electronic structure and geometry optimizations, without specific validation tests. In Chapter 6, we have tested the effect of the chosen semilocal (meta)GGA and description of van der Waals interaction and found no major differences between the methods, but only for a subset of cases involving *neutral* surfaces. Given the only fairly recently growing interest in charge-modulated catalytic processes, it might be useful to assess to what extent computational predictions depend on the applied level of theory.

Density functional theory is a very active research field, and has been exhaustively discussed in several excellent recent reviews.<sup>236–241</sup> DFT exchange-correlation functionals can be grouped by their mathematical form, where each incremental increase in complexity is a higher rung on Jacob’s ladder<sup>234</sup> towards the “heaven of chemical accuracy”—the higher the rung they are on, the better DFAs tend to perform.<sup>242</sup> The first-rung *local density approximation* (LDA) only depends on the electron density  $\rho$ , or the spin-labeled densities  $\rho_{\uparrow}$  and  $\rho_{\downarrow}$  (local spin density approximation, LSDA). It is in principle only suited for uni-

---

The results presented in this chapter will be published as:

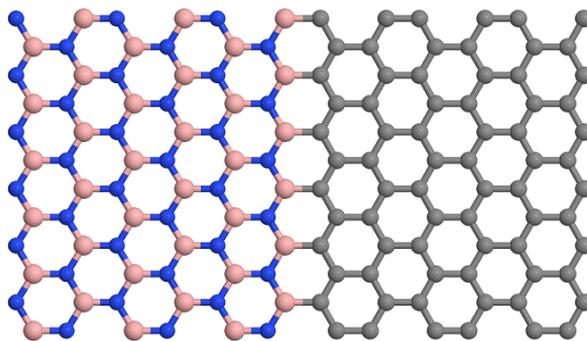
K. M. Bal and E. C. Neyts. On the applicability of density functional theory to charged electrocatalysts. (*in preparation*)

form or very slowly varying densities and, hence, performs poorly for atoms and molecules.

The *generalized gradient approximation* (GGA) adds a dependence on the density gradient  $\nabla\rho$  and hence becomes “semilocal” as it also includes some information on the curvature of  $\rho$ . The higher flexibility of the GGA form drastically improves its performance for atoms and molecules alike, but compromises still have to be made because not all desired properties of a DFA can be achieved with only  $\rho$  and  $\nabla\rho$ . In practice, this means that GGAs can be designed to achieve very good performance for certain properties or systems, at the cost of reducing their applicability to other problems. *Meta*-GGAs include the kinetic energy density  $\tau$ , and can exhibit very good across the board performance for the energetics and structures of solids, surfaces, and molecules.<sup>243,244</sup> Nevertheless, the first three (semi)local rungs still have some clear deficiencies. Band gaps in solids are consistently underestimated (sometimes by several eV), molecular systems tend to be overbound, and reaction barrier heights are too low.

*Hybrid functionals* (sometimes referred to as *hyper-GGA*) are on the fourth rung, and introduce an explicit dependence on the occupied Kohn–Sham orbitals by replacing a part of the exchange energy by a Hartree–Fock-type “exact” exchange term. Hybrids largely solve the issues of semilocal functionals and have since the late 1990s become the standard in chemistry. The HF exchange term is however very expensive to evaluate in the solid state, hampering the widespread adoption of hybrid functionals in materials science. A last rung, populated by *double hybrid* functionals, holds DFAs that also use *unoccupied* orbitals, in the form of a Møller–Plesset-type second order perturbation (MP2) term or the random phase approximation (RPA) that replaces part of the DFT correlation energy. Unlike DFAs on the other rungs, double hybrids are able to capture long-range van der Waals (vdW) interactions. They are very expensive to evaluate, while accurate treatments of the vdW interaction can just as well be added on top of semilocal DFAs or their hybrids in the form of an additional energy term<sup>245</sup> (as we have consistently done in the two previous chapters).

A specific point of interest to us in this discussion is the band gap problem of semilocal DFAs. Wide-gap materials are more difficult to charge up, and the work of charging might be correlated to their chemical properties. If semilocal DFAs cannot get the gap right, can they still be trusted for other properties of the charged material? In more general terms, to what extent does the choice of



**FIGURE 7.1:** Graphene/*h*-BN heterojunction used as narrow-gap derivative of *h*-BN. Atom color codes: gray, carbon; blue, nitrogen; pink, boron.

any DFA matter, i.e., do functionals within the same rung perform similarly? We therefore revisit CO<sub>2</sub> adsorption on charged *h*-BN—a wide-gap material—with different representative DFAs, and compare it to the same process on a graphene/*h*-BN heterojunction, which has the same surface chemistry but a greatly reduced gap.

## 7.2 Methodology

As in the previous chapters, all calculations were carried out with CP2K in the GPW formalism, and a Martyna-Tuckerman solver was used to impose mixed boundary conditions. Dispersion corrections were not available for all tested DFAs and have therefore been omitted, unless they are an integral part of the method. To minimize basis set incompleteness errors, the Kohn–Sham orbitals were expanded in a large doubly polarized triple- $\zeta$  (m-TZV2P)<sup>207</sup> basis set, while the cutoff for the density was 1000 Ry. Some functionals unavailable in CP2K were evaluated through the LIBXC 3.0.1 library,<sup>246</sup> and in those cases the cutoff was increased to 1500 Ry. Computation of exact exchange was made feasible by the auxiliary density matrix method (ADMM),<sup>247</sup> employing a small uncontracted polarized triple- $\zeta$  (pFIT3) basis in most cases.

The *h*-BN sheet was the same as in Chapter 5, and both charges ( $2e^-$ , excess electron density of about  $1.0 \times 10^{18} \text{ m}^{-2}$ ) and external fields (0.05 a.u.) were considered. The graphene/*h*-BN heterojunction consisted of a  $6 \times 6$  graphene sheet joined along the armchair edge with a same-sized *h*-BN counterpart,<sup>248</sup> carry-

ing a  $6e^-$  charge<sup>191</sup> ( $\sim 1.5 \times 10^{18} e^- \text{ m}^2$ , Figure 7.1). Energy gaps were calculated as the difference between the HOMO and LUMO Kohn–Sham orbitals, and all systems were closed-shell.

Structures, gaps, and energies were evaluated using several approximations of the exchange–correlation energy. Some of the most widely used workhorse semilocal functionals (PBE,<sup>208</sup> revPBE,<sup>230</sup> RPBE,<sup>232</sup> BLYP,<sup>249,250</sup> HCTH/407,<sup>251</sup> TPSS<sup>233</sup>) were selected and supplemented with more recent developments that claim improved performance for solids (AM05<sup>252</sup>, PBESol<sup>253</sup>, SOGGA<sup>254</sup>), molecules (GAM<sup>255</sup>), or both (revTPSS<sup>256</sup>, N12<sup>257</sup>, BEEF-vdW<sup>258</sup>). They are constructed with different application and principles in mind, and either derived from fundamental constraint satisfaction (PBE, RPBE, TPSS, revTPSS, AM05, PBESol, SOGGA), minor empiricism (revPBE, BLYP), or extensive fitting against reference energies and structures (HCTH/407, N12, BEEF-vdW, GAM). Only a small set of meta-GGAs (TPSS and revTPSS) is considered due to apparent difficulties of CP2K to handle numerically sensitive  $\tau$ -dependent modern functionals such as the Minnesota functionals<sup>259</sup> (of the MXY-L or MNXY-L type) or SCAN.<sup>260</sup>

Hybrid functionals were represented in the form of the PBE0<sup>261,262</sup> and HSE06<sup>263,264</sup> approaches. Both are based on PBE and include a theoretically justified 0.25 fraction of exact exchange.<sup>265</sup> PBE0 is a global hybrid, i.e., a fraction of the GGA exchange energy is completely replaced by an equivalent quantity of exact exchange. Such a functional form can be easily applied to molecular systems, but is ill-defined in a periodic solid, which is why the HF term was truncated for interelectronic distances beyond 6.5 Å (slightly shorter than half of the cell edge of the *h*-BN sheet); long range corrections at the GGA level were included, an approach known as PBE0-TR-LRC.<sup>266</sup> The HSE06 functional uses a screened exact exchange term that interpolates from PBE0-like behavior for short-range interactions, and standard PBE in the long range. The screening parameter  $\omega$  controls the effect of the exact exchange term: a smaller  $\omega$  makes its effective range longer and more expensive to evaluate. A value of  $\omega = 0.11$  a.u. is recommended for HSE06, mostly for practical reasons.<sup>264</sup> We tested the effect of  $\omega$  in a range 0.11–0.25 a.u., which is known to be small for thermochemistry, but large for gaps, and therefore an interesting complication to assess.<sup>264</sup>

We also tested two approaches that attempt to improve performance for band gaps at the semilocal level. First, the DFT+U method<sup>267</sup> applies a correction to

the Coulomb and exchange integrals of selected subshells and can cure electron overdelocalization and underestimated band gaps in semilocal DFT. Application of DFT+ $U$  in the commonly used form of Dudarev *et al.*<sup>268</sup> requires an empirical selection of the relevant subshells and the value of the parameter  $U$ , which is system- and property-dependent. Here, we followed Verma & Truhlar,<sup>269</sup> and applied the  $+U$  correction to the  $p$  orbitals of B, using a general value of  $U = 4$  eV, on top of PBE ( $U = 2$  eV gave very similar results). Second, increasing the weight of the exchange term can also open up the calculated gap. In this *high local exchange* (HLE) method, the larger exchange term (scaled by 1.25) is balanced by a reduced correlation contribution (scaled by 0.5).<sup>270</sup> We have tested HLE variants based on PBE (PBE-HLE) and TPSS (TPSS-HLE, previously published as HLE17<sup>269</sup>).

It must be stressed that these tests are not benchmarking, because there is no experimental (or high-level quantum mechanical) reference to which any of the computed adsorption energies can be compared. Instead of verifying whether a method gives the right answer, it is even more important to gauge the consistency and predictability of a method. If any systematic trends or general behavioral patterns can be discerned, the underlying physics will be meaningful, and it should be possible to establish good practices.

## 7.3 Results

### 7.3.1 Wide-gap material: $h$ -BN

The computed gaps and adsorption energies for  $h$ -BN in Table 7.1 show how all of these values have a rather large functional dependence. Some of these are well-known and could be anticipated. For example, methods known to cure the gap underestimation by semilocal functionals also predict larger gaps in our calculations: PBE+ $U$  and HLE methods open the gap by  $\sim 0.5$  eV, and hybrids even up to 2 eV. Benchmarks have indicated that global hybrids tend to overcorrect, and that the screened HSE06 (with  $\omega = 0.11$ ) appears to give the most reliable gaps.<sup>271</sup> This seems to be case here, too. The direct band gap of  $h$ -BN has been experimentally estimated to be 5.97 eV,<sup>272</sup> most closely matched by the value of 6.05 eV we obtained with the standard HSE06 functional.

It is instructive to first discuss CO<sub>2</sub> adsorption under influence of an external electric field. Such a process concerns neutral systems and is an electroni-

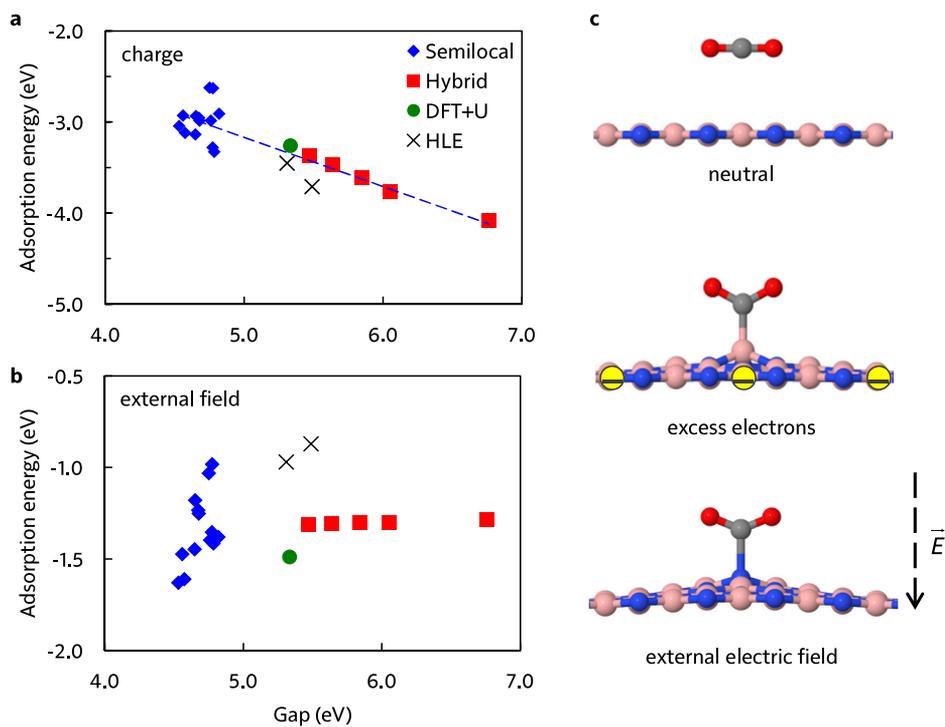
**TABLE 7.1:** Computed gaps and charge- and field-enhanced CO<sub>2</sub> adsorption energies on the wide-gap *h*-BN sheet using different density functional approximations. All energies are in eV.

	energy gap	$E_{\text{ads}}$ charged surface	$E_{\text{ads}}$ external field
<i>semilocal</i>			
PBE	4.65	-3.14	-1.45
revPBE	4.68	-2.96	-1.23
RPBE	4.68	-2.99	-1.25
BLYP	4.76	-2.99	-1.40
HCTH/407	4.78	-2.63	-0.98
AM05	4.56	-2.93	-1.47
PBEsol	4.57	-3.12	-1.61
SOGGA	4.53	-3.04	-1.63
N12	4.65	-2.94	-1.18
GAM	4.75	-2.63	-1.03
BEEF-vdW	4.82	-2.91	-1.38
TPSS	4.77	-3.28	-1.35
revTPSS	4.78	-3.33	-1.42
<i>DFT+U</i>			
PBE+U	5.33	-3.26	-1.49
<i>high local exchange</i>			
PBE-HLE	5.31	-3.45	-0.97
TPSS-HLE	5.49	-3.71	-0.87
<i>hybrid PBE</i>			
HSE ( $\omega = 0.25$ )	5.47	-3.37	-1.31
HSE ( $\omega = 0.20$ )	5.64	-3.47	-1.31
HSE ( $\omega = 0.15$ )	5.85	-3.61	-1.30
HSE ( $\omega = 0.11$ )	6.05	-3.76	-1.30
PBE0	6.76	-4.08	-1.29

cally simpler chemisorption problem that should hence expose the more obvious differences between methods. Most DFAs follow established trends, with PBE and its derivatives behaving as expected: revPBE, RPBE and PBE's hybrids exhibit reduced binding energies, while these are increased by the DFAs for solids PBESol and SOGGA; as mentioned earlier, GGAs cannot simultaneously be highly accurate for molecules, surface chemistry, and solid structures (but meta-GGAs can<sup>243,256</sup>). The screening parameter  $\omega$ , with its large impact on the gap, has little bearing on thermochemistry and, as shown, adsorption on the neutral surface. HCTH/407 and GAM, primarily fitted against molecular data, and the HLE functionals, designed for improved performance for gaps, show up as relative outliers, both predicting smaller adsorption energies (up to  $\sim 0.5$  eV below PBE). In other words, "specialized" GGA functionals give results that are inconsistent with more broadly conceived methods. Without these outliers (i.e., no DFAs specifically optimized for solids or molecules only), the spread on the semilocal methods is only 0.27 eV. Including all of them bumps it to 0.65 eV (or even 0.76 eV if the HLE functionals are also considered).

The spread between methods becomes even larger in the charged system, with adsorption energies occupying a range of almost 1.5 eV. Here, a functional's performance for the gap and charge-enhanced adsorption behavior is related, in the sense that bigger predicted gaps go hand in hand with stronger adsorption. Indeed, Figure 7.2a depicts a rather striking linear correlation between the two quantities, irrespective of the actual underlying mathematics and physics on which the functionals are based. Standard semilocal functionals are all clustered within a narrow range of predicted gaps and adsorption energies, irrespective of rung, whereas results from hybrids are much more spread out. Within the semilocal DFAs, even the solid state functionals PBESol and SOGGA are now indistinguishable from their more generally applicable siblings, although HCTH/407 and GAM remain outliers. PBE+U and the HLE functionals are stranded halfway between semilocal and hybrid DFAs with respect to predicted gaps and charge-enhanced adsorption energies.

The predicted *band gap* is strongly influenced by the self-interaction error of a DFA, and methods that correct for self-interaction (either through inclusion of exact exchange or empirical approaches such as DFT+U) tend to give more realistic values for the band gap. Self-interaction leads to additional problems upon charging: now, also *adsorption energies* are strongly affected. This is perhaps in-



**FIGURE 7.2:** Correlation between predicted gap and CO<sub>2</sub> adsorption energies on a wide-gap *h*-BN sheet. (a) Adsorption energies on the charged surface and (b) surface subjected to an external electric field. Density functional approximations are grouped by type, and a linear fit from all data points is shown as a dashed blue line. (c) Depiction of all relevant structures.

direct effect, in the sense that an incorrect band gap also leads to an incorrect energy of excess electrons that end up in the conduction band, explaining the correlated neutral band gap and chemisorption energies on the charged material.

Figure 7.2b shows that predicted properties for neutral surfaces are not correlated with the predicted gap: the fitted slope is only 0.07, with a similar-sized standard deviation, as opposed to the slope of  $-0.54 \pm 0.07$  that we find for the charged *h*-BN. The differences between PBE and any of its hybrids are much smaller than those within the group of semilocal DFAs, implying that the self-interaction error has little impact on the chemistry of the neutral material.

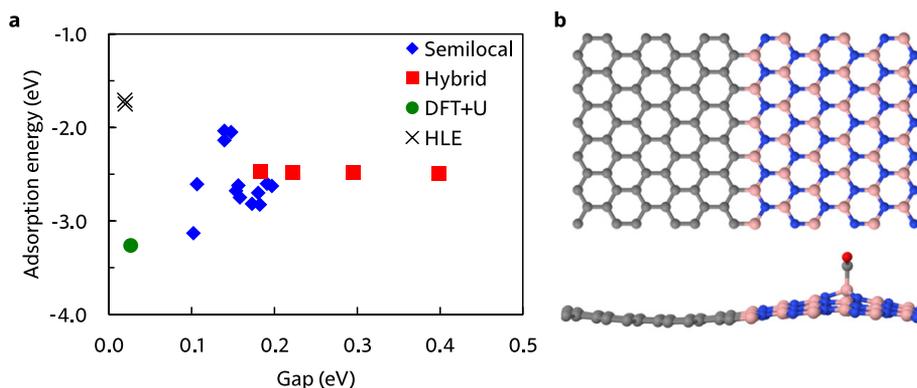
### 7.3.2 Narrow-gap material: graphene/*h*-BN heterojunction

A very small band gap of 0.01 eV has been reported for the particular graphene/*h*-BN heterojunction considered here, albeit at the PBE level.<sup>248</sup> Rather than computing the gap from the band structure, we approximate it has the HOMO-LUMO gap and obtain 0.16 eV as the PBE prediction, as given in Table 7.2. Other semilocal functionals only differ by at most 0.06 eV, while even hybrids do not add more than 0.3 eV to any of the (M)GGA values. (We were unable to converge any PBE0 calculation on this system.) The inclusion of exact exchange already has little impact on gaps, and its effect on charge-enhanced adsorption is even smaller. That is, the screening parameter of HSE somewhat influences the gap, but it does not change the adsorption energy. In that sense, its behavior is more similar to that observed for *neutral h*-BN—HSE also reduces the PBE binding energy, as one expects for neutral systems.

Figure 7.3 graphically corroborates these conclusions, and demonstrates that the variability within the group of semilocal functionals is much larger than the difference between PBE and its hybrid. Here, too, HCTH/407 and GAM are outliers, now joined by N12 (which is essentially an earlier version of GAM); BLYP is also an outlier, but in a different direction. Interestingly, all of these functionals are empirical—to some degree derived from fitting to reference data—and do not respect the LDA limit (i.e., reduce to LDA for  $\nabla\rho \rightarrow 0$ ). This is also true of BEEF-vdW, although its performance for this system (and also *h*-BN) is in line with the various LDA limit-respecting semilocal and hybrid approaches (most of them nonempirical). Unlike the other empirical methods, however, the fitting procedure leading to BEEF-vdW recognizes the limitations of the GGA form,

**TABLE 7.2:** Computed gaps and charge-enhanced CO<sub>2</sub> adsorption energies on the narrow-gap graphene/*h*-BN heterojunction using different density functional approximations. All energies are in eV.

	energy gap	$E_{\text{ads}}$ charged surface
<i>semilocal</i>		
PBE	0.16	-2.75
revPBE	0.16	-2.62
RPBE	0.15	-2.68
BLYP	0.10	-3.13
HCTH/407	0.14	-2.04
AM05	0.19	-2.60
PBEsol	0.18	-2.70
SOGGA	0.20	-2.63
N12	0.14	-2.14
GAM	0.15	-2.05
BEEF-vdW	0.11	-2.61
TPSS	0.17	-2.82
revTPSS	0.18	-2.83
<i>DFT+U</i>		
PBE+U	0.03	-3.26
<i>high local exchange</i>		
PBE-HLE	0.02	-1.71
TPSS-HLE	0.02	-1.75
<i>hybrid PBE</i>		
HSE ( $\omega = 0.25$ )	0.18	-2.48
HSE ( $\omega = 0.20$ )	0.22	-2.48
HSE ( $\omega = 0.15$ )	0.30	-2.48
HSE ( $\omega = 0.11$ )	0.40	-2.49

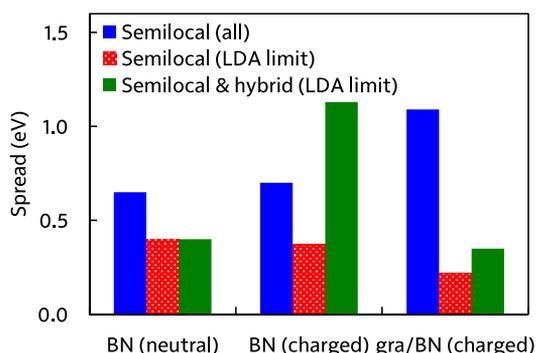


**FIGURE 7.3:** Correlation between predicted gap and charge-enhanced CO<sub>2</sub> adsorption energies on a narrow-gap graphene/*h*-BN heterojunction. (a) Adsorption energies on the charged surface and (b) depiction of all relevant structures.

and uses a consistent strategy to ensure that the final functional remains suitably balanced in its description of molecules, solids, and surface chemistry.<sup>258</sup> The explicit inclusion of a vdW-DF2-type nonlocal correlation component<sup>273</sup> does not appear to have a major impact on any of the properties calculated in this study.

PBE+U and the HLE functionals have shown reasonably consistent behavior on charged *h*-BN, but now they are not in line with any of trends followed by the semilocal or hybrid methods. Like on neutral *h*-BN, the HLE functionals underbind (compared to plain PBE or TPSS), while PBE+U overbinds. Remarkably, both methods counterintuitively tighten the gap, rather than open it. Functionals from both DFA classes show unpredictable behavior, and we cannot recommend them for charge-enhanced adsorption. All in all, the effect of self-interaction is minor in this system, as exemplified by the small difference between the semilocal PBE and its hybrid HSE06 for both band gaps and charge-enhanced adsorption. Deviations from the LDA limit have a significantly larger impact. It can be hypothesized that its narrow gap makes the graphene/*h*-BN heterostructure more metal-like, and a good DFT-based description of metals is usually associated with LDA-like DFAs.<sup>253</sup>

In general, any LDA limit-respecting performs very consistently for a given system, as shown in Figure 7.4. Inclusion of exact exchange has a major impact on the computed chemisorption energy only in the case of the charged wide-



**FIGURE 7.4:** Spread of predicted  $\text{CO}_2$  chemisorption energies (in eV) by various types of DFAs for the three different cases of charged and neutral materials considered in this work. The three groups of DFAs are: all semilocal functionals, only semilocal functionals that respect the LDA limit, and the latter functionals supplemented with the hybrids.

gap  $h$ -BN, where the effect closely tracks the predicted gap size. For all three systems, LDA limit-violating DFAs can be quite different from their asymptotically correct counterparts, but especially for the electronically challenging graphene/ $h$ -BN junction, where the tested semilocal DFAs give results over a range of 1.09 eV. In other words, the charged wide-gap  $h$ -BN sheet is a system in which the self-interaction error of semilocal DFAs has a decisive effect on chemisorption properties, whereas deviations from the LDA limit are the dominant potential source of error in the case of the charged narrow-gap graphene/ $h$ -BN heterostructure. As a consistent procedure, we therefore recommend a semilocal functional that respects the LDA limit, of which the predictions can be cross-checked against those of its hybrid in case of a wide-gap material. In other systems, semilocal methods can already offer a reliable window of results.

One additional remark, tying back to Chapter 5, should be made on the choice of material, charge state, and adsorption configuration, which were taken from the literature.<sup>191</sup> There, as in other studies of similar nature, full periodic boundary conditions were used instead of the appropriate partial periodicity. Single-point PBE calculations with a periodic 20 Å vacuum length on PBE-optimized geometries predict a charge-enhanced adsorption energy of only -1.24 eV, not even *half* of the converged result. The importance of a correct

**TABLE 7.3:** Computed adsorption energies of CO<sub>2</sub> on an Al<sub>2</sub>O<sub>3</sub>-supported Cu atom, comparing the plain PBE GGA and its hybrid HSE06 in the case of neutral and charged hydrated surfaces. All energies are in eV.

	$E_{\text{ads}}$ neutral surface	$E_{\text{ads}}$ charged surface
PBE	-0.14	-1.23
HSE ( $\omega = 0.11$ )	-0.36	-1.35

treatment of periodicity is hereby again demonstrated.

### 7.3.3 Revisiting plasma catalysis

The fact that charge-enhanced adsorption is highly sensitive to the employed DFT method, makes it important to verify whether our study of plasma catalysis in Chapter 6 is also affected. We can anticipate that the presence of transition metals will mitigate most of the potential issues: metal-containing systems have very small energy gaps (as also depicted in Figure 6.3), and the results in this chapter suggest that the impact of exact exchange will be small.

To test this assumption, we selected CO<sub>2</sub> adsorption on hydrated Cu/Al<sub>2</sub>O<sub>3</sub> as a sample to be reoptimized using both PBE and HSE06 ( $\omega = 0.11$ ). Just like the other systems in this study, a larger m-TZV2P basis set was used, but no additional dispersion corrections. The ADMM basis set for the exact exchange calculations was again pFIT3 for H, C, and O, and FIT9 and FIT12 for Al and Cu, respectively. Otherwise, the treatment of the surface charge was the same as in Chapter 6, using a neutralizing gas phase counterion.

At the GGA level, the gap of the supported catalyst model is only 0.24 eV and, as supported by Table 7.3, the impact of exact exchange is minimal. On the neutral surface, HSE predicts slightly stronger binding (by about 0.2 eV), and this difference is also found on the charged surface. It can therefore be inferred that our conclusions of Chapter 6 were not affected by our (computationally convenient) choice of DFT method, at least with respect to the presence of exact exchange.

## 7.4 Conclusions

The rather young discipline of charge-controlled surface science has not yet adopted a thoroughly verified set of standard practices, mostly due to the lack

of good benchmarking. Following our efforts in the areas of appropriate boundary conditions (Chapter 5) and charge distributions (Chapter 6), we have tackled the choice of density functional approximation. We find that computed adsorption energies on charged electrocatalytic materials are very sensitive to the level of theory employed, especially the wide-gap  $h$ -BN tested in this study. The performance of a DFA for the electronic gap of the neutral material is directly correlated with its predicted adsorption capabilities upon charging. However, such trends are not found in a narrow-gap analogue of the same material.

Interestingly, the variance within a class of DFT approximations can be very large, especially for empirical functionals geared more towards molecular systems and do not enforce the LDA limit. Furthermore, specialized semilocal approaches like DFT+U and HLE, which are intended to perform well for band gaps, exhibit very erratic behavior in some cases. Hence, it can be hypothesized that self-interaction is strong in charged wide-gap materials—explaining the fairly consistent behavior of hybrid, DFT+U, and even HLE methods—whereas consistent treatment of narrow-gap materials mostly benefits from correct enforcement of the LDA limit.

Due to lack of appropriate reference values, we cannot determine the “best” computational approach. However, the screened hybrid functional HSE06 is known to perform well for thermochemistry and solid state electronic structure alike, and its predictable and consistent improvements upon its semilocal parent PBE—which cannot be said about PBE+U and PBE-HLE—make it the most reliable, transparent, and affordable benchmark method. For future studies of charge-controlled adsorption, we recommend application of HSE06 as a consistency check of GGA calculations on any new system; the narrow-gap systems of Chapter 6 were anticipated to be not very sensitive to the effect of exact exchange, which was confirmed by a sample check with HSE06. On a broader note, we have demonstrated that the development of an appropriate computational strategy to a novel class of systems is not straightforward, and that the quality of well-worn standard approaches should never be simply taken for granted.

## CHAPTER 8

# General conclusions

Atomistic models are inherently small in scale and narrow in focus, while most practical scientific challenges are large and multifaceted. Such a mismatch must be addressed if a proper exchange between these two extremes is desired. A small-scale atomistic model cannot be used to answer big questions, so either the questions asked must be simplified to fit the model, or the model's scale must be somehow increased. The former approach is a core principle of scientific research, but still requires an intimate understanding of the model's strengths and limitations. The latter approach requires development of new modeling strategies, and is hence even more complicated. In this thesis, two case studies were presented, each exploring a particular attempt to bridge theory and experiment.

In the first part of this thesis, we introduce a new accelerated molecular dynamics method, dubbed *collective variable-driven hyperdynamics* (CVHD). MD simulation can provide explicit atomistic insights in chemical reactions, but are confined to microscopic length and time scales. Accelerated MD methods can in principle ameliorate this problem, but existing approaches are unfortunately limited to rather simple systems or specialized applications. These issues stem from overly ambitious goals, that is, extending the model (to longer timescales) while simultaneously “thinking big” (attempting to produce the full evolution of the system).

Much more widespread are the so-called enhanced sampling methods which can be applied to almost any system, but are designed to study specific processes which must be known beforehand, rather than entire reaction systems. In other words, the scope of the model is extended, which is made feasible by a deliberate simplification of the problem (in the form of a small set of CVs). CVHD is an approach that significantly improves the flexibility and performance of the

hyperdynamics accelerated MD method by recognizing and exploiting its similarities with the metadynamics enhanced sampling method. From metadynamics, CVHD borrows the self-consistent automatic generation of a bias potential and the generic representation of reaction coordinates in the form of collective variables. When spliced in the hyperdynamics algorithm, these building blocks form a method that needs little *a priori* user input, but can be applied to reactive process in different environments. That is, CVHD ends up following through on the promises of accelerated MD, but only thanks to more thorough simplification: complete reaction networks are obtained only by breaking up the complex problem into sequence of much simpler metadynamics-like simulations.

Initial validations on simple test systems demonstrate the flexibility, accuracy, and promise of the method: for processes as diverse as surface diffusion, conformational sampling and heterogeneous catalysis large accelerations—up to  $10^9$ , or macroscopic time scales of seconds—and excellent consistency with MD or other accelerated MD methods. The unrivaled abilities of CVHD are further established by its successful application to the pyrolysis and combustion of the *n*-dodecane model fuel: for the first time, an atomistic simulation is shown to directly produce the chemical evolution of a complex reacting system over macroscopic time scales in a fully self-consistent fashion. CVHD can efficiently handle reaction networks involving thousands of atoms, and capture their dependence on a wide range of process conditions, and has already found applications beyond the scope of this thesis, in fields as diverse as surface etching, catalysis, polymer degradation, and rate calculations. Application of CVHD to the simulation hence leads to a more balanced model, in which accurate models of the interatomic potential can finally be applied in dynamic atomistic studies that approach realistic reaction conditions.

All of the early applications of CVHD were based on empirical classical force fields. Although a computationally cheap choice, the limited transferability of these interaction models has somewhat hampered the full exploration of CVHD's application domain. For this reason, the method was also applied to simulations based on quantum mechanical (QM) forces, i.e.,  $\text{CO}_2$  splitting on alumina-supported metal catalysts modeled by DFT. These simulations revealed that the inherent time scale limitation of DFT-based MD simulations—a few tens of ps—cannot be easily cured by CVHD, which has an overhead of at least several ps taken up by bias growth and reset, and event detection. In the

hope of further improving the performance, a different biasing strategy based on variationally enhanced sampling was therefore tested—the integration of which was made trivial by the flexible nature of CVHD. With this variational version of CVHD, good performance was obtained within an even more flexible biasing strategy but the fact that single events still took ~100 ps of MD time remains problematic.

Thanks to the development of CVHD, the MD time scale has caught up with the development of interatomic potentials for the study of reaction systems: efficient force fields can now be used in high-quality simulations of realistic process conditions. Rather than directly attempting to further push the research field towards more general but also more expensive *ab initio* dynamics, it would make more sense to consider cheaper semi-empirical QM models. Indeed, there is a new trend in computational chemistry to replace force field approaches by novel “low cost” QM methods that combine a solid physical basis with limited empiricism and which are readily applicable to most elements of the periodic table. Only when more complex systems with more interesting chemistry, such as reactions in solution and processes at interfaces become more routinely accessible, further development of CVHD would make sense.

In the second part of this thesis, models of charged catalyst surfaces were constructed. Plasma catalysis is made up of many intertwined phenomena which are impossible to untangle experimentally. No atomistic model can represent the full plasma-catalyst interaction, and any modeling effort of this process will necessarily be an exercise in “thinking small,” which is exactly what an understanding of the experiment requires. A little-investigated effect, plasma-induced negative charging (i.e., injection of excess electrons by the plasma) can change the electronic structure of a catalyst and therefore also its reactivity. To be able to capture these effects, QM methods must be employed, *in casu* based on DFT.

A tightly focused computational foray into one specific plasma-catalytic phenomenon might appear straightforward, but looks are deceiving. The construction of an appropriate model of a plasma-charged catalyst has revealed four key difficulties, which all were successfully addressed in this thesis:

**Treatment of periodic boundaries is problematic.** Prior computational investigations demonstrating the potential of charge-tuned gas adsorption by electrocatalytic materials show the promise of both the charge effects, as

of QM modeling to study them. However, it was revealed that an incorrect periodicity of the simulation cell has a huge impact on predicted adsorption energies, and the qualitative trends across materials can be off. Such an error in the model is worse than a poor correspondence with experiments: it leads to a model inconsistent with itself. Simple corrective measures are available, though, and were shown to be sufficient.

A proper model of the charge distribution must be developed. A solid immersed in a plasma can never be modeled as an isolated charged object. Therefore, a simple atomistic model mimicking the interaction between the plasma-charged surface and the plasma sheath was developed.

Supported catalysts are difficult to deal with. Plasma catalysis is typically carried out with supported metal catalysts, which are very difficult to represent in an atomistic model. Our compromise to use single-atom catalysts allowed to capture the effect of charging on different metals in a comparatively simple system.

Common QM methods applied in catalysis are not good enough. DFT functionals give wildly different descriptions of charge-tuned gas adsorption processes. The problem is the most pronounced for wide-gap materials, for which a strong correlation between predicted band gaps and charge-tuned adsorption energies are observed. Given the well-documented problems of DFT in its description of gaps, its performance for the surface chemistry of charged materials should be vetted just as judiciously. Narrow-gap materials (such as metal catalysts) are not as problematic in this respect, but the choice of DFT method can still have an unreasonably large impact on the results. Hence, consistency checks must always be carried out to have an idea of the uncertainties associated with the calculation, and one should not rely on rote methodologies simply out of habit.

The ability of models to reduce real-life complexity does not mean the actual construction of these models is a simple task. However, when properly accounting for all possible pitfalls, exciting new insights become accessible: the importance of the plasma-induced surface charge has been made plausible. In particular, the fact that relative activities of catalyst material can be changed

through charging is a clear example of the speculated ability of plasma catalysis to break the scaling laws of traditional catalysis. Further extension of the model to larger chemistry sets will be useful to further assess the possible impact of plasma catalysis on the field. Along similar lines, other plasma effects could be incorporated, such as the presence of excited molecules. Such undertakings will not be straightforward, but as we have shown, existing methodology infused with some creativity goes a long way.

All science is transient by nature, building on top of previous work and unknowingly paving the way for future inventions. This thesis has presented two clear advances on the state of the art in computational chemistry and will hopefully inspire new developments. Possibilities are of course the extension of CVHD to new systems and processes, and the further refinement of atomistic models of plasma catalysis. No past, current, or future model will ever be perfect, and that's OK: as long as new ideas keep coming—and they will—we'll be just fine.



# Summary

Over the last decades, computational chemistry has become a mature and widely practiced discipline of the chemical sciences. Computational methods run the gamut from molecular dynamics (MD) simulations with approximate classical potentials to high-level quantum mechanical electronic structure calculations. At their best, computational models provide a fundamental peek into atomic-level properties of matter and greatly improve our understanding of processes all across chemistry. It is sometimes even possible to predict the outcome of an experiment, besides being able to give a rationalization of it afterward. However, atomic-level models are naturally limited to microscopic time and length scales. A key difficulty of computational investigations is hence their inability to match the size and complexity of most experiments. This poses an important challenge to computational chemists: how to achieve a useful interplay between model and experiment?

In this thesis, new ways to bridge theory and reality are investigated. In particular, two case studies are presented. In the first part of this thesis, a new *general method* to extend the short time scale of atomistic simulations is developed: *the model is brought closer to the experiment*. The second part of this thesis concerns the construction of a new model for the interpretation of a *specific* experiment: *the model simplifies the experiment*. These two approaches represent two sides of the same coin. On one hand, the limited scale of the model can be perceived as a weakness, and therefore approaches must be developed that allow to extrapolate simulated microscopic molecular dynamics to experimentally measurable macroscopic observables. On the other hand, the former idea can also be turned on its head: in this view, the model can be a “pure” representation of a phenomenon, whilst macroscopic experiments are only the sum of many different microscopic components, none of which can be studied on its own—in that sense, the model is the bottom-up inverse of a top-down experimental approach. This philosophy is applied to the interaction between a plasma and

catalyst: a model is constructed to specifically investigate the plasma-induced changes of the catalyst's electronic structure, i.e., how a plasma can change the chemical properties of a catalyst. Both cases, however, clearly show that model and experiment can and should co-exist: precisely because of their vastly differing scales, they offer different but complementary views of the same problem.

Chapter 1 discusses the current literature on the MD time problem—with its various solutions—and a discrepancy is pointed out. One class of long time scale methods (accelerated MD) has the goal of accurately producing the time evolution of a reactive process—so it can *predict* the outcome of an arbitrary process—but is generally limited to a handful of simple phenomena in the solid state. The other group of methods (enhanced sampling) is much more flexible, but can typically only be applied to specific processes rather than full reaction systems. Strong similarities between *hyperdynamics* (an accelerated MD method) and *metadynamics* (an enhanced sampling method) are pointed out, and the possibilities of combining their respective strengths and application domains is discussed. Hence, a missing element in the toolkit of molecular dynamics techniques is identified.

Chapter 2 describes a newly developed accelerated MD method, dubbed *collective variable-driven hyperdynamics* (CVHD). Like hyperdynamics, CVHD uses a bias potential to destabilize metastable states, and accelerate the occurrence of activated reactions. Like metadynamics, the reaction coordinates of the accelerated processes are represented in the form of a simple collective variable (CV), and an appropriate bias potential is self-consistently generated during simulation, requiring little *a priori* user input. Initial validations on simple test systems demonstrate the flexibility, accuracy, and promise of the method: for processes as diverse as surface diffusion, conformational sampling and heterogeneous catalysis large accelerations—up to  $10^9$ , or macroscopic time scales of seconds—and excellent consistency with MD or other accelerated MD methods. By reconciling metadynamics and hyperdynamics, a new generic accelerated MD method can thus be created.

Chapter 3 takes CVHD further than any of its competitors. It is applied to pyrolysis and combustion of the *n*-dodecane model fuel—unprecedented time scales of several seconds could be reached, leading to simulations of which the process conditions (temperature and pressure) closely match actual technological combustion applications. With CVHD realistic reaction networks and mech-

anisms can be directly simulated, establishing a strong link between simulation and experiment. Also, other applications of CVHD are summarized, demonstrating the broad usefulness of CVHD in fields as diverse as surface etching, catalysis, polymer degradation, and rate calculations. Application of CVHD can hence greatly extend the range of established modeling tools and techniques, and facilitate their validation against the experiment.

Chapter 4 further explores the flexibility of CVHD. A different bias strategy, based on variationally enhanced sampling rather than metadynamics, is tested in order to assess CVHD's applicability to even more complex processes and advanced quantum mechanical force fields. The performance of variational CVHD for rate calculations of a simple chemical reaction is found to be superior to other recent methodologies. These results inspire confidence in the ability of CVHD to be further adapted for new, challenging problems down the road.

Chapter 5 marks the start of this thesis' second part, and first introduces the complex reality of plasma catalysis. This combined application of heterogeneous catalysts with plasma technology holds great promise for greenhouse gas conversion and fuel production in a modern electrified society, but its fundamental underpinnings remain poorly understood. The many different types of plasma-catalyst interaction can be explicitly untangled through the construction of simplified models, each specifically dedicated to a much more limited subset of phenomena. Catalyst charging by plasma-supplied electrons is perhaps the most mysterious plasma-surface interaction, and it is decided to construct a model for it at the density functional level of theory (DFT). From a review of the literature, several studies of gas adsorption on charged low-dimensional nanomaterials are found. An attempt to reproduce these studies however reveals large errors in all literature results, stemming from poorly thought-out material models. It is shown that a correct treatment of the periodicity of the system, which removes an unphysical interaction of the charged surface with itself, is necessary to achieve internal consistency in the model. This first step towards a high-quality model of the plasma-induced surface charge hence already demonstrates the treacherous nature of atomistic models, and is only one of the many pitfalls along the way.

Chapter 6 extends the surface charge model to actual plasma-catalytic conditions. A simple atomistic model of the charge distribution around the surface—the plasma sheath—is developed and validated for internal consistency. Further-

more, the standard oxide-supported transition metal catalysts typically used in plasma catalysis are included in the model in the form of single atom catalysts rather than larger clusters: the precise structure of the catalyst is simplified in the model while still being able to capture most of catalyst-surface interaction. This model is specifically applied to investigate plasma-catalytic CO<sub>2</sub> activation on supported M/Al<sub>2</sub>O<sub>3</sub> (M = Ti, Ni, Cu) catalysts. It is found that (1) the presence of a negative surface charge dramatically improves the reductive power of the catalyst, strongly promoting the splitting of CO<sub>2</sub> to CO and oxygen, and (2) the relative activity of the investigated transition metals is also changed upon charging, suggesting that controlled surface charging is a powerful additional parameter to tune catalyst activity and selectivity. These results strongly point to plasma-induced surface charging of the catalyst as an important factor contributing to the plasma-catalyst synergistic effects frequently reported for plasma catalysis.

Chapter 7 delves deeper into the quality of the surface charge model. Here, the choice of electronic structure method is investigated. It is found that the choice of density functional approximation can have a very large impact on the predicted adsorption characteristics on the charged material. In contrast to standard adsorption chemistry on neutral surfaces, computed adsorption energies on wide-gap materials are correlated with the predicted band gap of the material. On the other hand, charge-enhanced adsorption on small-gap materials is treated inconsistently by many standard approaches. A practical and rather convenient validation of any standard density functional calculation is comparison to other common methods, in particular hybrid functionals: one can be fairly certain of a methodology if all these methods, although based on different physical principles, give similar answers. Most importantly, it must be stressed that the design of models for new problems cannot simply rely on standard solutions that might have worked in the past: it is sensible to use an existing approach as a starting point, but it should never be taken for granted, and extensive cross-checks will always be needed.

The general conclusion of this work is that there are plenty of possibilities to improve the quality of molecular modeling approaches, and extend the range of computational chemistry. New application domains and algorithmic innovations can always be found, even for mature techniques, if there is a willingness to look beyond established practices. This holds true for the investigations pre-

sented in this thesis. The first part shows how lateral thinking can push the simulation closer to the experiment, while the second part demonstrates the many complications arising from simplifying the experiment by translating it to an atomistic model. In the end, the enormous possibilities of atomistic models will continue to be explored. Research is inherently transient; this thesis draws from many influences that came before and hopefully, it will provide useful building blocks in future work.



# Samenvatting

Computationale chemie is geëvolueerd tot een volwassen en wijdbedreven subdiscipline van de chemie die een groot aantal technieken omvat, elk met hun eigen sterktes en toepassingsgebieden. Computationale modellen bieden een fundamenteel inzicht in de eigenschappen van materie, en een beter begrip van chemische processen. De uitkomst van een experiment kan in detail worden verklaard, en soms is het zelfs mogelijk om het precieze verloop ervan te voorspellen. Modellen met een dergelijk atomair detail zijn echter ook beperkt tot microscopische tijd- en lengteschalen. Er is dan ook vaak geen directe één-op-éénrelatie mogelijk tussen computationale studies en experimenten. Het bekomen van een goede wisselwerking tussen relatief eenvoudige microscopische modellen enerzijds, en de complexe macroscopische realiteit anderzijds, is om deze reden de grootste uitdaging voor de computationale chemie.

Deze thesis presenteert nieuwe manieren om de kloof tussen theorie en realiteit te overbruggen. Twee casussen worden in detail uitgewerkt. In het eerste deel van de thesis wordt een nieuwe *algemene methode* ontwikkeld om de tijdschaal van atomistische simulaties te verlengen: *het model wordt dichter naar het experiment gebracht*. Het tweede deel betreft de constructie van een nieuw model om een *specifiek* experiment te interpreteren: *het model vereenvoudigt het experiment*. Deze twee werkwijzen vertegenwoordigen twee kanten van dezelfde medaille. Aan de ene kant kan de beperkte schaal van het model als een zwakte worden gezien, zodat het noodzakelijk is om op moleculair niveau gesimuleerde dynamica met nieuwe technieken te extrapoleren naar experimenteel meetbare macroscopische grootheden. Aan de andere kant kan deze redenering ook worden omgedraaid: men kan stellen dat het model een meer “zuivere” voorstelling is van een fenomeen, terwijl macroscopische experimenten daarentegen steeds bestaan uit de som van zeer veel verschillende microscopische componenten, waarvan geen enkele kan worden geïsoleerd. In die zin is het model dan ook de “bottom-up” inversie van het “top-down” experiment. Concreet

wordt er in deze thesis een model ontwikkeld om te bestuderen hoe een plasma de elektronische eigenschappen, en dus de reactiviteit, van een katalysator kan modificeren: slechts een klein stukje van de totale plasma-katalysatorinteractie, maar één waarover tot nu toe zeer weinig geweten is. Beide casussen tonen echter duidelijk aan dat model en experiment perfect naast elkaar kunnen bestaan: precies omwille van hun sterk verschillende schaalgroottes geven beide dan ook complementaire perspectieven op hetzelfde probleem.

Hoofdstuk 1 bespreekt de huidige literatuur over het MD-tijdschaalprobleem—met de verschillende oplossingen ervoor—en een discrepantie wordt opgemerkt. Eén klasse aan lange tijdschaalmethoden (*versnelde MD*) kan de tijdsevolutie van een reactief systeem produceren—en in principe de uitkomst van een arbitrair proces *voorspellen*—maar is in de praktijk beperkt tot een handvol eenvoudige processen in de vaste stoffase. De andere methodegroep (*enhanced sampling*) is veel flexibeler, maar kan alleen worden toegepast op een enkel proces, in plaats van volledige reactiesystemen. Markante gelijkenissen tussen *hyperdynamics* (een methode voor versnelde MD) en *metadynamics* (voor enhanced sampling) worden onderlijnd, en de mogelijkheden om hun respectieve sterktes en toepassingsgebieden te combineren worden besproken. Aldus wordt een ontbrekend hulpmiddel binnen de moleculaire dynamica geïdentificeerd.

Hoofdstuk 2 beschrijft een nieuwontwikkelde techniek voor versnelde MD, *collective variable-driven hyperdynamics* (CVHD). Net als hyperdynamics gebruikt CVHD een externe *biaspotentiaal* om metastabiele toestanden te destabiliseren en zo geactiveerde processen te versnellen. Net als metadynamics worden de reactiecoördinaten van de versnelde processen versimpeld in de vorm van een eenvoudige collectieve variabele (CV), en wordt een geschikte bias op een zelfconsistente wijze gegenereerd gedurende de simulatie. CVHD vereist zeer weinig *a priori* input van de gebruiker en is hierdoor eenvoudig toe te passen. Initiële validatiestudies voor simpele testsystemen demonstreren de flexibiliteit, nauwkeurigheid en belofte van de methode: een brede waaier processen zoals oppervlaktediffusie, conformationele analyse en heterogene katalyse kan sterk worden versneld—tot  $10^9$ , of macroscopische tijdschalen van seconden—in uitstekende overeenkomst met MD en andere methoden voor versnelde MD. De vereniging van metadynamics en hyperdynamics leidt dus tot een nieuwe, breed toepasbare implementatie van versnelde MD.

Hoofdstuk 3 toont dat CVHD verder kan gaan dan eender welke ander simulatiemethode. Toegepast op pyrolyse en verbranding van *n*-dodecane worden ongeëvenaarde tijdschalen van seconden bereikt: in dergelijke simulaties kunnen experimentele condities (temperatuur en druk) worden gesimuleerd. CVHD kan realistische reactienetwerken en -mechanismen rechtstreeks simuleren, en geeft zo een directe link tussen simulatie en experiment. Andere toepassingen van CVHD worden verder ook samengevat, en demonstreren de brede toepasbaarheid van de CVHD in diverse onderzoeksgebieden zoals het etsen van oppervlakken, katalyse, degradatie van polymeren, en bepaling van kinetische parameters. Toepassing van CVHD kan het bereik van standaard simulatietechnieken dus sterk vergroten, en hun validatie tegen experimenten faciliteren.

Hoofdstuk 4 exploreert de flexibiliteit van CVHD. Om CVHD toe te kunnen passen op nog complexere processen en geavanceerde kwantummechanische krachtvelden wordt een andere biasstrategie getest, gebaseerd op variationally enhanced sampling in plaats van metadynamics. Op een eenvoudige testreactie blijkt deze “variationale CVHD” superieure prestaties te leveren, wat vertrouwen geeft in de mogelijkheden om CVHD verder aan te passen voor nieuwe, uitdagende toepassingen.

Hoofdstuk 5, de start van het tweede thesisdeel, introduceert een nieuwe uitdaging: de complexe realiteit van plasmakatalyse. Deze gecombineerde toepassing van heterogene katalyse met plasmatechnologie is een veelbelovende methode die een toekomst kan worden gebruikt om op duurzame wijze broeikasgassen om te zetten in hernieuwbare brandstoffen. Er is echter nog weinig geweten over de fundamentele principes en processen die aan de basis liggen van deze techniek. Modellen kunnen de verschillende factoren waaruit de plasmakatalysatorinteractie bestaat ontrafelen en zo de complexiteit van het probleem verlagen. Omdat het opladen van de katalysator door plasma-elektronen waarschijnlijk het meest mysterieuze plasma-effect is, werd het gekozen als onderwerp van een model op basis van dichtheidsfunctionaaltheorie (DFT). In de literatuur worden verschillende studies gevonden van gasadsorptie op nanomaterialen, die als basis voor verder werk kunnen dienen. Een poging tot reproductie van deze studies onthult echter grote systematische fouten in alle literatuurrresultaten, een resultaat van slecht geconstrueerde materiaalmodellen. Hierop wordt aangetoond dat een correcte behandeling van de periodiciteit van het systeem noodzakelijk is om een intern consistent model te verkrijgen, één

zonder fysisch onzinnige interacties van het geladen oppervlak met zichzelf. Deze eerste stap richting een hoogwaardig model van de plasma-geïnduceerde oppervlaktelading demonstreert dus al de verraderlijke aard van atomistische modellen, en is slechts één van de vele mogelijke valkuilen onderweg.

Hoofdstuk 6 breidt het model voor oppervlakteladingen verder uit naar feitelijke plasmakatalytische condities. Een eenvoudig atomistisch model voor de ladingsverdeling rond het oppervlak—de plasmashield—wordt ontwikkeld en gevalideerd. Verder worden standaard op oxides afgezette metaalkatalysatoren in het model geïntegreerd, in de vorm van enkele atomen in plaats van grotere clusters: de complexe structuur van de katalysator wordt vereenvoudigd in het model, dat echter nog steeds de voornaamste interacties tussen katalysator en oppervlak omvat. Dit model wordt vervolgens gebruikt om plasmakatalytische  $\text{CO}_2$ -activatie op afgezette  $\text{M}/\text{Al}_2\text{O}_3$  ( $\text{M} = \text{Ti}, \text{Ni}, \text{Cu}$ ) katalysatoren te onderzoeken. Er wordt aangetoond dat de aanwezigheid van een negatieve oppervlaktelading (1) de reductieve eigenschappen van de katalysator dramatisch versterkt en de splitsing van  $\text{CO}_2$  in  $\text{CO}$  en zuurstofgas promoot, en (2) de relatieve activiteit van de verschillende metalen wijzigt, wat suggereert dat via gecontroleerd opladen de activiteit en selectiviteit van een katalysator kan worden gecontroleerd. Deze resultaten wijzen erop dat plasma-geïnduceerde oppervlakteladingen een belangrijke rol kunnen spelen in het realiseren van synergie tussen plasma en katalysator.

Hoofdstuk 7 gaat dieper in op de kwaliteit van het model voor oppervlakteladingen. De keuze van kwantumchemische methode wordt onderzocht, en blijkt een grote impact te hebben. Het is geweten dat verschillende DFT-technieken niet in staat zijn om accurate elektronische excitatie-energieën te berekenen, maar dat deze zwakte geen impact heeft op hun vermogen om adsorptie te beschrijven. Echter, voor geladen oppervlakken wordt er in sommige omstandigheden wél een duidelijke correlatie tussen berekende excitatie- en adsorptie-energie gevonden. Verder wordt er in het algemeen gevonden dat allerlei standaardmethoden inconsistente resultaten leveren. Een praktische validatiemethode van eender welke DFT-techniek is daarom de vergelijking met andere methoden, in het bijzonder hybride functionalen: als al deze methoden, hoewel gebaseerd op andere fysische principes, gelijkaardige voorspellingen geven kan men tot op zekere hoogte vertrouwen hebben in hun kwaliteit. Het dient vooral te worden benadrukt dat het ontwerp van methodologieën voor nieuwe

problemen niet volledig mag vertrouwen op standaard oplossingen uit het verleden: het is zinvol om een bestaande methodiek als uitgangspunt te nemen, maar deze kan niet klakkeloos voor waar aangenomen worden. Uitgebreide validaties zullen altijd nodig zijn.

De algemene conclusie van dit werk is dat er veel mogelijkheden bestaan om de kwaliteit van moleculaire modellen te verbeteren, en zo de mogelijkheden van computationele chemie uit te breiden. Nieuwe toepassingsdomeinen en algoritmische innovaties zijn steeds te vinden, zelfs voor volwassen technieken, zolang er de wil is om voorbij gekende praktijken te kijken. Dit geldt ook voor het onderzoek in deze thesis. Het eerste deel laat zien hoe lateraal denken de simulatie dichter naar het experiment kan brengen; het tweede deel demonstreert de vele complicaties die voortvloeien uit de vereenvoudigingen van het experiment tot een atomistisch model. Hoe dan ook zullen de enorme mogelijkheden van atomistische modellen verder worden geëxploreerd. Elk inzicht is tijdelijk, en wordt vervangen door iets beters; deze thesis baseert zich op vele technieken en zal hopelijk een inspiratie zijn voor toekomstig werk.



# List of publications

## Related to this thesis

1. D. U. B. Aussems, **K. M. Bal**, T. W. Morgan, M. C. M. van de Sanden and E. C. Neyts. Mechanisms of elementary hydrogen ion-surface interactions during multilayer graphene etching at high surface temperature as a function of flux. *Carbon* **137**, 527–532 (2018).
2. **K. M. Bal** and E. C. Neyts. Modelling molecular adsorption on charged or polarized surfaces: a critical flaw in common approaches. *Phys. Chem. Chem. Phys.* **20**, 8456–8459 (2018).
3. **K. M. Bal**, S. Huygh, A. Bogaerts and E. C. Neyts. Effect of plasma-induced surface charging on catalytic processes: application to CO<sub>2</sub> activation. *Plasma Sources Sci. Technol.* **27**, 024001 (2018).
4. D. U. B. Aussems, **K. M. Bal**, T. W. Morgan, M. C. M. van de Sanden and E. C. Neyts. Atomistic simulations of graphite etching at realistic time scales. *Chem. Sci.* **8**, 7160–7168 (2017).
5. E. C. Neyts and **K. M. Bal**. Effect of electric fields on plasma catalytic hydrocarbon oxidation from atomistic simulations. *Plasma Process. Polym.* **14**, e1600158 (2017).
6. **K. M. Bal** and E. C. Neyts. Direct observation of realistic-temperature fuel combustion mechanisms in atomistic simulations. *Chem. Sci.* **7**, 5280–5286 (2016).
7. **K. M. Bal** and E. C. Neyts. Merging metadynamics into hyperdynamics: accelerated molecular simulations reaching time scales from microseconds to seconds. *J. Chem. Theory Comput.* **11**, 4545–4554 (2015).

## Other publications

1. S. Huygh, A. Bogaerts, **K. M. Bal**, E. C. Neyts. The high coke resistance of a TiO<sub>2</sub> anatase (001) catalyst surface during dry reforming of methane. *J. Phys. Chem. C* **122**, 9389–9396 (2018).
2. **K. M. Bal**, J. Cautereels and F. Blockhuys. Structures and spectroscopic properties of sulfur-nitrogen-pnictogen chains: R<sub>2</sub>P–N=S=N–PR<sub>2</sub> and R<sub>2</sub>P–N=S=N–AsR<sub>2</sub>. *J. Mol. Struct.* **1132**, 102–108 (2017).
3. **K. M. Bal** and E. C. Neyts. On the time scale associated with Monte Carlo simulations. *J. Chem. Phys.* **141**, 204104 (2014).
4. W. Somers, A. Bogaerts, A. C. T. van Duin, S. Huygh, **K. M. Bal** and E. C. Neyts. Temperature influence on the reactivity of plasma species on a nickel catalyst surface: An atomic scale study. *Catal. Today* **211**, 131–136 (2013).
5. E. C. Neyts, B. J. Thijsse, M. J. Mees, **K. M. Bal** and G. Pourtois. Establishing uniform acceptance in force biased Monte Carlo simulations. *J. Chem. Theory Comput.* **8**, 1865–1869 (2012).

# List of presentations

## Contributions to conferences

1. K. M. Bal and E. C. Neyts. Bottom-up disentanglement of plasma-catalytic effects from atomistic modelling. *2018 Plasma Processing Science Gordon Research Conference*, Bryant University, RI, USA, August 5–10, 2018. **(invited talk)**
2. K. M. Bal and E. C. Neyts. Bottom-up disentanglement of plasma-catalytic effects from atomistic modelling. *2018 Plasma Processing Science Gordon Research Seminar*, Bryant University, RI, USA, August 4–5, 2018. **(poster presentation)**
3. K. M. Bal and E. C. Neyts. Atomistic modeling of carbon nanostructures: challenges and opportunities. *2018 EMRS Spring Meeting*, Strasbourg, France, June 18–22, 2018. **(invited talk)**
4. K. M. Bal and E. C. Neyts. From microscopic simulations to macroscopic time scales in reactive MD. *ADF Developer Meeting 2018*, Software for Chemistry & Materials (SCM), Vrije Universiteit Amsterdam, The Netherlands, March 26–29, 2018. **(invited talk)**
5. K. M. Bal, S. Huygh and E. C. Neyts. Effect of plasma surface charging on catalytic activation of CO<sub>2</sub>: Fundamental insights from quantum chemical modeling. *23rd International Symposium on Plasma Chemistry*, Montreal, Canada. July 30–August 4, 2017. **(poster and oral flash presentation)**
6. K. M. Bal, S. Huygh and E. C. Neyts. Effect of plasma surface charging on the catalytic decomposition of carbon dioxide. *iPlasmaNano VIII* conference, University of Antwerp, Belgium, July 3–6, 2017. **(poster presentation)**

7. K. M. Bal and E. C. Neyts. Bridging time scales in atomistic simulations: from classical models to density functional theory. *VSC Users Day*, Royal Flemish Academy of Belgium for Science and the Arts, Brussels, Belgium, June 2, 2017. **(poster and oral flash presentation)**
8. K. M. Bal and E. C. Neyts. Collective variable-driven hyperdynamics: atomistic simulations of complex processes reaching long time scales. *Computational Molecular Science* conference, University of Warwick, UK, March 19–22, 2017. **(poster presentation)**
9. K. M. Bal and E. C. Neyts. Flexible accelerated molecular dynamics with a hybrid metadynamics/hyperdynamics method. *Bridging-Time Scale Techniques and their Applications in Atomistic Computational Science* focus workshop, Max Planck Institute for the Physics of Complex Systems in Dresden, Germany, September 12–15, 2016. **(oral presentation)**
10. K. M. Bal and E. C. Neyts. Flexible accelerated molecular dynamics with a hybrid metadynamics/hyperdynamics method. *AIP Publishing Horizons Future of Chemical Physics* conference, University of Oxford, UK, August 31–September 2, 2016. **(poster presentation)**
11. K. M. Bal and E. C. Neyts. The time scale associated with Monte Carlo simulations. *Spring School on Computational Tools for Materials Science*, Ghent University, Belgium, April 13–17, 2015. **(poster presentation)**

## Invited seminars

1. Atomistic simulations of large-scale chemical transformations with collective variable-based methods. *Institute of Computational Science*, ETH Zürich and Università della Svizzera italiana, Lugano, Switzerland, September 28, 2017.
2. Long time scale atomistic simulations of large reactive systems: from metadynamics to hyperdynamics (and back again). *Institute of Mineralogy, Materials Physics and Cosmochemistry*, Université Pierre et Marie Curie (Paris VI), Paris, France, June 19, 2017.

# Awards and grants

## Awards

1. **Best poster award** at the *23rd International Symposium on Plasma Chemistry*, Montreal, Canada, July 30–August 4, 2017. First prize out of ~200 posters. Awarded by IPCS/IOP, August 2017.
2. **JCP Editors' Choice** for the article “On the time scale associated with Monte Carlo simulations.” Awarded by the editors of *The Journal of Chemical Physics* in 2015.
3. **KVCV award** for the most meritorious Chemistry student graduating from the University of Antwerp. Awarded by the Royal Flemish Chemical Society (KVCV), September 2014.
4. **Oxford University Press Achievement in Chemistry Prize** for the best Bachelor student in Chemistry. Awarded by OUP, July 2010.

## Grants

1. **FWO travel grant** for a research visit to the ETH Zürich at Lugano, September–October 2017.
2. **FWO PhD fellowship** (2 + 2 years), October 2014–September 2018.



## APPENDIX A

# Timings & time scales: a time capsule

All calculations in this thesis were performed between 2015 and 2018 on the Hopper and Leibniz vsc clusters at the University of Antwerp, equipped with Intel processors of the Ivy Bridge and Broadwell generations, respectively. Various software tools and simulation methodologies were employed, each with their own computational cost. Given that the first part of this thesis boils down to the extension of the MD time scale, it is useful to re-stress the necessity of such an undertaking. Even for the simplest of systems, time scales of milliseconds, let alone seconds, are simply inaccessible.

While these kinds of statements can be found all over the literature (and this thesis), it is not always easy to find exact numbers: What is the accessible time scale for a given simulation approach? How large can a system be? What kind of hardware is required? For this reason, the hardware resources and timings of representative calculations as carried out in this thesis are summarized here. They again underline the need for long time scale simulation methods, but also serve as a time capsule of sorts. Indeed, the performance of computer hardware and algorithms is ever-improving, so it is useful to have a record of our contemporary computational abilities for future reference.

As the data in Table A.1 show, the “fastest” MD simulation needs almost a day to simulate a 1  $\mu$ s process, although it is for a very small system described by a simple short-range classical potential. A similarly cheap manybody potential like EAM can also not reach far beyond the  $\mu$ s time scale, and the problem only becomes worse for more sophisticated classical force fields like ReaxFF or the semi-empirical PM6 method. Evidently, *ab initio* methods are even more limited, with a throughput that is six orders of magnitude smaller than that

**TABLE A.1:** Time scales and computational parameters of the various MD simulation studies undertaken in this thesis. The simulation speed is expressed as throughput: MD time over wall time. Furthermore, used hardware resources and system sizes are included for fair comparison.

process	method	# atoms	# procs	throughput (ns/day)
Cu surface diffusion	EAM	300	1	$1.84 \times 10^2$
Chain folding	custom FF	50	1	$1.64 \times 10^3$
CH <sub>4</sub> /Ni	ReaxFF	384	4	0.15
<i>n</i> -C <sub>12</sub> H <sub>26</sub> pyrolysis	ReaxFF	912	8	0.79
CH <sub>3</sub> Cl + Cl <sup>-</sup>	PM6	6	1	1.61
CO <sub>2</sub> @ Ti/Al <sub>2</sub> O <sub>3</sub>	GGA-DFT	256	40	$1.42 \times 10^{-3}$

**TABLE A.2:** SCF times and computational parameters of the various DFT calculations carried out in this thesis. The time to complete a single SCF step is listed. Furthermore, used hardware resources, basis set size, and level of theory are included for fair comparison.

system	rung	basis	# procs	SCF time (s)
Cu/Al <sub>2</sub> O <sub>3</sub> neutral	GGA	DZVP	28	19.5
Cu/Al <sub>2</sub> O <sub>3</sub> charged	GGA	DZVP	56	23.3
Cu/Al <sub>2</sub> O <sub>3</sub> neutral	GGA	TZV2P	28	22.7
Cu/Al <sub>2</sub> O <sub>3</sub> neutral	hybrid	TZV2P	112	33.1
<i>h</i> -BN	GGA	DZVP	28	3.5
<i>h</i> -BN	GGA	TZV2P	28	3.2
<i>h</i> -BN	hybrid	TZV2P	56	6.5

of the simplest classical simulation in this study. Further parallelization will not improve matters. The systems are fairly small: the parallel efficiency of EAM evaluations already drops below 50% for 500 atoms/proc<sup>274</sup> and similar scalings have been reported for ReaxFF.<sup>84</sup> Similarly, DFT calculations with CP2K/QUICKSTEP do not scale well beyond ~5 atoms/proc.<sup>125</sup> In other words, these calculations are a fair representation of the intrinsic throughput of the computational techniques used in this thesis. Not only the evaluation speed of a potential determines the throughput, but also the allowed time step. This fact mostly affects ReaxFF, of which the rather rough PES requires time steps in the order of 0.1 fs, while all other methods allow at least 0.5 fs.

For completeness, our DFT-based structure optimizations are also discussed here. A large issue, especially of the calculations on  $\text{Al}_2\text{O}_3$  slabs, is memory: the large vacuum spaces in the countercharge-based calculations are very memory-consuming. For computer nodes with fairly small memory, this means that not the full number of processors can be used, as only large memory allocations ( $\sim 4$  GB/proc) are guaranteed to work smoothly. Table A.2 shows that GGA calculations can handle large systems quite well: a calculation on the 72-atom *h*-BN sheet is only about 3.5 times faster than on the 253-atom hydrated  $\text{Cu}/\text{Al}_2\text{O}_3$  slab, even though the latter represents a spin-unrestricted calculation (which should suffer from a  $\times 2$  performance penalty). In other words, the calculations on the small *h*-BN system exceed the linear strong scaling abilities of CP2K.

Countercharge-based calculation can be expected to double the computational overhead, while basis set size does not have a particularly large impact. However, calculations with many basis functions tend to be more difficult to converge, and might end up requiring much longer optimization runs. Hybrid calculations are significantly more costly, even with the many optimizations in CP2K: to the dense  $\text{Al}_2\text{O}_3$  slab a sixfold overhead is added, while it only increases hardware demands by about four times for the two-dimensional *h*-BN.



# Bibliography

- [1] A. F. Voter, F. Montalenti and T. C. Germann, *Annu. Rev. Mater. Res.* **32**, 321–346 (2002).
- [2] A. F. Voter, *Phys. Rev. B* **34**, 6819–6829 (1986).
- [3] G. Henkelman and H. Jónsson, *J. Chem. Phys.* **111**, 7010–7022 (1999).
- [4] N. Mousseau and G. T. Barkema, *Phys. Rev. E* **57**, 2419–2424 (1998).
- [5] G. Henkelman and H. Jónsson, *J. Chem. Phys.* **115**, 9657–9666 (2001).
- [6] H. Xu, Y. N. Osetsky and R. E. Stoller, *Phys. Rev. B* **84**, 132103 (2011).
- [7] F. El-Mellouhi, N. Mousseau and L. J. Lewis, *Phys. Rev. B* **78**, 153202 (2008).
- [8] L. K. Béland, Y. N. Osetsky, R. E. Stoller and H. Xu, *Comput. Mater. Sci.* **100**, 124–134 (2015).
- [9] M. Trochet, A. Sauv -Lacoursi re and N. Mousseau, *J. Chem. Phys.* **147**, 152712 (2017).
- [10] M. R. S rensen and A. F. Voter, *J. Chem. Phys.* **112**, 9599–9606 (2000).
- [11] G. Henkelman, B. P. Uberuaga and H. J nsson, *J. Chem. Phys.* **113**, 9901–9904 (2000).
- [12] G. H. Vineyard, *J. Phys. Chem. Solids* **3**, 121–127 (1957).
- [13] F. Montalenti, M. R. S rensen and A. F. Voter, *Phys. Rev. Lett.* **87**, 126101 (2001).
- [14] R. J. Zamora, B. P. Uberuaga, D. Perez and A. F. Voter, *Annu. Rev. Chem. Biomol. Eng.* **7**, 87–110 (2016).
- [15] Y. Shim, J. G. Amar, B. P. Uberuaga and A. F. Voter, *Phys. Rev. B* **76**, 205439 (2007).
- [16] V. Bochenkov, N. Suetin and S. Shankar, *J. Chem. Phys.* **141**, 094105 (2014).

## BIBLIOGRAPHY

- [17] A. F. Voter, *J. Chem. Phys.* **106**, 4665–4677 (1997).
- [18] A. F. Voter, *Phys. Rev. Lett.* **78**, 3908–3911 (1997).
- [19] M. M. Steiner, P.-A. Genilloud and J. W. Wilkins, *Phys. Rev. B* **57**, 10236–10239 (1998).
- [20] D. Hamelberg, J. Mongan and J. A. McCammon, *J. Chem. Phys.* **120**, 11919–11929 (2004).
- [21] D. Hamelberg, T. Shen and J. A. McCammon, *J. Chem. Phys.* **122**, 241103 (2005).
- [22] S. Pal and K. A. Fichthorn, *Chem. Eng. J.* **74**, 77–83 (1999).
- [23] R. A. Miron and K. A. Fichthorn, *J. Chem. Phys.* **119**, 6210–6216 (2003).
- [24] R. A. Miron and K. A. Fichthorn, *Phys. Rev. Lett.* **93**, 128301 (2004).
- [25] K. A. Fichthorn, R. A. Miron, Y. Wang and Y. Tiwary, *J. Phys.: Condens. Matter* **21**, 084212 (2009).
- [26] Y. Lin and K. A. Fichthorn, *Phys. Rev. B* **86**, 165303 (2012).
- [27] A. F. Voter, *Phys. Rev. B* **57**, R13985–R13988 (1998).
- [28] D. Perez, B. P. Uberuaga and A. F. Voter, *Comput. Mater. Sci.* **100**, 90–103 (2015).
- [29] D. Perez, E. D. Cubuk, A. Waterland, E. Kaxiras and A. F. Voter, *J. Chem. Theory. Comput.* **12**, 18–28 (2016).
- [30] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia and M. Parrinello, *Comput. Phys. Commun.* **180**, 1961–1972 (2009).
- [31] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni and G. Bussi, *Comput. Phys. Commun.* **185**, 604–613 (2014).
- [32] G. Fiorin, M. L. Klein and J. Hénin, *Mol. Phys.* **111**, 3345–3362 (2013).
- [33] U. H. Hansmann, *Chem. Phys. Lett.* **281**, 140–150 (1997).
- [34] Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.* **314**, 141–151 (1999).
- [35] L. Maragliano and E. Vanden-Eijnden, *Chem. Phys. Lett.* **426**, 168–175 (2006).

- [36] C. F. Abrams and E. Vanden-Eijnden, *Proc. Natl. Acad. Sci.* **107**, 4961–4966 (2010).
- [37] G. M. Torrie and J. P. Valleau, *J. Comput. Phys.* **23**, 187–199 (1977).
- [38] E. Darve and A. Pohorille, *J. Chem. Phys.* **115**, 9169–9183 (2001).
- [39] E. Darve, D. Rodríguez-Gómez and A. Pohorille, *J. Chem. Phys.* **128**, 144120 (2008).
- [40] J. Comer, J. C. Gumbart, J. Héning, T. Lelièvre, A. Pohorille and C. Chipot, *J. Phys. Chem. B* **119**, 1129–1151 (2015).
- [41] A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci.* **99**, 12562–12566 (2002).
- [42] A. Barducci, G. Bussi and M. Parrinello, *Phys. Rev. Lett.* **100**, 020603 (2008).
- [43] G. Bussi, A. Laio and M. Parrinello, *Phys. Rev. Lett.* **96**, 090601 (2006).
- [44] J. F. Dama, M. Parrinello and G. A. Voth, *Phys. Rev. Lett.* **112**, 240602 (2014).
- [45] A. Laio and F. L. Gervasio, *Rep. Prog. Phys.* **71**, 126601 (2008).
- [46] A. Barducci, M. Bonomi and M. Parrinello, *WIREs Comput. Mol. Sci.* **1**, 826–843 (2011).
- [47] S. Zheng and J. Pfendtner, *Mol. Simul.* **41**, 55–72 (2015).
- [48] O. Valsson, P. Tiwary and M. Parrinello, *Annu. Rev. Phys. Chem.* **67**, 159–184 (2016).
- [49] P. Tiwary and M. Parrinello, *Phys. Rev. Lett.* **111**, 230602 (2013).
- [50] M. Salvalaglio, P. Tiwary and M. Parrinello, *J. Chem. Theory Comput.* **10**, 1420–1425 (2014).
- [51] P. Tiwary, V. Limongelli, M. Salvalaglio and M. Parrinello, *Proc. Natl. Acad. Sci.* **112**, E386–E391 (2015).
- [52] K. L. Fleming, P. Tiwary and J. Pfendtner, *J. Phys. Chem. A* **120**, 299–305 (2016).
- [53] D. Branduardi, F. L. Gervasio and M. Parrinello, *J. Chem. Phys.* **126**, 054103 (2007).
- [54] P. Tiwary and B. J. Berne, *Proc. Natl. Acad. Sci.* **113**, 2839–2844 (2016).

## BIBLIOGRAPHY

- [55] J. McCarty and M. Parrinello, *J. Chem. Phys.* **147**, 204109 (2017).
- [56] P. Tiwary and B. J. Berne, *J. Chem. Phys.* **145**, 054113 (2016).
- [57] G. Piccini, D. Polino and M. Parrinello, *J. Phys. Chem. Lett.* **8**, 4197–4200 (2017).
- [58] C. Wehmeyer and F. Noé, *J. Chem. Phys.* **148**, 241703 (2018).
- [59] D. Mendels, G. Piccini and M. Parrinello, *J. Phys. Chem. Lett.* **9**, 2776–2781 (2018).
- [60] F. Pietrucci and W. Andreoni, *Phys. Rev. Lett.* **107**, 085504 (2011).
- [61] S. Zheng and J. Pfaendtner, *J. Phys. Chem. C* **118**, 10764–10770 (2014).
- [62] D. Polino and M. Parrinello, *J. Phys. Chem. A* **119**, 978–989 (2015).
- [63] P. Tiwary and A. van de Walle, *Phys. Rev. B* **87**, 094304 (2013).
- [64] J. Tersoff, *Phys. Rev. Lett.* **56**, 632–635 (1986).
- [65] S. Y. Kim, D. Perez and A. F. Voter, *J. Chem. Phys.* **139**, 144110 (2013).
- [66] W. K. Kim and M. L. Falk, *J. Chem. Phys.* **140**, 044107 (2014).
- [67] G. A. Tribello, M. Ceriotti and M. Parrinello, *Proc. Natl. Acad. Sci.* **107**, 17509–17514 (2010).
- [68] S. A. Paz and E. P. M. Leiva, *J. Chem. Theory Comput.* **11**, 1725–1734 (2015).
- [69] H. Hirai, *J. Chem. Phys.* **141**, 234109 (2014).
- [70] D. Perez, B. P. Uberuaga, Y. Shim, J. G. Amar and A. F. Voter, *Annu. Rep. Comput. Chem.* **5**, 79–98 (2009).
- [71] S. Plimpton, *J. Comput. Phys.* **117**, 1–19 (1995).
- [72] G. Bussi and M. Parrinello, *Phys. Rev. E* **75**, 056707 (2007).
- [73] G. Boisvert and L. J. Lewis, *Phys. Rev. B* **56**, 7643–7655 (1997).
- [74] M. S. Daw and M. I. Baskes, *Phys. Rev. B* **29**, 6443–6453 (1984).
- [75] S. M. Foiles, M. I. Baskes and M. S. Daw, *Phys. Rev. B* **33**, 7983–7991 (1986).
- [76] Z. Guo and D. Thirumalai, *J. Mol. Biol.* **263**, 323–343 (1996).

- [77] H. S. Bengaard, J. K. Nørskov, J. Sehested, B. S. Clausen, L. P. Nielsen, A. Molenbroek and J. R. Rostrup-Nielsen, *J. Catal.* **209**, 365–384 (2002).
- [78] J. E. Mueller, A. C. T. van Duin and W. A. Goddard, III, *J. Phys. Chem. C* **114**, 5675–5685 (2010).
- [79] B. Liu, M. T. Lusk and J. F. Ely, *Surf. Sci.* **606**, 615–623 (2012).
- [80] Y. Shibuta, R. Arifin, K. Shimamura, T. Oguri, F. Shimojo and S. Yamaguchi, *Chem. Phys. Lett.* **565**, 92–97 (2013).
- [81] W. Somers, A. Bogaerts, A. C. T. van Duin and E. C. Neyts, *J. Phys. Chem. C* **116**, 20958–20965 (2012).
- [82] W. Somers, A. Bogaerts, A. C. T. van Duin, S. Huygh, K. M. Bal and E. C. Neyts, *Catal. Today* **211**, 131–136 (2013).
- [83] A. C. T. van Duin, S. Dasgupta, F. Lorant and W. A. Goddard, III, *J. Phys. Chem. A* **105**, 9396–9409 (2001).
- [84] H. M. Aktulga, J. C. Fogarty, S. A. Pandit and A. Y. Grama, *Parallel Comput.* **38**, 245–259 (2012).
- [85] J. E. Mueller, A. C. T. van Duin and W. A. Goddard, III, *J. Phys. Chem. C* **114**, 4939–4949 (2010).
- [86] A. K. Rappé and W. A. Goddard, III, *J. Phys. Chem.* **95**, 3358–3363 (1991).
- [87] J. E. Mueller, A. C. T. van Duin and W. A. Goddard, III, *J. Phys. Chem. C* **113**, 20290–20306 (2009).
- [88] M. Yao, Z. Zheng and H. Liu, *Prog. Energy Combust. Sci.* **35**, 398–437 (2009).
- [89] R. D. Reitz and G. Duraisamy, *Prog. Energy Combust. Sci.* **46**, 12–71 (2015).
- [90] F. Battin-Leclerc, E. Blurock, R. Bounaceur, R. Fournet, P.-A. Glaude, O. Herbinet, B. Sirjean and V. Warth, *Chem. Soc. Rev.* **40**, 4762–4782 (2011).
- [91] C. K. Westbrook, W. J. Pitz, O. Herbinet, H. J. Curran and E. J. Silke, *Combust. Flame* **156**, 181–199 (2009).
- [92] K. Kohse-Höinghaus, P. Oßwald, T. Cool, T. Kasper, N. Hansen, F. Qi, C. Westbrook and P. Westmoreland, *Angew. Chem. Int. Ed.* **49**, 3572–3597 (2010).

## BIBLIOGRAPHY

- [93] L. S. Tran, B. Sirjean, P.-A. Glaude, R. Fournet and F. Battin-Leclerc, *Energy* **43**, 4–18 (2012).
- [94] K. Chenoweth, A. C. T. van Duin and W. A. Goddard, III, *J. Phys. Chem. A* **112**, 1040–1053 (2008).
- [95] K. Chenoweth, A. C. T. van Duin, S. Dasgupta and W. A. Goddard, III, *J. Phys. Chem. A* **113**, 1740–1746 (2009).
- [96] Q.-D. Wang, J.-B. Wang, J.-Q. Li, N.-X. Tan and X.-Y. Li, *Combust. Flame* **158**, 217–226 (2011).
- [97] J. Ding, L. Zhang, Y. Zhang and K.-L. Han, *J. Phys. Chem. A* **117**, 3266–3278 (2013).
- [98] T. Qi, C. W. Bauschlicher, J. W. Lawson, T. G. Desai and E. J. Reed, *J. Phys. Chem. A* **117**, 11115–11125 (2013).
- [99] C. Zou, S. Raman and A. C. T. van Duin, *J. Phys. Chem. B* **118**, 6302–6315 (2014).
- [100] E. Salmon, A. C. T. van Duin, F. Lorant, P.-M. Marquaire and W. A. Goddard, III, *Org. Geochem.* **40**, 416–427 (2009).
- [101] F. Castro-Marciano, A. M. Kamat, M. F. Russo, A. C. T. van Duin and J. P. Mathews, *Combust. Flame* **159**, 1272–1285 (2012).
- [102] M. Zheng, X. Li, J. Liu and L. Guo, *Energy Fuels* **27**, 2942–2951 (2013).
- [103] O. Kum, B. M. Dickson, S. J. Stuart, B. P. Uberuaga and A. F. Voter, *J. Chem. Phys.* **121**, 9808–9819 (2004).
- [104] K. L. Joshi, S. Raman and A. C. T. van Duin, *J. Phys. Chem. Lett.* **4**, 3792–3797 (2013).
- [105] T. Cheng, A. Jaramillo-Botero, W. A. Goddard, III and H. Sun, *J. Am. Chem. Soc.* **136**, 9434–9442 (2014).
- [106] G. J. Martyna, M. L. Klein and M. Tuckerman, *J. Chem. Phys.* **97**, 2635–2643 (1992).
- [107] G. J. Martyna, D. J. Tobias and M. L. Klein, *J. Chem. Phys.* **101**, 4177–4189 (1994).

- [108] M. E. Tuckerman, J. Alexandre, R. López-Rendón, A. L. Jochim and G. J. Martyna, *J. Phys. A: Math. Gen.* **39**, 5629 (2006).
- [109] M. P. B. Musculus, P. C. Miles and L. M. Pickett, *Prog. Energy Combust. Sci.* **39**, 246–283 (2013).
- [110] S. Banerjee, R. Tangko, D. A. Sheen, H. Wang and C. T. Bowman, *Combust. Flame* **163**, 12–30 (2016).
- [111] W. Tsang, *J. Phys. Chem. Ref. Data* **19**, 1–68 (1990).
- [112] D. U. B. Aussems, K. M. Bal, T. W. Morgan, M. C. M. van de Sanden and E. C. Neyts, *Chem. Sci.* **8**, 7160–7168 (2017).
- [113] D. U. B. Aussems, K. M. Bal, T. W. Morgan, M. C. M. van de Sanden and E. C. Neyts, *Carbon* **137**, 527–532 (2018).
- [114] E. C. Neyts and K. M. Bal, *Plasma Process. Polym.* **14**, 1600158 (2017).
- [115] C. Ashraf, S. Shabnam, A. Jain, Y. Xuan and A. C. T. van Duin, *Fuel* **235**, 194–207 (2019).
- [116] B. Arash, B. J. Thijsse, A. Pecenko and A. Simone, *Polym. Degrad. Stab.* **146**, 260–266 (2017).
- [117] C. D. Fu, L. F. L. Oliveira and J. Pfaendtner, *J. Chem. Theory Comput.* **13**, 968–973 (2017).
- [118] O. Valsson and M. Parrinello, *Phys. Rev. Lett.* **113**, 090601 (2014).
- [119] J. McCarty, O. Valsson and M. Parrinello, *J. Chem. Theory Comput.* **12**, 2162–2169 (2016).
- [120] P. Shaffer, O. Valsson and M. Parrinello, *Proc. Natl. Acad. Sci.* **113**, 1150–1155 (2016).
- [121] O. Valsson and M. Parrinello, *J. Chem. Theory Comput.* **11**, 1996–2002 (2015).
- [122] J. McCarty, O. Valsson, P. Tiwary and M. Parrinello, *Phys. Rev. Lett.* **115**, 070601 (2015).
- [123] G. Piccini, J. McCarty, O. Valsson and M. Parrinello, *J. Phys. Chem. Lett* **8**, 580–583 (2017).
- [124] J. J. P. Stewart, *J. Mol. Model.* **13**, 1173–1213 (2007).

## BIBLIOGRAPHY

- [125] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing and J. Hutter, *Comput. Phys. Commun.* **167**, 103–128 (2005).
- [126] J. Hutter, M. Iannuzzi, F. Schiffmann and J. VandeVondele, *WIREs Comput. Mol. Sci.* **4**, 15–25 (2014).
- [127] L. Genovese, T. Deutsch, A. Neelov, S. Goedecker and G. Beylkin, *J. Chem. Phys.* **125**, 074105 (2006).
- [128] G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.* **126**, 014101 (2007).
- [129] E. C. Neyts, K. K. Ostrikov, M. K. Sunkara and A. Bogaerts, *Chem. Rev.* **115**, 13408–13446 (2015).
- [130] F. Abild-Pedersen, J. Greeley, F. Studt, J. Rossmeisl, T. R. Munter, P. G. Moses, E. Skúlason, T. Bligaard and J. K. Nørskov, *Phys. Rev. Lett.* **99**, 016105 (2007).
- [131] A.-J. Zhang, A.-M. Zhu, J. Guo, Y. Xu and C. Shi, *Chem. Eng. J.* **156**, 601–606 (2010).
- [132] X. Tu and J. C. Whitehead, *Appl. Catal. B: Environ.* **125**, 439–448 (2012).
- [133] D. Mei, X. Zhu, C. Wu, B. Ashford, P. T. Williams and X. Tu, *Appl. Catal. B: Environ.* **182**, 525–532 (2016).
- [134] G. Chen, V. Georgieva, T. Godfroid, R. Snyders and M.-P. Delplancke-Ogletree, *Appl. Catal. B: Environ.* **190**, 115–124 (2016).
- [135] G. J. van Rooij, H. N. Akse, W. A. Bongers and M. C. M. van de Sanden, *Plasma Phys. Control. Fusion* **60**, 014019 (2018).
- [136] D. Cheng, F. R. Negreiros, E. Aprà and A. Fortunelli, *ChemSusChem* **6**, 944–965 (2013).
- [137] Y.-A. Zhu, D. Chen, X.-G. Zhou and W.-K. Yuan, *Catal. Today* **148**, 260–267 (2009).
- [138] L. C. Grabow and M. Mavrikakis, *ACS Catal.* **1**, 365–384 (2011).
- [139] Y. Yang, M. G. White and P. Liu, *J. Phys. Chem. C* **116**, 248–256 (2012).
- [140] C. Liu, T. R. Cundari and A. K. Wilson, *J. Phys. Chem. C* **116**, 5681–5688 (2012).
- [141] Z. Wang, X.-M. Cao, J. Zhu and P. Hu, *J. Catal.* **311**, 469–480 (2014).

- [142] L. Dietz, S. Piccinin and M. Maestri, *J. Phys. Chem. C* **119**, 4959–4966 (2015).
- [143] J. Ko, B.-K. Kim and J. W. Han, *J. Phys. Chem. C* **120**, 3438–3447 (2016).
- [144] L. Foppa, M.-C. Silaghi, K. Larmier and A. Comas-Vives, *J. Catal.* **343**, 196–207 (2016).
- [145] Y.-X. Pan, C.-J. Liu and Q. Ge, *Langmuir* **24**, 12410–12419 (2008).
- [146] D. C. Sorescu, W. A. Al-Saidi and K. D. Jordan, *J. Chem. Phys.* **135**, 124701 (2011).
- [147] S. Huygh, A. Bogaerts and E. C. Neyts, *J. Phys. Chem. C* **120**, 21659–21669 (2016).
- [148] D. Cornu, H. Guesmi, J.-M. Krafft and H. Lauron-Pernot, *J. Phys. Chem. C* **116**, 6645–6654 (2012).
- [149] J. Ye, C. Liu and Q. Ge, *J. Phys. Chem. C* **116**, 7817–7825 (2012).
- [150] K. R. Hahn, M. Iannuzzi, A. P. Seitsonen and J. Hutter, *J. Phys. Chem. C* **117**, 1701–1711 (2013).
- [151] O. I. Mayi, P. Thiyam, M. Bostrom and C. Persson, *Phys. Chem. Chem. Phys.* **17**, 20125–20133 (2015).
- [152] H. S. Moon, S. Kwon, S. H. Kwon, M. Cho, J. G. Seo and S. G. Lee, *RSC Adv.* **6**, 28607–28611 (2016).
- [153] A. K. Mishra, A. Roldan and N. H. de Leeuw, *J. Phys. Chem. C* **120**, 2198–2214 (2016).
- [154] Y. Zeng, X. Zhu, D. Mei, B. Ashford and X. Tu, *Catal. Today* **256**, 80–87 (2015).
- [155] Y.-X. Pan, C.-J. Liu, T. S. Wiltowski and Q. Ge, *Catal. Today* **147**, 68–76 (2009).
- [156] M.-C. Silaghi, A. Comas-Vives and C. Copéret, *ACS Catal.* **6**, 4501–4505 (2016).
- [157] R. Zhang, B. Wang, H. Liu and L. Ling, *J. Phys. Chem. C* **115**, 19811–19818 (2011).
- [158] J. Ye, C.-J. Liu, D. Mei and Q. Ge, *J. Catal.* **317**, 44–53 (2014).

## BIBLIOGRAPHY

- [159] J. A. Rodriguez, P. Liu, D. J. Stacchiola, S. D. Senanayake, M. G. White and J. G. Chen, *ACS Catal.* **5**, 6696–6706 (2015).
- [160] E. C. Neyts and A. Bogaerts, *J. Phys. D: Appl. Phys.* **47**, 224010 (2014).
- [161] S. Huygh and E. C. Neyts, *J. Phys. Chem. C* **119**, 4908–4921 (2015).
- [162] S. Huygh, A. Bogaerts, K. M. Bal and E. C. Neyts, *J. Phys. Chem. C* **122**, 9389–9396 (2018).
- [163] M. Shirazi, E. C. Neyts and A. Bogaerts, *Appl. Catal. B* **205**, 605–614 (2017).
- [164] M. Shirazi, A. Bogaerts and E. C. Neyts, *Phys. Chem. Chem. Phys.* **19**, 19150–19158 (2017).
- [165] R. L. Heinisch, F. X. Bronold and H. Fehske, *Phys. Rev. B* **85**, 075323 (2012).
- [166] P. F. Ambrico, M. Ambrico, L. Schiavulli, T. Ligonzo and V. Augelli, *Appl. Phys. Lett.* **94**, 051501 (2009).
- [167] P. F. Ambrico, M. Ambrico, A. Colaianni, L. Schiavulli, G. Dilecce and S. D. Benedictis, *J. Phys. D: Appl. Phys.* **43**, 325201 (2010).
- [168] F. J. J. Peeters, R. F. Rumphorst and M. C. M. van de Sanden, *Plasma Sources Sci. Technol.* **25**, 03LT03 (2016).
- [169] R. Tschiersch, S. Nemschokmichal, M. Bogaczyk and J. Meichsner, *J. Phys. D: Appl. Phys.* **50**, 105207 (2017).
- [170] J. Kim, M. S. Abbott, D. B. Go and J. C. Hicks, *ACS Energy Lett.* **1**, 94–99 (2016).
- [171] J. Rossmeisl, E. Skúlason, M. E. Björketun, V. Tripkovic and J. K. Nørskov, *Chem. Phys. Lett.* **466**, 68–71 (2008).
- [172] E. Skúlason, V. Tripkovic, M. E. Björketun, S. Gudmundsdóttir, G. Karlberg, J. Rossmeisl, T. Bligaard, H. Jónsson and J. K. Nørskov, *J. Phys. Chem. C* **114**, 18182–18197 (2010).
- [173] N. A. Deskins, R. Rousseau and M. Dupuis, *J. Phys. Chem. C* **114**, 5891–5897 (2010).
- [174] W.-J. Yin, B. Wen, S. Bandaru, M. Krack, M. Lau and L.-M. Liu, *Sci. Rep.* **6**, 23298 (2016).
- [175] X. Tan, H. A. Tahini and S. C. Smith, *Ener. Storage Mater.* **8**, 169–183 (2017).

- [176] Q. Sun, Z. Li, D. J. Searles, Y. Chen, G. Lu and A. Du, *J. Am. Chem. Soc.* **135**, 8246–8253 (2013).
- [177] X. Tan, L. Kou, H. A. Tahini and S. C. Smith, *Sci. Rep.* **5**, 17636 (2015).
- [178] X. Tan, L. Kou, H. A. Tahini and S. C. Smith, *ChemSusChem* **8**, 3626–3631 (2015).
- [179] X. Li, T. Guo, L. Zhu, C. Ling, Q. Xue and W. Xing, *Chem. Eng. J.* **338**, 92–98 (2018).
- [180] G. Qin, A. Du and Q. Sun, *Energy Technol.* **6**, 205–212 (2018).
- [181] G. Qin, Q. Cui, B. Yun, L. Sun, A. Du and Q. Sun, *Int. J. Hydrogen Energy* **43**, 9895–9901 (2018).
- [182] G. Qin, Q. Cui, W. Wang, P. Li, A. Du and Q. Sun, *ChemPhysChem* **19** (2018), in press.
- [183] B. Xiao, Y. Li, X. Yu and J. Cheng, *Sens. Actuators B: Chem.* **235**, 103–109 (2016).
- [184] Y. Jiao, Y. Zheng, S. C. Smith, A. Du and Z. Zhu, *ChemSusChem* **7**, 317–317 (2014).
- [185] X. Tan, H. A. Tahini and S. C. Smith, *ACS Appl. Mater. Interfaces* **8**, 32815–32822 (2016).
- [186] X. Tan, H. A. Tahini and S. C. Smith, *APL Mater.* **4**, 053202 (2016).
- [187] X. Li, L. Zhu, X. Chang, D. He, Q. Xue and W. Xing, *J. Mater. Chem. A* **6**, 12404–12410 (2018).
- [188] X. Tan, H. A. Tahini and S. C. Smith, *ACS Appl. Mater. Interfaces* **9**, 19825–19830 (2017).
- [189] G. Qin, A. Du and Q. Sun, *Phys. Chem. Chem. Phys.* **20**, 16216–16221 (2018).
- [190] X. Tan, L. Kou and S. C. Smith, *ChemSusChem* **8**, 2987–2993 (2015).
- [191] X. Tan, H. A. Tahini and S. C. Smith, *Chem. Phys.* **478**, 139–144 (2016).
- [192] J. Zhou, Q. Wang, Q. Sun, P. Jena and X. S. Chen, *Proc. Natl. Acad. Sci.* **107**, 2801–2806 (2010).
- [193] S. Lee, M. Lee and Y.-C. Chung, *Phys. Chem. Chem. Phys.* **15**, 3243–3248 (2013).

## BIBLIOGRAPHY

- [194] H. Guo, W. Zhang, N. Lu, Z. Zhuo, X. C. Zeng, X. Wu and J. Yang, *J. Phys. Chem. C* **119**, 6912–6917 (2015).
- [195] Q. Sun, G. Qin, Y. Ma, W. Wang, P. Li, A. Du and Z. Li, *Nanoscale* **9**, 19–24 (2017).
- [196] A. A. Koverga, S. Frank and M. T. Koper, *Electrochim. Acta* **101**, 244–253 (2013).
- [197] F. Che, J. T. Gray, S. Ha and J.-S. McEwen, *ACS Catal.* **7**, 551–562 (2017).
- [198] A. Kakekhani and S. Ismail-Beigi, *Phys. Chem. Chem. Phys.* **18**, 19676–19695 (2016).
- [199] G. Makov and M. C. Payne, *Phys. Rev. B* **51**, 4014–4022 (1995).
- [200] J. S. Hub, B. L. de Groot, H. Grubmüller and G. Groenhof, *J. Chem. Theory Comput.* **10**, 381–390 (2014).
- [201] H.-P. Komsa and A. Pasquarello, *Phys. Rev. Lett.* **110**, 095505 (2013).
- [202] H.-P. Komsa, N. Berseneva, A. V. Krashennnikov and R. M. Nieminen, *Phys. Rev. X* **4**, 031044 (2014).
- [203] D. Vinichenko, M. G. Sensoy, C. M. Friend and E. Kaxiras, *Phys. Rev. B* **95**, 235310 (2017).
- [204] G. Lippert, J. Hutter and M. Parrinello, *Mol. Phys.* **92**, 477–488 (1997).
- [205] S. Goedecker, M. Teter and J. Hutter, *Phys. Rev. B* **54**, 1703–1710 (1996).
- [206] M. Krack, *Theor. Chem. Acc.* **114**, 145–152 (2005).
- [207] J. VandeVondele and J. Hutter, *J. Chem. Phys.* **127**, 114105 (2007).
- [208] J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
- [209] S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.* **132**, 154104 (2010).
- [210] S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.* **32**, 1456–1465 (2011).
- [211] P. Bultinck, C. Van Alsenoy, P. W. Ayers and R. Carbó-Dorca, *J. Chem. Phys.* **126**, 144111 (2007).
- [212] G. J. Martyna and M. E. Tuckerman, *J. Chem. Phys.* **110**, 2810–2821 (1999).

- [213] X. Tan, H. A. Tahini, P. Seal and S. C. Smith, *ACS Appl. Mater. Interfaces* **8**, 10897–10903 (2016).
- [214] X.-F. Yang, A. Wang, B. Qiao, J. Li, J. Liu and T. Zhang, *Acc. Chem. Res.* **46**, 1740–1748 (2013).
- [215] J. Liu, *ACS Catal.* **7**, 34–59 (2017).
- [216] M. Digne, P. Sautet, P. Raybaud, P. Euzen and H. Toulhoat, *J. Catal.* **226**, 54–68 (2004).
- [217] R. Wischert, P. Laurent, C. Copéret, F. Delbecq and P. Sautet, *J. Am. Chem. Soc.* **134**, 14430–14449 (2012).
- [218] J. A. Bittencourt. *Fundamentals of Plasma Physics* (Springer, New York, 2004), 3 edn.
- [219] S. F. J. Hackett, R. M. Brydson, M. H. Gass, I. Harvey, A. D. Newman, K. Wilson and A. Lee, *Angew. Chem. Int. Ed.* **46**, 8593–8596 (2007).
- [220] T. K. Ghosh and N. N. Nair, *ChemCatChem* **5**, 1811–1821 (2013).
- [221] E. J. Peterson, A. T. DeLaRiva, S. Lin, R. S. Johnson, H. Guo, J. T. Miller, J. H. Kwak, C. H. F. Peden, B. Kiefer, L. F. Allard, F. H. Ribeiro and A. K. Datye, *Nat. Commun.* **5**, 4885 (2014).
- [222] R. Sabatini, T. Gorni and S. de Gironcoli, *Phys. Rev. B* **87**, 041108 (2013).
- [223] O. A. Vydrov and T. Van Voorhis, *J. Chem. Phys.* **133**, 244103 (2010).
- [224] E. D. Murray, K. Lee and D. C. Langreth, *J. Chem. Theory Comput.* **5**, 2754–2762 (2009).
- [225] W. Hujo and S. Grimme, *J. Chem. Theory Comput.* **7**, 3866–3871 (2011).
- [226] G. Miceli, S. de Gironcoli and A. Pasquarello, *J. Chem. Phys.* **142**, 034501 (2015).
- [227] T. Björkman, *Phys. Rev. B* **86**, 165109 (2012).
- [228] H. Peng, Z.-H. Yang, J. P. Perdew and J. Sun, *Phys. Rev. X* **6**, 041005 (2016).
- [229] K. Patkowski, G. Murdachaew, C.-M. Fou and K. Szalewicz, *Mol. Phys.* **103**, 2031–2045 (2005).
- [230] Y. Zhang and W. Yang, *Phys. Rev. Lett.* **80**, 890–890 (1998).

## BIBLIOGRAPHY

- [231] L. Goerigk and S. Grimme, *Phys. Chem. Chem. Phys.* **13**, 6670–6688 (2011).
- [232] B. Hammer, L. B. Hansen and J. K. Nørskov, *Phys. Rev. B* **59**, 7413–7421 (1999).
- [233] J. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, *Phys. Rev. Lett.* **91**, 146401 (2003).
- [234] J. P. Perdew and K. Schmidt. Jacob’s ladder of density functional approximations for the exchange–correlation energy. In V. Van Doren, C. Van Alsenoy and P. Geerlings (eds.) *Density functional theory and its application to materials*, 1–20 (American Institute of Physics, Melville, New York, 2001).
- [235] D. C. Sorescu, J. Lee, W. A. Al-Saidi and K. D. Jordan, *J. Chem. Phys.* **134**, 104707 (2011).
- [236] J. P. Perdew, A. Ruzsinszky, J. Tao, V. N. Staroverov, G. E. Scuseria and G. I. Csonka, *J. Chem. Phys.* **123**, 062201 (2005).
- [237] K. Burke, *J. Chem. Phys.* **136**, 150901 (2012).
- [238] A. J. Cohen, P. Mori-Sánchez and W. Yang, *Chem. Rev.* **112**, 289–320 (2012).
- [239] A. D. Becke, *J. Chem. Phys.* **140**, 18A301 (2014).
- [240] R. O. Jones, *Rev. Mod. Phys.* **87**, 897–923 (2015).
- [241] H. S. Yu, S. L. Li and D. G. Truhlar, *J. Chem. Phys.* **145**, 130901 (2016).
- [242] L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi and S. Grimme, *Phys. Chem. Chem. Phys.* **19**, 32184–32215 (2017).
- [243] J. Sun, R. C. Remsing, Y. Zhang, Z. Sun, A. Ruzsinszky, H. Peng, Z. Yang, A. Paul, U. Waghmare, X. Wu, M. L. Klein and J. P. Perdew, *Nat. Chem.* **8**, 831–836 (2016).
- [244] F. Tran, J. Stelzl and P. Blaha, *J. Chem. Phys.* **144**, 204120 (2016).
- [245] S. Grimme, A. Hansen, J. G. Brandenburg and C. Bannwarth, *Chem. Rev.* **116**, 5105–5154 (2016).
- [246] M. A. Marques, M. J. Oliveira and T. Burnus, *Comput. Phys. Commun.* **183**, 2272–2281 (2012).
- [247] M. Guidon, J. Hutter and J. VandeVondele, *J. Chem. Theory Comput.* **6**, 2348–2364 (2010).

- [248] Q. Sun, Y. Dai, Y. Ma, W. Wei and B. Huang, *RSC Adv.* **5**, 33037–33043 (2015).
- [249] A. D. Becke, *Phys. Rev. A* **38**, 3098–3100 (1988).
- [250] C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B* **37**, 785–789 (1988).
- [251] A. D. Boese and N. C. Handy, *J. Chem. Phys.* **114**, 5497–5503 (2001).
- [252] R. Armiento and A. E. Mattsson, *Phys. Rev. B* **72**, 085108 (2005).
- [253] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou and K. Burke, *Phys. Rev. Lett.* **100**, 136406 (2008).
- [254] Y. Zhao and D. G. Truhlar, *J. Chem. Phys.* **128**, 184109 (2008).
- [255] H. S. Yu, W. Zhang, P. Verma, X. He and D. G. Truhlar, *Phys. Chem. Chem. Phys.* **17**, 12146–12160 (2015).
- [256] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, L. A. Constantin and J. Sun, *Phys. Rev. Lett.* **103**, 026403 (2009).
- [257] R. Peverati and D. G. Truhlar, *J. Chem. Theory Comput.* **8**, 2310–2319 (2012).
- [258] J. Wellendorff, K. T. Lundgaard, A. Møgelhøj, V. Petzold, D. D. Landis, J. K. Nørskov, T. Bligaard and K. W. Jacobsen, *Phys. Rev. B* **85**, 235149 (2012).
- [259] R. Peverati and D. G. Truhlar, *Phil. Trans. R. Soc. A* **372**, 20120476 (2014).
- [260] J. Sun, A. Ruzsinszky and J. P. Perdew, *Phys. Rev. Lett.* **115**, 036402 (2015).
- [261] M. Ernzerhof and G. E. Scuseria, *J. Chem. Phys.* **110**, 5029–5036 (1999).
- [262] C. Adamo and V. Barone, *J. Chem. Phys.* **110**, 6158–6170 (1999).
- [263] J. Heyd, G. E. Scuseria and M. Ernzerhof, *J. Chem. Phys.* **118**, 8207–8215 (2003).
- [264] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov and G. E. Scuseria, *J. Chem. Phys.* **125**, 224106 (2006).
- [265] J. P. Perdew, M. Ernzerhof and K. Burke, *J. Chem. Phys.* **105**, 9982–9985 (1996).
- [266] M. Guidon, J. Hutter and J. VandeVondele, *J. Chem. Theory Comput.* **5**, 3010–3021 (2009).
- [267] V. I. Anisimov, J. Zaanen and O. K. Andersen, *Phys. Rev. B* **44**, 943–954 (1991).

## BIBLIOGRAPHY

- [268] S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys and A. P. Sutton, *Phys. Rev. B* **57**, 1505–1509 (1998).
- [269] P. Verma and D. G. Truhlar, *J. Phys. Chem. C* **121**, 7144–7154 (2017).
- [270] P. Verma and D. G. Truhlar, *J. Phys. Chem. Lett.* **8**, 380–387 (2017).
- [271] A. J. Garza and G. E. Scuseria, *J. Phys. Chem. Lett.* **7**, 4165–4170 (2016).
- [272] K. Watanabe, T. Taniguchi and H. Kanda, *Nat. Mater.* **3**, 404–409 (2004).
- [273] K. Lee, E. D. Murray, L. Kong, B. I. Lundqvist and D. C. Langreth, *Phys. Rev. B* **82**, 081101 (2010).
- [274] S. J. Plimpton and A. P. Thompson, *MRS Bull.* **37**, 513–521 (2012).

#### A NOTE ABOUT THE TYPE

The main body text font is Equity and the default sans serif typeface as used in headings and captions is Concourse. Textual labels in graphs and diagrams are set in `mVB Solitaire`, and chapter headings in Elena.

For mathematical typesetting, additional characters were taken from Minion Pro (Greek letters) and MnSymbol (mathematical symbols) to supplement Equity.

This thesis was prepared with XeLaTeX in the KOMA-script document class.

