Memory and fluctuations in chemical dynamics

by

Ardavan Farahvash

B.S. Chemical Physics
University of California, Davis

Submitted to the Department of Chemistry
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN CHEMISTRY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2024

© 2024 Ardavan Farahvash. This work is licensed under a CC BY-NC-ND 4.0 license.

The author hereby grants to MIT a nonexclusive, worldwide, irrevocable, royalty-free license to exercise any and all rights under copyright, including to reproduce, preserve, distribute and publicly display copies of the thesis, or release the thesis under an open-access license.

Authored by:	Ardavan Farahvash Department of Chemistry April 11, 2024
Certified by:	Adam P. Willard Professor of Chemistry, Thesis Supervisor
Accepted by:	Adam P. Willard Department of Chemistry Graduate Officer Professor of Chemistry

This doctoral thesis has been examined by a committee of the Department of Chemistry as follows:

Professor Jianshu Cao: .....
Thesis Committee Chair
Professor of Chemistry

Professor Adam P. Willard: .....
Thesis Supervisor
Professor of Chemistry

Professor Heather J. Kulik: .....
Thesis Committee Member
Associate Professor of Chemical Engineering

Memory and fluctuations in chemical dynamics

by

Ardavan Farahvash

Submitted to the Department of Chemistry
on April 11, 2024 in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN CHEMISTRY

ABSTRACT

This thesis describes the development and application of theories that elucidate both the static and time-dependent responses of various condensed-phase environments to molecular systems.

Part I, the cornerstone of this thesis, explores the role of surface vibrations in gas-phase heterogeneous catalysis. Utilizing the Mori-Zwanzig projection operator formalism, I have developed a theory that maps surface vibrations to a generalized Langevin equation (GLE). Two projection schemes are considered. The first scheme projects the motion of the entire solid substrate onto the motion of molecular adsorbates. The second scheme projects onto both the motion of adsorbates and of surface adsorption sites. Through the first approach, I demonstrate that physisorbed species primarily couple with acoustic phonons, while chemisorbed species couple with dispersionless local vibrations. I also use this scheme to examine how phonons affect reactions rates, both in ensembles near and far-from thermal equilibrium. Using the second approach, I study how energy is dissipated in simulations of molecule-surface scattering. I demonstrate that phonon confinement effects from nanoscale simulations can significantly impact calculated surface sticking coefficients.

Part II considers the role of solvent in adsorption and desorption at liquid-solid interfaces. Specifically, I employ enhanced sampling methods to study a model system of carbon monoxide at a water/platinum interface. Using these methods, I show that the local coordination number around a CO molecule plays a crucial role in the transition states of the adsorption/desorption process, and that CO tends to increase its coordination number before desorbing.

Part III develops a machine learning and electronic structure framework for the computationally efficient parametrization of Frenkel Hamiltonians from snapshots of molecular dynamics simulations of organic semiconductors. Direct electronic structure calculations on these snapshots encode the nuclear fluctuations of the chromophores in the material and how they couple to excitons, but at enormous cost. I discuss how the strategic application of machine learning methods can drastically reduce the number of electronic structure calculations needed to produce a complete exciton trajectory. Critically, I demonstrate that by decomposing the two-molecule excitonic coupling into interactions between one-molecule transition monopoles, a more accurate and less data-intensive machine learning scheme can be devised.

Thesis supervisor: Adam P. Willard
Title: Professor of Chemistry

Acknowledgments

Saffron is a spice used throughout Iranian cooking, made from a delicate flower that requires the perfect humidity and detail-oriented care. Scientific research is similar to saffron cultivation in that it too is a highly rewarding endeavor, but one that requires ideal conditions made possible by the support of others. The individuals I mention herein are those who created the environment that made it possible for me to complete my doctorate. They are the cultivators that provided the nourishment necessary for my scientific endeavors to flourish.

I cannot thank my research advisor, Professor Adam Willard, enough for all he has done for me. From the moment I met him during my first semester of 5.70 (Statistical Thermodynamics) and throughout my time as a graduate student, Adam has been an invaluable source of guidance and a bedrock of support. He is an incredibly kind and generous advisor, and I am beyond grateful for his past and continuing support.

My direct collaborators were a cornerstone of my PhD work. I am especially grateful to Professors Andrew (Andy) Peterson and Keith Nelson. Both were invaluable in providing ideas, encouragement, and in critiquing my results. Andy is a true master of computational catalysis whose guidance allowed me to transition from my work on exciton dynamics early in my PhD, to my work in catalysis later on. Keith is an unbelievably deep-thinking physical chemist, who largely molded my thoughts and experiences on when and how to draw connections between theory and experiment. I would also like to thank Professor Liang Shi for teaching me about exciton dynamics. I thank Tristan Shi for being an incredible experimental collaborator; you taught me how rewarding it can be to work in close collaboration with experimentalists. I thank Mayank Agarwal for your crucial support on my projects studying the dynamics of adsorbates on surfaces. I would also like to acknowledge Chee-Kong Lee, Weiwei Sun, and Professor Moungi Bawendi.

I am grateful to Professors Jianshu Cao and Heather Kulik for being my thesis committee members. Thank you for your feedback and support.

I could not have asked for better colleagues that my fellow members of the Willard group: Maria Castellanos, Henrik Heelweg, Caelan Brooks, Yoon Park, Dina Sharon, Linsey Nowak, Julia Nauman, Dylan Suvulu, Amr Dodin, Michiel Niessen, Yizhi Shen, Kaitlyn Dwelle, Kyaw Myint, Hamish Hiscock, Professor Wendu Ding, and Aditya Limaye. Many of the best ideas I had during my PhD were inspired by questions and informal conversations at group and subgroup meetings.

I am extremely grateful for the financial support and opportunities afforded to me by the National Science Foundation Graduate Research Fellowship. I would not have been able to pursue many of my more inventive ideas without it.

My time at MIT and my conviction to pursue science as a career would not have happened without the terrific mentors I had as an undergraduate: Professors Alexei Stuchebrukhov and Bill McCurdy. Thank you for believing in me, inspiring me, and indulging my insufferable curiosity.

I must give thanks to the many great friends I made in Boston/Cambridge. Thank you to Daniel, Jet, Amogh, Cliff, Natalie, Jinyi, Mikhaila, Bryan, Yosh, Dalit, and

Dharashri. I will always cherish the memories we made together. And a special thank is given to my best friend Maria, who showed me that true love does exist.

Finally, and most importantly, I thank my parents, to whom this thesis is dedicated. My parents sacrificed so much and worked so incredibly hard to immigrate to the United States and raise me here. From my childhood through graduate school, they have given me their unconditional love and support. They instilled in me the value of discipline, humility, creativity, and excellence. Words cannot truly express my appreciation.

Dedication

To my mother, Yasaman Ameri, and my father, Shayan Farahash, whose love and sacrifice made it possible for me to pursue my dreams.

Contents

Title page **1**
Committee Signatures 2

Abstract **3**

Acknowledgments **5**
Dedication 7

List of Figures **11**

List of Tables **16**

I Phonon-induced memory in surface chemistry **17**

1 Introduction to heterogeneous catalysis **19**
1.1 The Sabatier principle 19
1.2 Surface vibrations in heterogeneous catalysis 21

2 Theoretical preliminaries **25**
2.1 Mori-Zwanzig theory 25
2.2 Properties of memory kernels 30
2.3 Reaction rate theory 33

3 Phonon-induced memory on adsorbates **38**
3.1 Introduction 38
3.2 Memory for physisorbed and chemisorbed species 40
3.3 Perturbation Theory 44
3.3.1 Strong coupling/chemisorbed limit 46
3.3.2 Weak-coupling/physisorbed limit 48
3.4 Dispersion 49
3.5 Anharmonic effects 52
3.6 Ramifications for reaction rates 55
3.6.1 Surface desorption 57
3.6.2 Comparison to experiment 59
3.7 Phonon response functions 61

3.8 Conclusions . . . 64

4 Phonon-induced memory on surface sites 65
4.1 Introduction . . . 65
4.2 Methods . . . 67
4.2.1 Extended variable transformation . . . 67
4.2.2 Simulation details . . . 69
4.3 Memory kernels for surface sites in metal lattices . . . 70
4.3.1 Finite-size effects . . . 73
4.3.2 Solvation effects . . . 74
4.4 How memory affects scattering . . . 75
4.5 Conclusions . . . 78

II Solvent fluctuations at surfaces 79

5 The influence of solvent on surface adsorption and desorption 80
5.1 Introduction . . . 80
5.2 Transition path sampling . . . 82
5.3 Simulation methods . . . 84
5.4 Results and discussion . . . 86
5.4.1 TPS . . . 86
5.4.2 Free-energy surface . . . 90
5.5 Conclusions . . . 92

III Exciton dynamics 93

6 Machine learning Frenkel Hamiltonians 94
6.1 Introduction . . . 94
6.2 The Frenkel-Davydov Model . . . 98
6.3 Computational Details . . . 100
6.3.1 Data Generation . . . 100
6.3.2 Data Preparation and Machine Learning . . . 101
6.4 ML Model For Excitation Energies . . . 102
6.5 Direct ML Model For Couplings . . . 106
6.6 Indirect ML Model For Couplings via Transition Densities . . . 108
6.6.1 Coupling Calculation Benchmarking . . . 108
6.6.2 Machine Learning Model . . . 112
6.6.3 Trajectory Analysis . . . 114
6.7 Conclusions . . . 115

IV Conclusions 119

7 Summary and outlook 120

A Non-linear Mori-Zwanzig theory **123**

B Calculating memory kernels from time-correlation functions **128**

C Elastic continuum theory **131**

D Derivation of extended variable transformation **136**
D.1 Ornstein-Uhlenbeck processes 136
D.2 Mapping to a multivariate OU bath 139

E Continuum limit of 1D harmonic chain **141**

F Derivation of Frenkel-Davydov model **143**
F.1 Diagonal Elements 144
F.2 Off-Diagonal Elements 146

G 1-electron density matrices and transition density matrices **148**

Bibliography **151**

List of Figures

1.1 Schematic of Sabatier principle. 20
1.2 Schematic of models developed in Chapter 3 and Chapter 4. The equations shown illustrate how projections map the deterministic Newtonian motion to stochastic, non-Markovian motion. 24

2.1 Bare potential energy (U) and potential of mean force (W) for a Morse oscillator system coupled to a single harmonic mode bath. The bath mode widens the width of the bare Morse potential, such that the frequency at the minimum shifts from $\omega_{\mathrm{A}} = 500$ cm^{-1} to $\omega_{\mathrm{eff}} = 167$ cm^{-1}. 31
2.2 Depiction of characteristic forms of the (A) memory kernel and (B) spectral density. The overdamped form is common in systems coupled to liquid environments, while the underdamped form common in systems coupled to solid environments. 33
2.3 Characteristic potential energy curve for barrier-crossing problem. . . 34
2.4 Grote-Hynes rate constants versus numerical simulation results for a 1D barrier crossing problem with exponential memory, $K(t) = Ae^{-\frac{t}{\tau}}$. 35

3.1 2D schematic illustrating the dominant phonon modes in terms of their coupling to the adsorbate or contribution to the memory kernel. ω_{as} is the frequency of the adsorbate-surface bond and the ω_{D} is the solid's Debye frequency. 39
3.2 (A) Memory kernel, (B) spectral density, and (C) density of states for Pt(111) and three values of ω_{as}. The density of states in (C) was calculated using the 4x4x8 surface slab. 42
3.3 (A) Memory kernel and (B) spectral density for Ru(0001) surface. . . 44
3.4 Spectral density for Pt(111) surface calculated using exact diagonalization, 1st order, and 2nd order perturbation theory in the strong-coupling scheme. 48
3.5 Spectral density for Pt(111) surface calculated using exact diagonalization, 1st order, and 2nd order perturbation theory in the weak-coupling scheme. 49

3.6 Phonon dispersion curves. (A) Bulk Pt dispersion curves calculated using an EMT forcefield and a 10x10x10 atom supercell. (B) Dispersion curves for a 4x4x8 atom surface slab replicated in a 6x6 surface cell (C) Same as middle but with an adsorbed CO molecule corresponding to $\omega_{as} = 480$ cm^{-1}. 51
3.7 Memory kernel for a model adsorbate on Pt(111) in a finite system (black) and in the infinite system size limit (blue). The finite system is equivalent to the 4x4x8 results presented in Figure 1. 52
3.8 (A) Memory kernel and (B) spectral density for a 4x4x4 Pt surface calculated via Eqs. 3.1 and 3.2 versus with time-correlation functions calculated with MD simulations (Eq. B.1). 53
3.9 Dependence of (A) memory kernel and (B) spectral density on hopping rate with $\omega_{as} = 480$ cm^{-1}. 54
3.10 Rate constants for desorption from a Pt(111) surface. (A) CO desorption. Grey squares are experimental data which mixed contributions from both steps and terraces. Black circles refer experimental data where the kinetics of terrace desorption was isolated. (B) Xe desorption. 60
3.11 Schematics for (A) extensional forces, (B) shear forces, and (C) shockwave forces used for computing χ. (D) Time-domain response functions for extensional, shear, and shockwave forces. (F) Imaginary part of frequency-domain response function. Response functions were calculated using a 25x25x10 Pt film with an EMT forcefield[71]. 63

4.1 Memory kernel and random force power spectrum for surface sites of a Pt(111) lattice computed using an EMT forcefield. (A) Memory kernel for fluctuations in x/y (in-surface plane) and z (out of plane) directions. (B) z component of the memory kernel. Red and blue lines are two exponential sinusoids optimized to fit the computed memory kernel (grey line). (C) Power spectrum of z component of the memory kernel. The grey dashed vertical line corresponds to the experimental Debye frequency. (D) Depiction of lattice normal modes most associated with red and blue lines. τ is the period of the respective normal mode. The yellow-highlighted sphere in the middle of the lattice represents the chosen adsorption site, which is fixed in space, while the surrounding grey atoms represent the fluctuating bulk lattice. 71
4.2 (A) Memory kernel and (B) spectral densities for surface site fluctuations of Pt(111) simulated using three different atomistic models: Effective Medium Theory, Embedded-Atom Method, and a Lennard-Jones model. The acoustic peak at 20 cm^{-1} of the spectral density has been scaled by a factor of 0.3 for the purposes of visual clarity. 72
4.3 Memory kernel and spectral densities for surface site fluctuations of Pt(111) lattices of different sizes. The plots on the right zoom-in on the regions near to acoustic and Debye peaks respectively 74

4.4 (A) Memory kernel and (B) spectral densities for surface site fluctuations of Pt(111) simulated using a LJ model with and without SPC/E solvent. 75
4.5 Sticking probabilities S as a function of the ratio of the incident kinetic energy to the well depth $\mathrm{KE}(t=0)/D$. (A) Results for Morse PES with $D = 6.62$eV. (B) Results for Morse PES with an increased well-depth, $D = 30.62$eV. (C) Power spectra for a 4x4x4 Pt(111) lattice calculated using the CF method overlayed with the fits used to calculate results in (A) and (B). 77

5.1 Snapshot of simulation system consisting of a three-layer Pt(100) surface with CO molecule adsorbed atop a central site and 57 water molecules in 1650 $\mathrm{Å}^3$ volume above surface to achieve a water density of roughly 1 $\mathrm{gm/cm}^3$. 84
5.2 (A) Committor as a function of distance to surface. The black dashed line is a windowed average over the bar heights with a window of 5. The red vertical dashed line corresponds to the configuration closest to $p_{\mathrm{D}} = 0.5$. (B) Committor as a function of water coordination number. (C) Committor as a function of CO orientation. (D) Committor histogram (grey bars) and beta distribution model (black dashed line) for configurations near $z_{\mathrm{C}}^{\ddagger}$. (E) Committor histogram and beta distribution model for configurations near $\eta_{\mathrm{W}}^{\ddagger}$. (F) Committor histogram and beta distribution model for configurations near $\phi_{\mathrm{CO}}^{\ddagger}$. 87
5.3 (A) 2D committor probability as a function of the distance to surface and the water coordination number. The black vertical line corresponds to $z_{\mathrm{C}} = z_{\mathrm{C}}^{\ddagger}$, the brown horizontal line corresponds to $\eta_{\mathrm{W}} = \eta_{\mathrm{W}}^{\ddagger}$, and the red dashed line to states along $\eta_{\mathrm{W}} = 1.92 z_{\mathrm{C}} + 17.18$. (B) Committor histogram and beta distribution model taken for states near the red dashed line in 3A. 88
5.4 (A) 2D FES along z_{C} and η_{W} calculated using metadynamics. The black contour corresponds to the path of minimum local free-energy stage (steepest descent) between reactants and products. A minor windowed average smoothing, with a window of 3 was applied to this contour to remove noise. The green contour represents a path of constant η_{W}. The red dashed line represents the transition state surface shown in Figure 2A found using TPS. (B) 1D cross-section of free-energy as a function of z_{C} along the χ and $\eta_{\mathrm{W}} = 6.2$ contours, as well as the gas phase desorption free-energy surface. (C) 1D cross-section of free-energy as a function of η_{W} along the χ contour. The red vertical lines show the $p_{\mathrm{D}} = 0.5$ states found with TPS. 90

6.1 Depiction of machine learning schemes described by this work. Three machine learning models are tested, one for the excitation energy and two for the excitonic couplings. The deep neural network (DNN) model for direct prediction of couplings is shown to perform worse than the transition charge based kernel ridge regression (KRR) model. 96
6.2 Snapshots of different molecular morphologies used in this study. Panels (A) and (B) show the monolayer and amorphous morphologies, respectively. 101
6.3 (A) Differential performance of using several different charge models with the Coulomb Matrix. Performances reported here are using the 10,000 example monolayer PPP/CIS training set. atom-dist = pairwise atomic distances (no charges), atm# = atomic numbers, uff = Universal Forcefield [232], gastg = Gasteiger charges[231], opls = OPLS/2005 Forcefield [206]. (B) Mean absolute error in E_m predictions versus training set size used. (C) 2D principal component analysis of pairwise atomic distances and coulomb matrices taken from T6 conformations. (D) Cross-morphology error analysis on predictions of E_m from TDDFT. Dashed lines indicate fluctuations (standard deviation) in E_m from MD bath. 104
6.4 V_{mn} prediction performance using direct model. (A) Mean absolute error in PPP/CIS V_{mn} predictions versus the number of one-molecule PPP/CIS excited-state calculations used to generate the training set. (B) Cross-morphology error analysis on predictions of V_{mn} using 10,000 ES calculations = 260,000 coupling training sets. 107
6.5 Parity plots for benchmarking coupling calculations. All values are from monolayer simulation. (A) Exact PPP/CIS versus TDDFT couplings. (B) TDDFT couplings with and without exchange. (C) TDDFT Coulomb couplings from Mulliken transition charges. (D) TDDFT Coulomb couplings from Löwdin transition charges. 109
6.6 Charges and density visualizations for a typical molecule taken from monolayer simulation. (A) Isosurface of transition density using a value of 0.07, violet lobes are positive, red lobes are negative. (B) Mulliken charges, the size of each sphere indicates the magnitude of the assigned charge. (C) Löwdin charges. 109
6.7 Parity plot comparing exact Coulomb couplings to couplings from using a single, average value of the Löwdin transition charges. The MAE here is 1.74 meV 111
6.8 (A) Parity plot comparing exact Coulomb couplings to transition charge coupling from natural population analysis. All values are from monolayer simulation. (B) Parity plot comparing NAO and Lowdin transition charge couplings. 111

6.9 Comparing the error between actual transition charge couplings and machine-learned transition charge couplings. (A) Mean absolute error in V_{mn} predictions versus training set size used. (B) Cross-morphology error analysis with TDDFT data. (C) Average coupling versus center of geometry distance, solid lines are from exact transition charges, dashed lines are from ML predicted charges. (D) Couplings distributions. . . 113
6.10 Parity plots for comparing accuracy of ML predictions to (A) the exact values, and (B) the values computed with TDDFT Coulomb couplings from Löwdin transition charges. All values are from monolayer simulation. ML model was trained using 10,000 molecular conformations. 114
6.11 PPP/CIS trajectory statistics from monolayer and amorphous simulations. Solid lines are using exact calculations dashed lines are using ML model trained with 10,000 training examples selected only from the monolayer simulation. (A) Mean square displacement. (B) Lowest energy eigenvalue. 116

B.1 Memory kernel calculated with by using midpoint quadrature, trapezoidal quadrature, fast Fourier transform methods to solve the Volterra equation versus the input (target) memory kernel used to conduct the simulation. The trapezoidal quadrature used the derivative of Eq. B.1. 129

C.1 Square of the effective (phonon-renormalized) surface-adsorbate bond frequency as a function of the solid's Debye frequency. 135

E.1 Comparison of memory kernel (A) and power spectra (B) for single site fluctuations of 1D harmonic chains of various lengths with periodic boundary conditions. 142

List of Tables

3.1 Parameters used for computing desorption rate constants shown in Figure 3.11. 59

6.1 Relative computational cost for different methods of computing coupling and TDDFT calculations. All CPU times were measured on an Intel(R) Xeon(R) CPU E5-2620 @ 2.10GHz. The costs shown are for the entire 100ps trajectory, but are estimated using the cost of a single calculation, and account for the 20Åcoupling cutoff. 112

Part I

Phonon-induced memory in surface chemistry

‘

Chapter 1

Introduction to heterogeneous catalysis

1.1 The Sabatier principle

Catalysis is the science of accelerating chemical reactions by introducing a non-interacting intermediary. Perhaps the most renowned example is the Haber-Bosch process, in which an iron metal catalyst is used to convert atmospheric nitrogen into ammonia, a key ingredient in fertilizer[1, 2]. Without the Haber-Bosch process, Europe would have faced a massive famine in the early 20th century. Today, from plastics to pharmaceuticals, almost 30% of the world's industrial production relies on catalysis[3, 4]. Moreover, many of the most promising research programs for renewable energy and green chemistry hinge on the discovery of better catalysts[4–6]. For example, zeolites, naturally occurring aluminum and silicon crystals, have attracted significant interest as an alternative to the several environmentally toxic catalysts used in the petrochemical industry [6, 7].

Catalysts are divided into two categories based on the phase of the catalyst substance: homogeneous and heterogeneous. Homogeneous catalysts exist in the same phase as the reactants and include examples such as biological enzymes or acid catalysts. Heterogeneous catalysts, the focus of this thesis, exist in a different phase than the reactants and typically consist of a solid surface to which the reactants bind

(adsorb).

Due to the past and continuing importance of heterogeneous catalysts, understanding the physical principles governing their effectiveness and selectivity — particularly at the molecular scale — is a central challenge for physical chemistry and materials science. Significant progress has been made in this endeavor over the past century. Of particular note is the *Sabatier principle*, the fundamental thermodynamic principle that governs magnitude of the rate enhancement of a given heterogeneous catalyst.

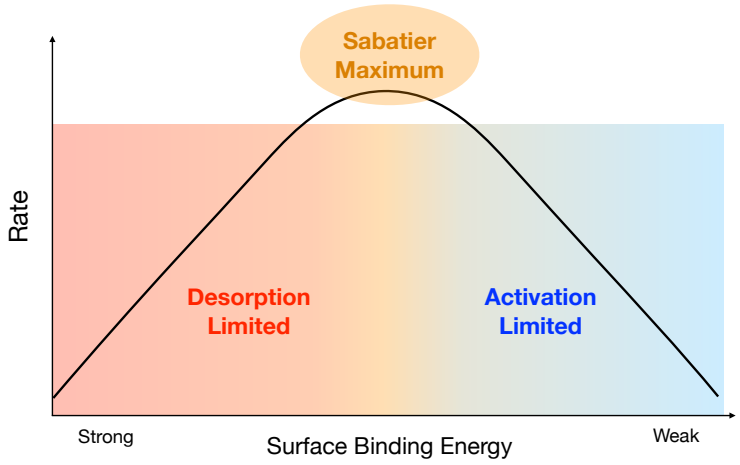

Figure 1.1: Schematic of Sabatier principle.

In Sabatier's seminal treatise, for which he would later win the Nobel Prize, he details the reactions of organic molecules across a wide variety of metal surfaces under different conditions [8]. Through his analysis of these reactions, he postulates that the optimal rate enhancement for a heterogeneous catalyst occurs when the binding energy between the surface and the reagent molecules is neither too high nor too low. If the binding energy is too high, the products will be unable to leave (desorb from) the surface. Conversely, if the binding energy is too low, the catalyst will be unable to sufficiently reduce the activation barrier of the reaction. Sabatier's principle directly leads to the familiar volcano-shaped plot of catalytic rates versus binding strengths (Figure 1) and has been verified across a wide variety of reactions using both experimental and computational methods [9–15].

Despite its many successes and explanatory value, the Sabatier principle is not a

complete theory. It does not address the physical properties that determine the binding strength, nor can it explain the selectivity of a given catalyst toward a particular product. In order to solve such problems, more detailed theories are needed, which specifically examine the interplay between the catalyst and molecular reagent degrees of freedom. Additionally, several experiments have demonstrated that when subject to time-dependent external forces, some catalysts can exhibit rate enhancements orders of magnitude larger than the static Sabatier limit [16, 17]. The mechanisms behind such dynamic effects are still poorly understood, and significant technological advancements may be possible if descriptive theories are developed.

Attempts to develop theories that build beyond Sabatier's principle typically focus on describing either the electronic or the nuclear degrees of freedom of the catalyst. Both are crucially important, but the two can often be separated owing to the validity of the Born-Oppenheimer approximation for heavy atoms at room temperature. In Part I of this thesis concerns the nuclear vibrations (phonons) of the catalyst surface and their coupling to the molecular degrees of freedom. The remainder of Chapter 1 is devoted to giving an overview of the study of phonons in heterogeneous catalysis and an outline of the major results presented thereafter.

1.2 Surface vibrations in heterogeneous catalysis

The study of the effects of surface vibrations in heterogeneous catalysis has its origins in molecular beam scattering experiments. In such an experiment, a high-pressure nozzle directs a beam of molecules towards a surface, and the flux of scattered products from the surface is subsequently measured, often as a function of the incident angle and velocity. As early as 1975, molecular beam experiments illustrated a sharp increase in the dissociation probability of methane (CH_4) when the catalyst surface was heated above a certain critical temperature [18, 19]. It is important to note that such effects are not observed when only increasing the temperature of the molecular beam, and cannot not be explained by the naive $\exp\left(-\frac{E_\ddagger}{k_B T}\right)$ scaling of Arrhenius theory. Further experimental and theoretical studies revealed that the

surface-temperature dependence arises from a strong coupling of the dissociation reaction's transition state to the displacement of the metal surface atoms[19–25]. Similar observations have been made for nitrogen dissociation on a variety of metal surfaces as well [26–29].

In order to gain insight into the timescales of energy transfer between metal vibrations and molecular adsorbates, many scattering experiments began to use laser pulses in order to heat the metal in specific locations and at specific rates [30–37]. A particularly notable example of such an experiment was done by Bonn et al.[38] In that study, the oxidation of CO to CO_2 on Ru(0001) was observed shortly after a near-infrared laser was used to heat the surface. It is worth noting that conductive heating of the surface does not increase the reactive flux. Time of flight spectra revealed a spike in CO_2 production occurring simultaneously with a spike in CO desorption, and thus the mechanism of the reaction is considered to be a concerted effect between hot electrons coupling to the oxidation step and hot phonons coupling to the desorption step.

Another very important class of experiments which demonstrate the coupling between surface vibrations and molecular adsorbates involves the use of acoustic waves to generate forced oscillations of the catalyst surface [16, 39–42]. In such experiments, a thin (10-100nm) catalyst surface is placed on top of a piezoelectric material (typically $LiNbO_3$). An oscillating electric field applied to this material creates a sound (acoustic) wave, which then propagates into the thin catalyst material. Several studies have demonstrated that these acoustic waves, when applied at specific frequencies and polarizations, can cause a substantial increase in the rates of certain reactions. However, the mechanisms behind this effect have been hotly debated in recent years.

One explanation is that the displacements caused by the acoustic waves alter the electronic structure of metal and therefore the binding strength to molecular adsorbates. Oscillation between high and low binding strengths allows for the adsorbate bonds to be highly activated, without poisoning of the surface, leading to reaction rates above the Sabatier limit. This explanation was supported by photoelectron emission microscopy (PEEM) experiments showing that acoustic waves produced re-

versible changes in the work function of the metal in tandem with an increase in reaction products [16, 39, 43]. However, it remained unclear whether the work function changes were themselves the causal factor, or just another consequence of an underlying physical mechanism. In 2020, a critical study published by von-Boehn et al.[42] quantified the magnitude of the local work-function changes and demonstrated that they were only on the order of a few μeV, far too small to substantially impact surface chemistry. Instead, von-Boehn et al. attributed the increased rate to surface roughening/restructuring, which they illustrated by showing that the SAW can cause the intermixing between platinum and rhodium thin films. However, von-Boehn et al. also admit that even this explanation is far from a definitive theory.

The abundance of experiments delineating the myriad of ways that surface vibrations affect catalytic reactions has spurned enormous interest in developing theoretical and computational methods which can reproduce said experiments and give insight into their atomistic mechanisms. In particular, the ability to calculate and predict how and when surface acoustic waves activate a given reaction could lead to development of finely-tunable and selective catalysts that could unlock chemical synthesis pathways previously impossible [16, 17]. However, relevant surface vibrations can couple motion spanning the nano-, meso-, and macroscopic scales, and therefore the numerical integration of equations of motion for model systems is costly and can exhibit spurious finite-size effects. A more prudent approach is to project the collective vibrations of a solid surface either onto the adsorbate itself or onto a collective of surface active sites, giving an effective equation of motion for these degrees of freedom [44–48]. This effective equation of motion, called a generalized Langevin equation, is *non-Markovian* with a time-dependent friction constant called a *memory kernel*. The memory kernel is of central importance to determining the timescales of energy exchange between adsorbates and the metal surface.

The following chapters develop a general theory for the coupling of phonons to molecular adsorbates within the projection operator / generalized Langevin equation framework. Chapter 2 provides an overview of Mori-Zwanzig projection operator theory and then describe the application of that theory to solid phonons. The formal

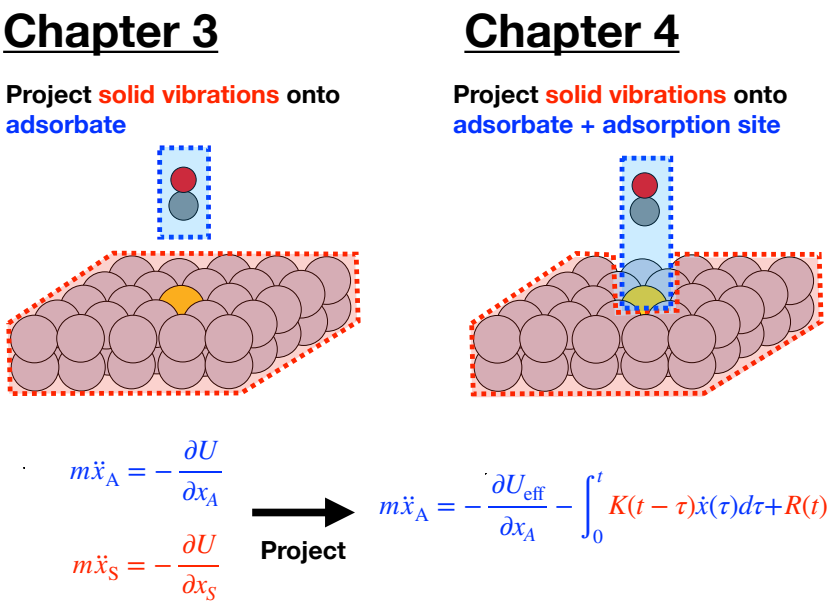

Figure 1.2: Schematic of models developed in Chapter 3 and Chapter 4. The equations shown illustrate how projections map the deterministic Newtonian motion to stochastic, non-Markovian motion.

properties of the memory kernel are described, including theories for how memory affects reaction rates near thermal equilibrium. Chapter 3 describes results for projecting vibrations directly onto the motion of adsorbates. It is shown that adsorbates couple to different phonons depending on the relative value of the molecule-surface bond frequency and solid Debye frequency. Using this theory phonon-contributions to transition state theory rates are calculated and compared with experiment. Chapter 4 describes results for projecting vibrations directly onto the motion of individual surface sites. Notably, a subtle, but universal source of error in many calculations of molecule-surface scattering is described.

Chapter 2

Theoretical preliminaries

Catalytic reactions are an intricate dance between the reagent and catalyst degrees of freedom. However, we are rarely interested in the motion of the catalyst itself, but instead the effect of the catalyst on the motion of the molecular reagent. It is therefore natural to desire theories which *project* the catalyst degrees of freedom onto the molecular reagent.

This chapter is devoted to a description of the mathematical background necessary to develop and understand such theories within classical statistical mechanics. First, an introduction to Mori-Zwanzig theory is given and subsequently used to derive a generalized Langevin equation (GLE) for a system interacting with a harmonic environment. Subsequently, the mathematical and physical properties of GLEs are explored, particularly focusing on the properties of the memory kernel and its Fourier transform. Finally, an overview of the role of memory in reaction rate theory is given.

2.1 Mori-Zwanzig theory

The theory of projection operators for dynamical equations of motion is called Mori-Zwanzig theory, after physicists Hazime Mori and Robert Zwanzig who developed the theory in the 1960s [49, 50]. The motivation for the theory is to use projection operators to reduce the dynamics of a complex, many-body system (either classical or quantum) onto effective equations of motion for a set of collective variables

of interest. In the most general formulation of the theory, neither the dynamics nor the projection operators need be linear. However, the fully non-linear Mori-Zwanzig equations are often just as difficult to analyze and solve as the original equations of motion[51]. Interested readers may find a discussion of the non-linear Mori-Zwanzig equations in Appendix A.

Fortunately, when considering the nuclear vibrations of a solid we can approximate the potential energy as harmonic. Furthermore, since our collective variables (the adsorbate) are just a subset of the total degrees of freedom, the projection operators are also linear. When both the dynamics and the projection operators are linear, an exact form of the generalized Langevin equation for the collective variables be derived. The derivation given below essentially follows the one given by Zwanzig himself in Ref. [52].

Let $\mathbf{q}$ denote a vector containing the coordinates of a classical many-body system. We separate these degrees of freedom into a subspace corresponding to the degrees of freedom of interest $\mathbf{q}_\mathrm{A}$ and the remainder $\mathbf{q}_\mathrm{B}$. Henceforth, we term the degrees of freedom of interest as *system*, and the remainder as the *bath*. For example, in Chapter 3 the system will correspond to the adsorbate degrees of freedom, and the bath correspond to the solid degrees of freedom.

It is convenient to also separate the total potential energy $V(\mathbf{q}_\mathrm{A}, \mathbf{q}_\mathrm{A})$ into two parts,

$$V(\mathbf{q}_\mathrm{A}, \mathbf{q}_\mathrm{B}) = V_\mathrm{A}(\mathbf{q}_\mathrm{A}) + V_\mathrm{AB}(\mathbf{q}_\mathrm{A}, \mathbf{q}_\mathrm{B}). \tag{2.1}$$

where $V_\mathrm{A}(\mathbf{q}_\mathrm{A})$ contains any terms that only depend on $\mathbf{q}_\mathrm{A}$, and $V_\mathrm{AB}(\mathbf{q}_\mathrm{A}, \mathbf{q}_\mathrm{B})$ contains the remaining terms. The purpose of this separation is to preserve any anharmonicites in $V_\mathrm{A}(\mathbf{q}_\mathrm{A})$. Expanding V_AB to second-order around the equilibrium positions $\bar{\mathbf{q}}$ gives,

$$V_\mathrm{AB}(\mathbf{q}_\mathrm{A}, \mathbf{q}_\mathrm{A}) \approx \sum_{ij} H_{ij}(q_i - \bar{q}_i)(q_j - \bar{q}_j), \tag{2.2}$$

where the summation indices go over all the degrees of freedom both of the system and the bath, and the matrix elements H_{ij} correspond to the Hessian of V_AB evaluated

at the equilibrium positions,

$$H_{ij} = \frac{1}{2}\frac{\partial^2 V_{\mathrm{AB}}}{\partial q_i \partial q_j}\bigg|_{\bar{q}_i,\bar{q}_j} . \tag{2.3}$$

Using Eq. 2.1 and Eq. 2.2 we may express the total system and bath equations of motion as,

$$\begin{pmatrix} \mathbf{m}_{\mathrm{A}}\ddot{\mathbf{q}}_{\mathrm{A}}(t) \\ \\ \mathbf{m}_{\mathrm{B}}\ddot{\mathbf{q}}_{\mathrm{B}}(t) \end{pmatrix} = - \begin{pmatrix} \dfrac{\partial V_{\mathrm{A}}}{\partial \mathbf{q}_{\mathrm{A}}}(t) \\ \\ 0 \end{pmatrix} - \begin{pmatrix} \\ \mathbf{H} \\ \\ \end{pmatrix} \begin{pmatrix} \mathbf{q}_{\mathrm{A}}(t) - \bar{\mathbf{q}}_{\mathrm{A}} \\ \\ \mathbf{q}_{\mathrm{B}}(t) - \bar{\mathbf{q}}_{\mathrm{B}} \end{pmatrix} . \tag{2.4}$$

It is convenient to introduce mass-weighted coordinates $x_i = \sqrt{m_i}(q_i - \bar{q}_i)$, such that the equations of motion may be written as,

$$\begin{pmatrix} \ddot{\mathbf{x}}_{\mathrm{A}}(t) \\ \\ \ddot{\mathbf{x}}_{\mathrm{B}}(t) \end{pmatrix} = - \begin{pmatrix} \dfrac{\partial V_{\mathrm{A}}}{\partial \mathbf{x}_{\mathrm{A}}}(t) \\ \\ 0 \end{pmatrix} - \begin{pmatrix} \\ \mathbf{D}^2 \\ \\ \end{pmatrix} \begin{pmatrix} \mathbf{x}_{\mathrm{A}}(t) \\ \\ \mathbf{x}_{\mathrm{B}}(t) \end{pmatrix} , \tag{2.5}$$

where $\mathbf{D}^2$ is the *mass-weighted* Hessian,

$$D^2_{ij} = \frac{1}{\sqrt{m_i m_j}}\frac{\partial^2 V_{\mathrm{AS}}}{\partial q_i \partial q_j}\bigg|_{\bar{q}_i,\bar{q}_j} . \tag{2.6}$$

We now introduce two projection operators $\mathcal{A}$ and $\mathcal{B}$. $\mathcal{A}$ is a matrix with ones on the diagonal corresponding to the indices of the system degrees of freedom such that $\mathcal{A}\mathbf{x} = \mathbf{x}_{\mathrm{A}}$, and $\mathcal{B} = \mathbb{1} - \mathcal{A}$ is it's orthogonal complement,

$$\mathcal{A} = \left(\begin{array}{c|c} \mathbb{1} & 0 \\ \hline \\ 0 & 0 \\ \\ \end{array}\right) \qquad \mathcal{B} = \left(\begin{array}{c|c} 0 & 0 \\ \hline \\ 0 & \mathbb{1} \\ \\ \end{array}\right) \tag{2.7}$$

With these projection operators we can separate equation Eq. 2.6 into,

$$\ddot{\mathbf{x}}_{\mathrm{A}}(t) = -\frac{\partial V_{\mathrm{A}}}{\partial \mathbf{x}_{\mathrm{A}}}(t) - \mathbf{D}^2_{\mathrm{AA}}\mathbf{x}_{\mathrm{A}}(t) - \mathbf{D}^2_{\mathrm{AB}}\mathbf{x}_{\mathrm{B}}(t), \tag{2.8}$$

and,

$$\ddot{\mathbf{x}}_{\mathrm{B}}(t) = -\mathbf{D}^2_{\mathrm{BA}}\mathbf{x}_{\mathrm{A}}(t) - \mathbf{D}^2_{\mathrm{BB}}\mathbf{x}_{\mathrm{B}}(t), \tag{2.9}$$

where $\mathbf{D}^2_{\mathrm{AA}} = \mathcal{A}\mathbf{D}^2\mathcal{A}$, $\mathbf{D}^2_{\mathrm{AB}} = \mathcal{A}\mathbf{D}^2\mathcal{B}$, $\mathbf{D}^2_{\mathrm{BA}} = \mathcal{B}\mathbf{D}^2\mathcal{A}$, and $\mathbf{D}^2_{\mathrm{BB}} = \mathcal{B}\mathbf{D}^2\mathcal{B}$.

In order to solve Eq. 2.9 we must introduce a set of normal modes which diagonalize $\mathbf{D}^2_{\mathrm{BB}}$. Let $\mathbf{u}_{\mathrm{B}} = \mathbf{U}^T\mathbf{x}_{\mathrm{B}}$, where $\mathbf{U}$ is a matrix with the eigenvectors of $\mathbf{D}^2_{\mathrm{BB}}$ as columns. In terms of these normal modes Eq. 2.9 may be written as,

$$\ddot{\mathbf{u}}_{\mathrm{B}}(t) = -\mathbf{U}^T\mathbf{D}^2_{\mathrm{BA}}\mathbf{x}_{\mathrm{A}}(t) - \boldsymbol{\Omega}^2\mathbf{u}_{\mathrm{B}}(t), \tag{2.10}$$

where $\boldsymbol{\Omega}^2 = \mathbf{U}^T\mathbf{D}^2_{\mathrm{BB}}\mathbf{U}$ is a diagonal matrix of the squared frequencies of the normal modes. Each row of Eq. 2.10 is now 2nd-order inhomogeneous differential equation that may be solved analytically to yield,

$$\mathbf{u}_{\mathrm{B}}(t) = \cos(\boldsymbol{\Omega}t)\mathbf{u}_{\mathrm{B}}(0) + \frac{\sin(\boldsymbol{\Omega}t)}{\boldsymbol{\Omega}}\dot{\mathbf{u}}_{\mathrm{B}}(0) - \int_0^t d\tau \frac{\sin(\boldsymbol{\Omega}(t-\tau))}{\boldsymbol{\Omega}}\mathbf{C}^T\mathbf{x}_{\mathrm{A}}(\tau), \tag{2.11}$$

where $\mathbf{C}^T = \mathbf{U}^T\mathbf{D}^2_{\mathrm{BA}}$ is the matrix of the couplings between $\mathbf{x}_{\mathrm{A}}$ and each normal mode. Substituting Eq. 2.11 into Eq. 2.8 yields a GLE,

$$\ddot{\mathbf{x}}_{\mathrm{A}}(t) = -\frac{\partial V_{\mathrm{A}}}{\partial \mathbf{x}_{\mathrm{A}}}(t) - \mathbf{D}^2_{\mathrm{AA}}\mathbf{x}_{\mathrm{A}}(t) - \int_0^t d\tau \boldsymbol{\Theta}(t-\tau)\mathbf{x}_{\mathrm{A}}(t) + \mathbf{R}_{\Theta}(t), \tag{2.12}$$

where $\boldsymbol{\Theta}(t)$ and the $\mathbf{R}_{\Theta}(t)$ are the memory kernel and random forces respectively, satisfying,

$$\boldsymbol{\Theta}(t) = -\mathbf{C}\frac{\sin(\boldsymbol{\Omega}t)}{\boldsymbol{\Omega}}\mathbf{C}^T, \tag{2.13}$$

$$\mathbf{R}_{\Theta}(t) = -\mathbf{C}\left(\cos(\boldsymbol{\Omega}t)\mathbf{u}_{\mathrm{B}}(0) + \frac{1}{\boldsymbol{\Omega}}\sin(\boldsymbol{\Omega}t)\dot{\mathbf{u}}_{\mathrm{B}}(0)\right). \tag{2.14}$$

Although $\mathbf{R}_{\Theta}(t)$ is not truly random, it is entirely independent of the position of the

system. Furthermore, it can be shown that for a very large bath comprised of $\sim 10^{23}$ normal modes $\mathbf{R}_\Theta(t)$ limits to random Gaussian noise [53]. The random forces and memory kernel satisfy a fluctuation-dissipation relationship. In order to derive this relationship, note that the equilibrium correlation between any two normal modes is, $\langle u_{\mathrm{B},i} u_{\mathrm{B},j} \rangle = \dfrac{k_\mathrm{B}T}{\omega_i^2}\delta_{ij}$, where ω_i is the frequency of the ith mode. Thus, taking the autocorrelation function of $\mathbf{R}_\Theta(t)$ we have,

$$\left\langle \mathbf{R}_\Theta(t)\mathbf{R}_\Theta^T(0)\right\rangle = k_\mathrm{B}T\mathbf{C}\frac{\cos(\boldsymbol{\Omega}t)}{\boldsymbol{\Omega}^2}\mathbf{C}, \tag{2.15}$$

$$\boldsymbol{\Theta}(t) = \frac{1}{k_\mathrm{B}T}\frac{d}{dt}\left\langle \mathbf{R}_\Theta(t)\mathbf{R}_\Theta^T(0)\right\rangle \tag{2.16}$$

The memory integral in Eq. 2.12 depends on the displacement of the system from its equilibrium position, and such a dependence is problematic in cases where the position of the system is unbound, leading the memory integral to diverge. Fortunately, this dependence can be removed by integrating the memory integral by parts, leading to a different GLE,

$$\ddot{\mathbf{x}}_\mathrm{A}(t) = -\frac{\partial V_\mathrm{A}}{\partial \mathbf{x}_\mathrm{A}}(t) - \left[\mathbf{D}_{\mathrm{AA}}^2 - \mathbf{K}(t=0)\right]\mathbf{x}_\mathrm{A}(t) - \int_0^t d\tau \mathbf{K}(t-\tau)\dot{\mathbf{x}}_\mathrm{A}(t) + \mathbf{R}(t), \tag{2.17}$$

$$\mathbf{K}(t) = \mathbf{C}\frac{\cos(\boldsymbol{\Omega}t)}{\boldsymbol{\Omega}^2}\mathbf{C}^T, \tag{2.18}$$

$$\mathbf{R}(t) = -\mathbf{C}\left(\cos(\boldsymbol{\Omega}t)\mathbf{u}_\mathrm{B}(0) + \frac{\sin(\boldsymbol{\Omega}t)}{\boldsymbol{\Omega}}\dot{\mathbf{u}}_\mathrm{B}(0) + \frac{\cos(\boldsymbol{\Omega}t)}{\boldsymbol{\Omega}^2}\mathbf{C}^T\mathbf{x}_\mathrm{A}(0)\right). \tag{2.19}$$

$\mathbf{K}$ and $\mathbf{R}$ also satisfy a fluctuation-dissipation theorem,

$$\mathbf{K}(t) = \frac{1}{k_\mathrm{B}T}\left\langle \mathbf{R}(t)\mathbf{R}^T(0)\right\rangle. \tag{2.20}$$

Eq. 2.17 is the form of the generalized Langevin equation that we will utilize in subsequent chapters. We see that by integrating out the bath degrees of freedom, we are left with a non-Markovian differential equation for the system, which depends not only on the instantaneous value of the position/velocity, but also on the entire history of past velocity values. This non-Markovianity or *memory* arises because

of the time-lag between the force of the system on the bath, and the response of the bath. Since we assumed that the bath potential energy V_{AB} was harmonic, we have found that the non-Markovian response can be encoded in terms of a linear-operator $\mathbf{K}(t)$. The properties of this memory kernel are crucial for understanding the dynamics induced by Eq. 2.17. In the subsequent section, the general mathematical and physical properties of memory kernels are discussed.

2.2 Properties of memory kernels

The memory kernel is the most important quantity in Eq. 2.17 as it encapsulates both how the bath affects both the equilibrium and the non-equilibrium statistics of the system. For the equilibrium statistics, the dynamics of the bath widen the local fluctuations of the system around its equilibrium position. This widening can be seen from the total potential energy in Eq. 2.17,

$$V_{S,\mathrm{eff}}(\mathbf{x}_{\mathrm{A}}) = V_S(\mathbf{x}_{\mathrm{A}}) + \mathbf{x}_{\mathrm{A}}^T \left(\mathbf{D}_{AA}^2 - \mathbf{K}(t=0)\right) \mathbf{x}_{\mathrm{A}}, \tag{2.21}$$

where the mass-weighted Hessian of the system $\mathbf{D}_{AA}^2$ is reduced by the instantaneous component of memory kernel, $\mathbf{K}(t=0)$. It is important to note that for an general, anharmonic bath, the equilibrium statistics of the system are determined by the potential of mean force,

$$W(\mathbf{x}_{\mathrm{A}}) = -k_{\mathrm{B}}T \ln \left[\int d\mathbf{x}_{\mathrm{B}} V_{\mathrm{A}}(\mathbf{x}_{\mathrm{A}}, \mathbf{x}_{\mathrm{B}}) \right], \tag{2.22}$$

and that $V_{S,\mathrm{eff}}$ is just a local approximation to the potential of mean force. An illustrative example of how the potential energy widens due to the presence of the bath is provided in Figure 2.1. For the non-equilibrium statistics, since $\langle \mathbf{R}(t) \rangle = 0$, if $\mathbf{R}(t)$ is modeled as Gaussian noise its statistics are purely determined by the two-time correlation function, and therefore the memory kernel by virtue of Eq. 2.20.

The properties of of the memory kernel are often more transparent in frequency space than in real-time. In order to analyze the Fourier transform of the memory

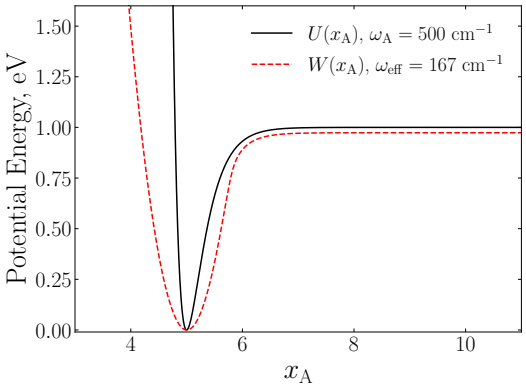

Figure 2.1: Bare potential energy (U) and potential of mean force (W) for a Morse oscillator system coupled to a single harmonic mode bath. The bath mode widens the width of the bare Morse potential, such that the frequency at the minimum shifts from $\omega_A = 500$ cm^{-1} to $\omega_{eff} = 167$ cm^{-1}.

kernel, it is easiest to first express $\mathbf{K}(t)$ in index form,

$$K_{ij}(t) = \sum_k \frac{C_{ik}C_{jk}}{\omega_k^2} \cos(\omega_k t), \tag{2.23}$$

where C_{ik} is an element of the coupling matrix $\mathbf{C}$. Taking the Fourier transform of Eq.2.23 leads to,

$$\tilde{K}_{ij}(\omega) = \pi \sum_k \frac{C_{ik}C_{jk}}{\omega_k^2} \left(\delta(\omega - \omega_k) + \delta(\omega + \omega_k)\right). \tag{2.24}$$

As the negative frequency components of Eq. 2.24 are mere a reflection of the positive frequency components, it is common drop them leading to,

$$\bar{K}_{ij}(\omega) = \sum_k \frac{C_{ik}C_{jk}}{\omega_k^2} \delta(\omega - \omega_k). \tag{2.25}$$

$\bar{K}_{ij}(\omega)$ is called the *power spectral density* or simply *spectral density* of the bath. The name arises from the Wiener-Khinchin theorem, which states that the Fourier transform of the autocorrelation function of a stationary random process is equivalent

to the power spectrum of that process. Alternatively we can re-express Eq. 2.25 as,

$$\bar{K}_{ij}(\omega) = \frac{C_i(\omega)C_j(\omega)}{\omega}\rho(\omega), \tag{2.26}$$

where $\rho(\omega) = \sum_k \delta(\omega - \omega_k)$ is the bath density of states. Equations 2.25 and 2.26 reveal that the Fourier transform of $\mathbf{K}(t)$ can be interpreted as the density of states of the bath reweighted by the value of the system-bath coupling $\mathbf{C}$. Eq. 2.26 is a more appropriate definition of the spectral density when the bath is made up of a continuum of normal modes.

We can learn much about the physical properties of a given bath and how it interacts with the system by analyzing the shape of the memory kernel and spectral density. For a Markovian (memoryless) bath, the memory kernel is simply a delta function, and thus the spectral density is a uniform distribution across all frequencies. For a particle moving in a liquid, the memory kernel tends to decay exponentially, $K(t) \sim e^{-\alpha t}$, and the spectral density is very broad as there are no distinct normal modes in liquids. The broad, flat spectral density supports why Markovian approximations are often reasonable when simulating the dynamics of particles of liquids. In contrast, in a glassy environment both the memory kernel and the spectral density can behave like a power-law, $K(t) \sim t^{-\alpha}$, with long-tails due to the slow rate of relaxation. In fact, the presence of long tails in the memory kernel is one way to quantify the ergodicity breaking. Finally, many environments behave like damped harmonic oscillators; that is, they have one or more central frequencies but also dissipate energy at a certain rate. In such cases, the memory kernel will behave like a damped sinusoid $K(t) \sim e^{-\alpha t}\cos(\omega_0 t)$ and the spectral density will be a Lorentzian centered at a finite frequency. In Chapters 3 and 4 we will show that the vibrations of a solid surface indeed behave in such a manner.

One may wonder why, given the fact that the GLE in Eq. 2.17 was derived under the assumption that the bath was harmonic, the GLE can and often is be applied to baths which are highly anharmonic. The answer to this question is subtle and relies on linear response theory. One can show that the fluctuation-dissipation theorem

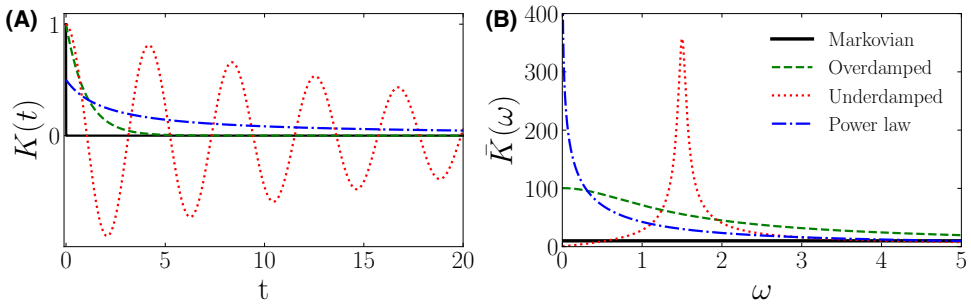

Figure 2.2: Depiction of characteristic forms of the (A) memory kernel and (B) spectral density. The overdamped form is common in systems coupled to liquid environments, while the underdamped form common in systems coupled to solid environments.

Eq. 2.20 does not require the bath itself to be harmonic, only that the statistics of the bath change linearly due to forces from the system. Thus, the true test of the validity of the GLE is whether one can safely say that the bath remains near thermal equilibrium regardless of the state of the system. Further insight can be found in Ayaz et al.[54]

2.3 Reaction rate theory

One of the most important physical consequences of memory is how it can affect reaction rates. For a 1D system embedded in a Markovian bath, Kramers[55] showed that the rate constant may be expressed as,

$$k_{\mathrm{Kr}} = \frac{\omega_0}{2\pi}\frac{1}{\omega^{\ddagger}}\left(-\frac{\gamma}{2}+\sqrt{\frac{\gamma^2}{4}+(\omega^{\ddagger})^2}\right)e^{-\beta E^{\ddagger}}, \tag{2.27}$$

where ω_0 is the frequency of the reactant well (see Figure 2.3), $\omega^{\ddagger}$ is the frequency of the barrier, $E^{\ddagger}$ is the height of the barrier, and γ is the Markovian friction constant. γ can be expressed as a limit of the memory kernel integrated across time,

$$\gamma = \int_0^{\infty} dt K(t). \tag{2.28}$$

Kramers' theory reveals that while friction/memory affects the rate prefactor, but the height of the barrier is left unchanged. To derive Eq. 2.27, Kramers exploited the

equivalence between the Langevin equation and the Fokker-Planck equation in order to solve for the probability flux towards the products at the top of the barrier.

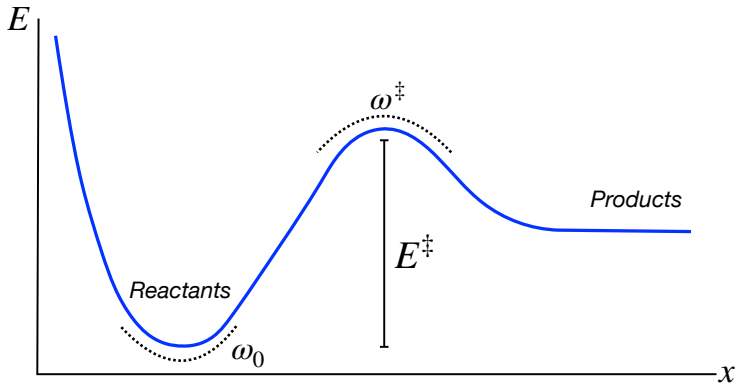

Figure 2.3: Characteristic potential energy curve for barrier-crossing problem.

Grote and Hynes generalized Kramers formula for a non-Markovian bath[56],

$$k_{\mathrm{GH}} = \frac{\omega_0}{2\pi}\frac{\lambda}{\omega^{\ddagger}}e^{-\beta E^{\ddagger}}, \tag{2.29}$$

where λ is the unique real, positive solution to the algebraic equation,

$$\lambda^2 + \lambda K(\lambda) + (\omega^{\ddagger})^2 = 0, \tag{2.30}$$

and $K(\lambda)$ is the Laplace transform of the memory kernel. If one takes the memory kernel to be a delta function, the Grote-Hynes rate constant is equivalent to the Kramers rate constant.

Both Kramers and Grote-Hynes rate constants rely on the assumption that the attempts to make it over the barrier are generated by a Gaussian distribution, or put more simply, that the reactants are in thermal equilibrium. This assumption is not valid if the friction constant is very large and/or the decay rate of the memory kernel is very slow (see Figure 2.4). Indeed Kramers himself noted that there is a turnover in the dependence of the rate constant with friction that is not captured by Eq. 2.27 and Kramers turnover has been verified experimentally[55, 57, 58]. Unfortunately, for non-Markovian environments, there is no general form for the rate-constant for for an arbitrary memory kernel. For an overview of perturbative solutions that are

possible for exponential memory, see Ref. 59 and Ref. 60.

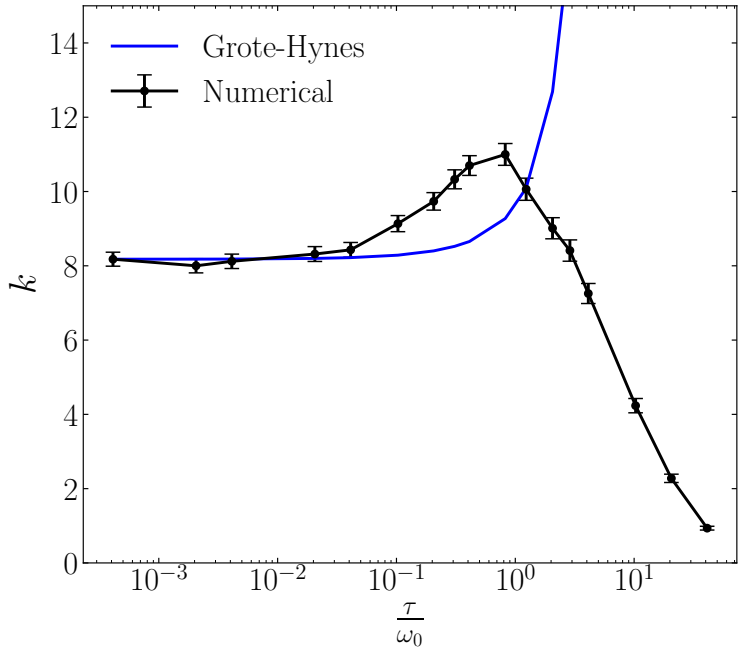

Figure 2.4: Grote-Hynes rate constants versus numerical simulation results for a 1D barrier crossing problem with exponential memory, $K(t) = Ae^{-\frac{t}{\tau}}$.

Another approach to calculating rate constants in chemistry is called transition state theory (TST), first developed by Eyring and Wigner [61–63]. TST is a philosophically different approach to the Kramers-Grote-Hynes (KGH) theory. KGH theory presupposes a one-dimensional reaction coordinate and captures the effects of other degrees of freedom implicitly through the friction constant/memory kernel. In contrast, TST posits the existence of a high-dimensional dividing surface in coordinate space, $f(\mathbf{q}) = 0$ that seperates reactants and products. The transition state is defined as a saddle point along this dividing surface such that if a trajectory crosses the saddle point, it cannot re-cross. TST thus calculates the rate constant as the equilibrium flux through this saddle point,

$$k_{\mathrm{TST}} = \frac{1}{Z_R} \int d\mathbf{q} d\mathbf{p}\; e^{-\beta H(\mathbf{q},\mathbf{p})} \delta\left[f(\mathbf{q})\right] \left(\nabla f \cdot \mathbf{p}\right) \Theta\left(\nabla f \cdot \mathbf{p}\right), \tag{2.31}$$

where $H = \sum_i^{3N} \frac{p_i^2}{2m_i} + V(\mathbf{q})$ is the Hamiltonian, ∇f is the normal vector to the surface, Θ is the Heaviside step function, and Z_R is the reactant partition function.

Z_R is defined as,

$$Z_R = \frac{1}{h^{3N}} \int_R d\mathbf{q} d\mathbf{p}\ e^{-\beta(H(\mathbf{q},\mathbf{p})-E_R)}. \tag{2.32}$$

where the subscript R denotes that the integral is taken only over the reactant region in position space. Eq. 2.31 is often simplified by carrying out the momentum and position integrals along the surface normal vector in the numerator giving,

$$k_{\mathrm{TST}} = \frac{k_{\mathrm{B}}T}{h}\frac{Z^{\ddagger}}{Z_R}e^{-\beta\Delta E^{\ddagger}}, \tag{2.33}$$

where $\Delta E^{\ddagger} = E^{\ddagger} - E_R$ is energy difference between the transition and reactant state, and $Z^{\ddagger}$ is a partition function for configurations taken near the transition state,

$$Z^{\ddagger} = \frac{1}{h^{3N-1}} \int_R d\mathbf{q}^{3N-1} d\mathbf{p}^{3N-1}\ e^{-\beta(H(\mathbf{q}^{3N-1},\mathbf{p}^{3N-1})-E^{\ddagger})}. \tag{2.34}$$

In Eq. 2.34, $\mathbf{q}^{3N-1}$ $\mathbf{p}^{3N-1}$ refer the $3N-1$ position and momentum coordinates that orthogonal to ∇f. Eq. 2.33 is the form of the TST rate constant found in most modern texts. Often the partitions functions Z_R and $Z^{\ddagger}$ in Eq. 2.33 are approximated by products of harmonic partition functions,

$$k_{\mathrm{TST}} \approx \frac{1}{2\pi}\frac{\prod_{i=1}^{3N}\omega_i}{\prod_{i=1}^{3N-1}\omega_i^{\ddagger}}e^{-\beta\Delta E^{\ddagger}}, \tag{2.35}$$

where ω_i are the eigenfrequencies of the mass-weighted Hessian evaluated at the reactant minimum, and $\omega^{\ddagger}$ are the real, positive eigenfrequencies of the mass-weighted Hessian evaluated at the transition state. Note that because the transition state is defined to be a saddle point, one of the eigenfrequencies at the transition state is always imaginary.

Both TST and KGH theory assume that the reactants remain in thermal equilibrium and that there are no recrossings over the barrier. However, Eq. 2.33 is far more expressive and generally applicable than KGH theory. Surprisingly though, Eq. 2.35 can be shown to be equivalent to KGH theory for a specific class of harmonic system-

bath Hamiltonians, as was shown by Pollak in his seminal paper [64]. Despite this equivalence, we emphasize that even Eq. 2.35 is still more general than KGH theory, as KGH theory assumes that the system-bath coupling remains the same in the reactant state and the transition state, while Eq. 2.35 does not necessarily. In Section 3.1, we will apply Eq. 2.35 to derive corrections factors for desorption constants due to phonons.

Chapter 3

Phonon-induced memory on adsorbates [1]

3.1 Introduction

As discussed in Chapter 1, the vibrations of a heterogeneous catalyst can significantly influence the rates and mechanisms of surface chemical processes.[12, 66–68] Unraveling these influences experimentally is difficult, motivating the development of theoretical tools. In this chapter, we introduce a theory for the coupling between the nuclear dynamics of surface-bound (adsorbed) molecules and the vibrations of the underlying surface. The theory projects the collective surface vibrations onto the motion of the adsorbate via the Mori-Zwanzig formalism discussed in Section 2.1, and in doing so, describes the motion of surface-bound molecules via a generalized Langevin equation (GLE).

Using this formalism, we demonstrate that the influence of surface vibrations on the dynamics of an adsorbed molecule depends significantly on the magnitude of the adsorbate-surface coupling. Specifically, we find that chemisorbed species couple primarily to dispersionless local vibrations, while physisorbed species couple primarily to acoustic phonons. The key parameter that determines whether the primary influence of surface vibrations is via extended or localized phonon modes is the ratio

[1] Based on work published in Ref. 65.

of the solid's Debye frequency to the frequency of the adsorbate-surface bond (Figure 3.1). By combining these observations with harmonic transition state theory, we derive equations that describe how phonons alter reaction rates at solid surfaces and demonstrate that these phononic corrections agree with experimental measurements of desorption rate constants. Below we summarize the topic and main findings of each remaining section of this chapter.

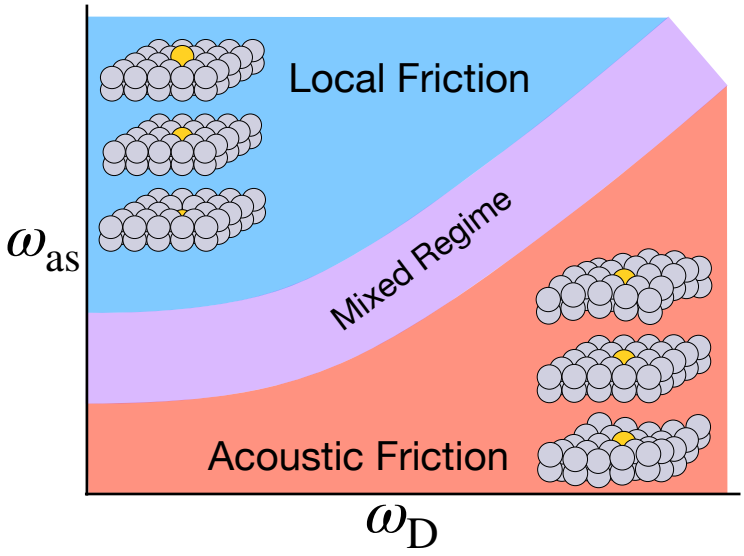

Figure 3.1: 2D schematic illustrating the dominant phonon modes in terms of their coupling to the adsorbate or contribution to the memory kernel. ω_{as} is the frequency of the adsorbate-surface bond and the ω_D is the solid's Debye frequency.

In Section 3.2, we begin by presenting the physical and computational model we used to compute the phonon-induced memory kernel. Subsequently, we analyze results for the memory kernel across a range of values for the adsorbate-surface bond frequency ω_{as}, illustrating that the mode contributing to the memory kernel is an acoustic phonon for physisorbed species, and a local phonon for chemisorbed species. We show this result is qualitatively independent of the details of the atomistic model used.

In Section 3.3, we use perturbation theory to interrogate the generality of the results presented in the previous section. We show that in the chemisorbed regime, the bulk phonons are a small perturbation to the motion of the adsorption site, while in the physisorbed regime, the local vibrations of the adsorption site are a small

perturbation to the bulk phonons. We derive analytical results for the memory kernel in the chemisorbed regime.

In Section 3.4, we analyze the role phonon dispersion plays in the memory kernel. Since the acoustic mode has a linear dispersion, summing over k-space broadens the acoustic peak in the spectral density. In contrast, the localized phonon modes are dispersionless, and thus spectral density in the chemisorbed limit remains roughly the same.

In Section 3.5, we analyze the effect of anharmonicity in the metal potential energy surface on the shape of the memory kernel. It is well known that anharmonicities couple phonon modes together, broadening the spectral density. We show that this is indeed the case for our models, but further illustrate that surface diffusion, if sufficiently fast, also broadens the spectral density. We conclude by illustrating that chemisorbed species can form surface polarons.

In Section 3.6, we use transition state theory to develop a general methodology for computing the phononic contribution to reaction rates at catalyst interface. We subsequently apply that methodology to surface desorption and validate our theory by demonstrating an improved agreement with experiments of CO desorption from Pt(111).

Finally, in Section 3.6 we discussing the relationship between the memory kernel and response functions to external forces. We demonstrate that adsorbates couple strongly to applied extensional and surface shockwave forces, but do not couple strongly to applied shear forces. We contextualize these results in terms of experiments analyzing surface chemistry under applied acoustic waves.

3.2 Memory for physisorbed and chemisorbed species

We have studied the phonon memory kernel acting on the surface normal (desorption) coordinate for a simple model of CO on Pt(111). The CO was modeled as a single adatom and interacting with a single adsorption site, as experimental structures show that CO adsorbs primarily atop Pt(111) sites [69, 70]. The metal potential

energy surface V_S was modeled using an effective medium theory (EMT) forcefield developed by Norskov et al.[71, 72] All calculations were performed using the Atomic Simulation Environment (ASE) python package [73–78].

In Equations 2.23 and 2.25 we gave the general formula for the memory kernel, and the spectral density (its Fourier transform) in terms of the normal mode frequencies of the bath and their coupling to the system. In our model, these normal mode frequencies correspond to the phonon modes of the solid surface, and the couplings correspond to the interaction, mediated by the adsorbate-surface bond, between each phonon mode of the solid and the adsorbate. Since we have assumed that CO interacts with a single surface Pt atom, Eqs. 2.23 and 2.25 may be simplified to,

$$K(t) = \frac{\mu^2}{mM}\omega_{\text{as}}^4 \sum_j \frac{U_{sj}^2}{\omega_j^2}\cos(\omega_j t), \tag{3.1}$$

$$\bar{K}(\omega) = \frac{\mu^2}{mM}\omega_{\text{as}}^4 \sum_j \frac{U_{sj}^2}{\omega_j^2}\delta(\omega - \omega_j), \tag{3.2}$$

where m, M, and μ are the CO mass, Pt mass, and reduced CO-Pt mass respectively, U_{sj} is the expansion coefficient of the adsorption site s in the jth normal mode, and ω_{as} is the adsorbate-surface bond frequency. In the remainder of this section, we will artificially vary ω_{as} over a range of physically motivated values in order to assess how the strength of the adsorbate-surface interaction affects the properties of the memory kernel.

Figure 3.2 illustrates the results of our calculations for a selected set of ω_{as} values ranging from those characteristic of weakly adsorbed species to those characteristic of strongly adsorbed species. We plot the memory kernel computed in the direction normal to the surface in systems with either a 4x4x8 or 8x8x8 Pt(111) slab. The lower values of ω_{as} that we consider lie below that of the platinum's Debye frequency of $\omega_{\text{D}} = 156\text{cm}^{-1}$. We associate these values with the physisorbed regime, as they are characteristic of weakly adsorbed systems such as noble gases on Pt(111).[79] The higher values of ω_{as} that we consider are associated with the chemisorbed regime, corresponding to physical systems such as CO on Pt(111), which has $\omega_{\text{as}} \approx 480\ \text{cm}^{-1}$).[70,

80, 81] Delta function peaks in $\bar{K}(\omega)$ and $\rho(\omega)$ were broadened to thin Lorentzians of width 1cm^{-1} for ease of visualization.

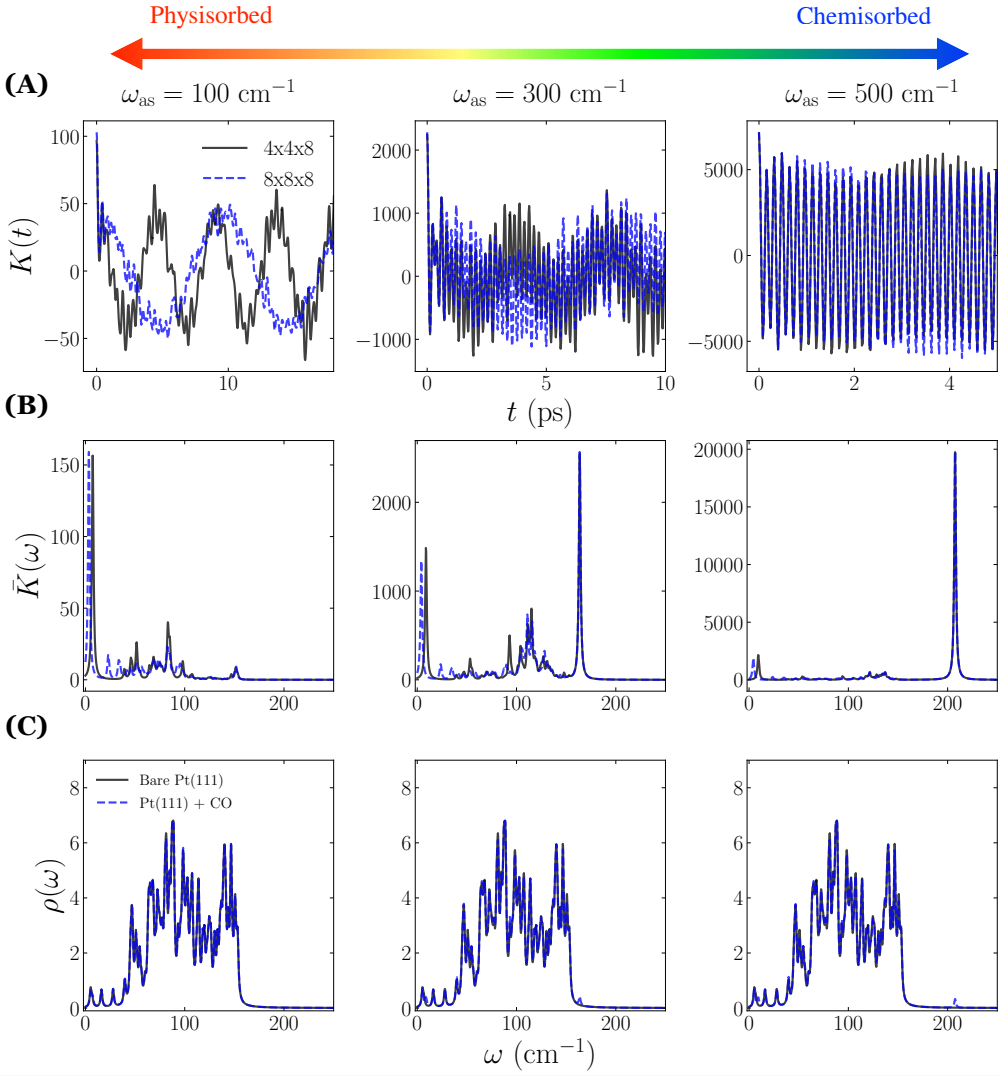

Figure 3.2: (A) Memory kernel, (B) spectral density, and (C) density of states for Pt(111) and three values of ω_{as}. The density of states in (C) was calculated using the 4x4x8 surface slab.

For all cases of ω_{as}, the adsorbate couples strongly to a surface acoustic mode appearing around $\omega = 10\text{cm}^{-1}$. The location of this peak is highly dependent on the dimensions of the simulated surface slab, consistent with the behavior of an acoustic phonon. In Section 3.5, we give a more detailed theoretical and numerical analysis of the size effects observed here, fully accounting for phonon dispersion. In the physisorbed regime, the memory kernel is dominated by this acoustic mode,

resulting in highly non-Markovian behavior. The phonon mode associated with this peak is associated with the flexing of the lattice in the direction perpendicular to the surface. Snapshots of this motion are provided in the insets in Figure 3.1.

In the chemisorbed regime, the memory kernel is dominated by a high frequency mode around $\omega = 210$ cm^{-1}. This mode arises from the local oscillations of the surface site the adsorbate is bound to. The frequency of this mode is independent of simulation cell size, suggesting that it is dispersionless. The large amplitude of this mode signifies that in the chemisorbed limit, the adsorbate is primarily sensitive to the local oscillations of the surface binding site (which are shifted in frequency due to the presence of the adsorbate). Many published models for reactive scattering on solid surfaces will describe solid vibrations using only a single, harmonically bound surface atom [19, 27, 82–84]. These results explain why such a method is successful for strongly coupled species.

Despite this large shift in the spectral density when varying ω_{as}, in Figure 3.2C we demonstrate that the phonon density of states is nearly identical to that of a bare surface. The only significant change in $\rho(\omega)$ is the presence of the aforementioned surface site local mode present around $\omega = 210$ cm^{-1}. These results make physical sense given that the low frequency acoustic modes of a solid should be unaffected by gaseous species, especially at low pressures and surface coverages.

To evaluate the effect of the solid force field and structure on these results, we have tested several different crystal structures, facets, elemental compositions, and force fields. While the quantitative properties of the memory kernel vary across different systems, the qualitative dependence of the memory kernel on ω_{as} is quite general. In Figure 3.3 we present the memory kernel and spectral density for CO adsorbed atop Ru(0001) sites, where interactions between Ruthenium atoms were calculated using an embedded atom method (EAM) forcefield [85]. Since Ru has a much higher Debye Frequency than Pt (225 cm^{-1} for Ru and 156 cm^{-1} for Pt) the dependence of the spectral density on ω_{as} is shifted, but the modes associated with the primary peaks in the physisorbed and chemisorbed regimes remains the same.

This independence with respect to the details of the atomistic model is appeal-

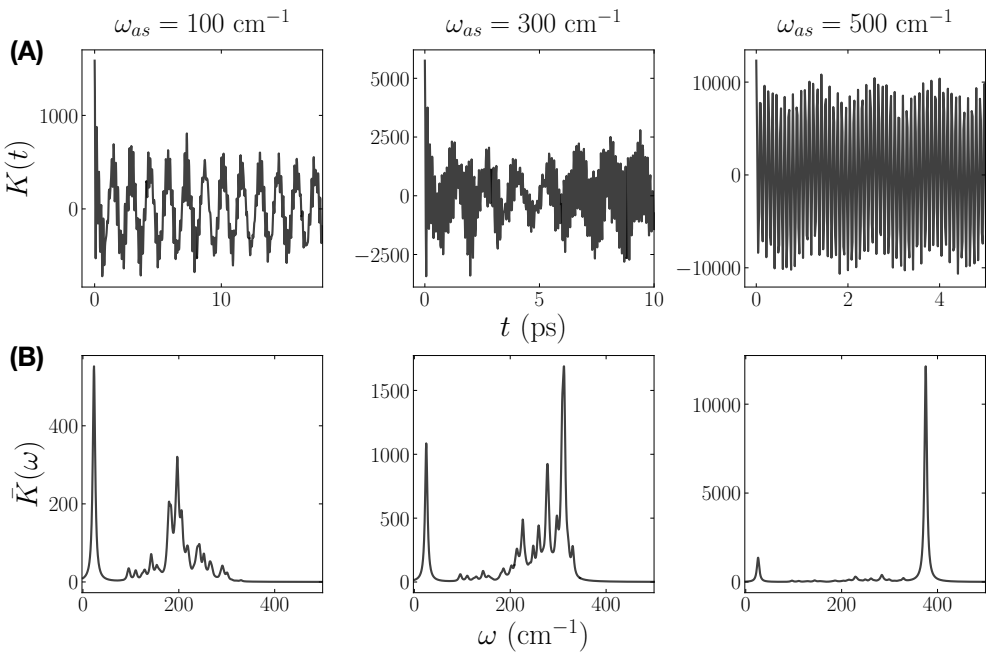

Figure 3.3: (A) Memory kernel and (B) spectral density for Ru(0001) surface.

ing and can be understood from a perturbative perspective. Specifically, in the physisorbed limit, the motion of the surface binding site is a small perturbation to the bulk phonon modes, while in the chemisorbed limit, the bulk phonon modes are a small perturbation to the motion of the surface binding sites. In the next section, we rigorously examine this statement by comparing perturbative schemes to the exact results for $K(t)$ and $\bar{K}(\omega)$.

Finally, we note that in this section we have only studied the effects of phonons for an ideal clean surface. Presumably, sources of surface heterogeneity, such as steps and defects, may add further richness and complexity to the picture we have provided here. We also note that we have only studied elemental solids, and the optical modes of polyatomic crystals could add another interesting dimension to this physical picture.

3.3 Perturbation Theory

In this section, we compare the results for the memory kernels and spectral densities calculated using exact diagonalization to two different perturbative schemes: one

which agrees well with exact results in the physisorbed limit (when ω_{as} is small), and the other which agrees well with exact results when in the chemisorbed limit (when ω_{as} is large).

The formulas for the perturbative corrections are the familiar Rayleigh-Schrodinger perturbation theory equations. In this section, we will only use 1st and 2nd order corrections to the eigenvalues and 1st order corrections to the eigenvectors. We provide the formulas below for completeness. For the eigenvalues we have,

$$\lambda_i \approx \lambda_i^{(0)} + \lambda_i^{(1)} + \lambda_i^{(2)} \tag{3.3}$$

where λ_i is equal to the square phonon frequency ω_i^2 and $\lambda_i^{(0)}$, $\lambda_i^{(1)}$, and $\lambda_i^{(2)}$ are the 0th, 1st, and 2nd order corrections respectively.

$$\lambda_i^{(1)} = \mathbf{P}_i^{(0)} \cdot \delta\mathbf{H} \cdot \mathbf{P}_i^{(0)}, \tag{3.4}$$

where $\mathbf{P}_i^{(0)}$ is the ith eigenvector of the unperturbed Hessian $\mathbf{H}_0$ and $\delta\mathbf{H}$ is the perturbation term.

$$\lambda_i^{(2)} = \sum_{j \neq i} \frac{\left(\mathbf{P}_j^{(0)} \cdot \delta\mathbf{H} \cdot \mathbf{P}_i^{(0)}\right)^2}{\lambda_i^{(0)} - \lambda_j^{(0)}}. \tag{3.5}$$

For the eigenvectors we have,

$$\mathbf{P}_i \approx \mathbf{P}_i^{(0)} + \mathbf{P}_i^{(1)}, \tag{3.6}$$

where,

$$\mathbf{P}_i^{(1)} = \sum_{j \neq i} \frac{\mathbf{P}_j^{(0)} \cdot \delta\mathbf{H} \cdot \mathbf{P}_i^{(0)}}{\lambda_i^{(0)} - \lambda_j^{(0)}} \mathbf{P}_j^{(0)}. \tag{3.7}$$

In order to gain insight from perturbation theory, we must judiciously choose how to separate the solid's mass-weighted Hessian (denoted as $\mathbf{H}_{\text{B}}$) into the reference Hessian $\mathbf{H}_0$ and the perturbation $\delta\mathbf{H}$. In order to do so, we first structure $\mathbf{H}_{\text{B}}$ into blocks corresponding to the surface adsorption site(s) $\mathbf{H}_{\text{X}}$, the remaining bulk atoms

$\mathbf{H}_{\mathrm{Y}}$, and off-diagonal blocks between the two $\mathbf{H}_{\mathrm{XY}}$,

$$\mathbf{H}_{\mathrm{B}} = \left(\begin{array}{c|c} \mathbf{H}_{\mathrm{X}} & \mathbf{H}_{\mathrm{XY}} \\ \hline \mathbf{H}_{\mathrm{XY}}^{T} & \mathbf{H}_{\mathrm{XY}} \end{array}\right), \tag{3.8}$$

Note that the size of $\mathbf{H}_{\mathrm{Y}}$ should be much larger than $\mathbf{H}_{\mathrm{X}}$ as most atoms in the solid can be treated as not interacting with the adsorbate. We may diagonalize the $\mathbf{H}_{\mathrm{Y}}$ to arrive the following form,

$$\mathbf{H}_{\mathrm{B}} = \left(\begin{array}{c|c} \mathbf{H}_{\mathrm{X}} & \mathbf{G} \\ \hline \mathbf{G}^{T} & \boldsymbol{\Omega}_{\mathrm{Y}}^{2} \end{array}\right), \tag{3.9}$$

where $\mathbf{G}$ is the coupling between the adsorption site(s) and each bulk phonon mode and $\boldsymbol{\Omega}_{\mathrm{Y}}^{2}$ is a diagonal matrix containing the square frequencies of these bulk modes. With this setup we are ready to perform perturbation theory. We will illustrate the results of perturbation theory on a 4x4x4 Pt(111) lattice with the Hessian evaluated using the EMT[72] forcefield.

3.3.1 Strong coupling/chemisorbed limit

In the chemisorbed limit, we set the reference Hessian to where the adsorption site(s) and the phonon modes of the bulk are uncoupled,

$$\mathbf{H}_{0} = \left(\begin{array}{c|c} \mathbf{H}_{\mathrm{X}} & 0 \\ \hline 0 & \boldsymbol{\Omega}_{\mathrm{Y}}^{2} \end{array}\right), \tag{3.10}$$

and therefore the perturbation is coupling,

$$\delta\mathbf{H} = \left(\begin{array}{c|c} 0 & \mathbf{G} \\ \hline \mathbf{G}^T & 0 \end{array}\right). \tag{3.11}$$

If we assume adsorption site to be a single atom, then the following analytical forms can be found for the memory kernel to 0th, 1st, and 2nd order respectively:

$$K^{(0)}(t) = \frac{\mu^2}{mM}\frac{\omega_{\text{as}}^4}{\tilde{\omega}_s^2}\cos(\tilde{\omega}_s t), \tag{3.12}$$

$$K^{(1)}(t) = K^{(0)}(t) + \frac{\mu^2}{mM}\omega_{\text{as}}^4 \sum_{i\neq s} \frac{g_i^2}{(\tilde{\omega}_s^2 - \omega_{\text{Y},i}^2)^2\omega_{\text{Y},i}^2}\cos(\omega_{\text{Y},i}t), \tag{3.13}$$

$$K^{(2)}(t) = K^{(0)}(t) + \frac{\mu^2}{mM}\omega_{\text{as}}^4 \sum_{i\neq s} \frac{g_i^2}{(\tilde{\omega}_s^2 - \omega_{\text{Y},i}^2)^2\omega_{\text{Y},i}^2 + (\tilde{\omega}_s^2 - \omega_{\text{Y},i}^2)g_i^2}\cos\left(\sqrt{\omega_{\text{Y},i}^2 + \frac{g_i^2}{\tilde{\omega}_s^2 - \omega_{\text{Y},i}^2}}t\right), \tag{3.14}$$

where g_i is the ith element $\mathbf{G}$, $\omega_{i,\text{Y}}$ is ith element of $\mathbf{\Omega}_\text{Y}$, and $\tilde{\omega}_s$ is the frequency of motion of the adsorption site with the adsorbate bound,

$$\tilde{\omega}_s = \sqrt{\frac{\mu}{M}\omega_{\text{as}}^2 + \omega_s^2}. \tag{3.15}$$

In this scheme, we see that to 0th order the surface adsorption site does not interact with the other modes of the solid, which results in the phonon memory kernel being a single sinusoid with frequency $\bar{\omega}_s$ due to solely the motion of the surface adsorption site. The perturbation introduces coupling between the surface adsorption site(s) and the bulk solid, allowing for the bulk phonons to contribute to the memory kernel. Figure 3.4 illustrates the spectral density calculated using this scheme to first and second order, and compares it to the results from exact diagonalization. The 1st

order results qualitatively match the exact results, however they underestimate the frequency of the adsorption-site local mode. This mismatch is because the first order corrections to the eigenvalues of the Hessian are zero. Introducing second order corrections removes this discrepancy, leading to excellent agreement with the exact results when $\bar{\omega}_s > \omega_{\mathrm{D}}$. As the effective frequency of motion of the adsorption site approaches ω_{D}, this perturbative scheme qualitatively fails to describe the memory kernel/spectral density.

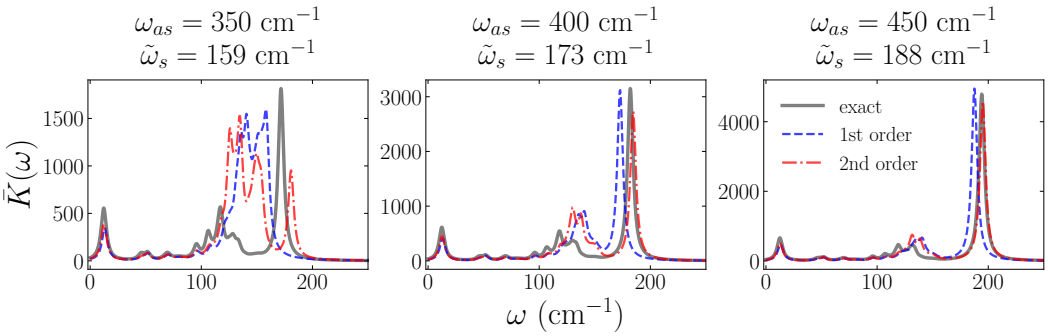

Figure 3.4: Spectral density for Pt(111) surface calculated using exact diagonalization, 1st order, and 2nd order perturbation theory in the strong-coupling scheme.

3.3.2 Weak-coupling/physisorbed limit

In the physisorbed limit, we treat $\mathbf{H}_{\mathrm{X}}$ as the perturbation, $\delta\mathbf{H} = \mathbf{H}_{\mathrm{X}}$, and the remaining Hessian as the reference,

$$\mathbf{H}_0 = \left(\begin{array}{c|c} 0 & \mathbf{G} \\ \hline \mathbf{G}^T & \boldsymbol{\Omega}_{\mathrm{Y}}^2 \end{array}\right). \tag{3.16}$$

This scheme assumes that the contribution of the adsorption sites to the bulk phonon modes is small. Figure 3.5 illustrates the spectral density calculated using this scheme, to first and second order, and compares it to the results from exact diagonalization. The 1st order results for $K(\omega)$ use 1st order corrections for both the eigenvectors and eigenvalues, while the 2nd order results add 2nd order corrections to the eigenvalues

while keeping the eigenvectors at 1st order. Both 1st and 2nd order results agree well with the exact results when ω_{as} is less than platinum's Debye frequency $\omega_{\text{D}} = 156$ cm^{-1} as expected, and even qualitatively capture results at $\omega_{\text{as}} \approx \omega_{\text{D}}$. However, as the effective frequency of motion of the adsorption site, approaches ω_{D}, this weak-coupling perturbative scheme fails.

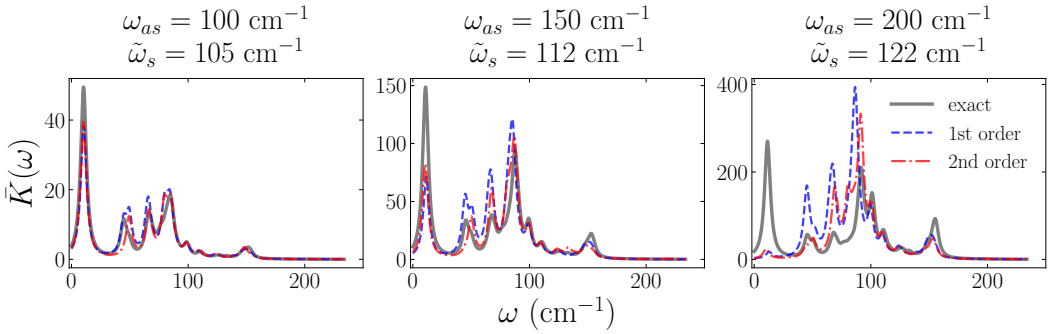

Figure 3.5: Spectral density for Pt(111) surface calculated using exact diagonalization, 1st order, and 2nd order perturbation theory in the weak-coupling scheme.

The success of these perturbative schemes at reproducing the exact results within the appropriate regime suggests that the behavior we have observed is not a feature of a particular atomistic model, but a fundamental consequence of the response of phonon modes in different physical regimes.

3.4 Dispersion

The dependence of the acoustic peak frequency on simulation size, such as illustrated in Figure 3.2, is an artifact that has the potential to prevent straightforward comparison between experiment and simulation. With or without periodic boundary conditions, the size of the simulated solid limits the maximum wavelength of the phonon modes. In this section, we illustrate how to generalize the theory presented in Chapter 2 by computing the memory kernel in the limit of an infinite surface through integration of frequencies and eigenmodes across the first Brillouin zone.

In an infinite crystalline solid, displacement by a lattice vector returns the same solid. This symmetry can be leveraged to calculate the phonon frequencies and dis-

placements of the bulk solid using a spatial Fourier transform of the mass-weighted Hessian. This approach leads to the well-known Bloch's theorem, which is summarized as follows. Let a and b be the indices of two primitive unit cells, and let $\mathbf{H}_S(a,b)$ be the mass-weighted Hessian of the crystal, $H_{S;ij}(a,b) = \frac{\partial^2 V}{\partial x_{S,i}(a)\partial x_{S,j}(b)}$. The Fourier transform of this matrix may be expressed as,

$$\mathbf{D}(\mathbf{k}) = \sum_{a,b} \mathbf{H_S}(a,b) e^{i\mathbf{k}\cdot(\mathbf{r}_a - \mathbf{r}_b)}, \tag{3.17}$$

where $\mathbf{r}_a$ is the origin of the ath cell. $\mathbf{D}(\mathbf{k})$ is known as the *dynamical matrix* satisfying,

$$\mathbf{D}(\mathbf{k})\mathbf{U}_j(\mathbf{k}) = -\omega_j^2(\mathbf{k})\mathbf{U}_j(\mathbf{k}). \tag{3.18}$$

where ω_j is the jth phonon band frequency and $\mathbf{U}_j$ is the corresponding polarization vector. For 3D monatomic crystals, the primitive (Wigner-Seitz) cell consists of a single atom with three degrees of freedom, thus producing three phonon bands. These bands, illustrated in Figure3.6A for FCC platinum, are the acoustic transverse and longitudinal modes, and are characterized by linear dispersion at low wavenumber.

In a surface slab, the presence of anisotropy breaks the symmetry of the 3D crystal. This symmetry breaking leads to additional surface mode bands, which can be acoustic (Rayleigh waves) or non-acoustic in character[86–88]. In Figure 3.6B and 3.6C we demonstrate the phonon dispersion of a Pt(111) surface calculated via EMT. These results were calculated using a 4x4x8 surface replicated in a 6x6 super-cell. We verified convergence with respect to supercell size. Of course, the number of surface phonon bands depends on the size of the surface unit cell used in computing $\mathbf{D}(\mathbf{k})$, however, we verified that the results are qualitatively similar across different unit cell sizes.

Figure 3.6B illustrates the phonon dispersion for a bare surface, while Figure 3.6C illustrates the dispersion for a surface with an adsorbed CO. The bands are colored based on their mean slope between the Γ and K high-symmetry points. The three lowest frequency bands are all surface acoustic modes, for example, the flexing mode depicted in the inset of Figure 3.1, confirming that the corresponding peak in

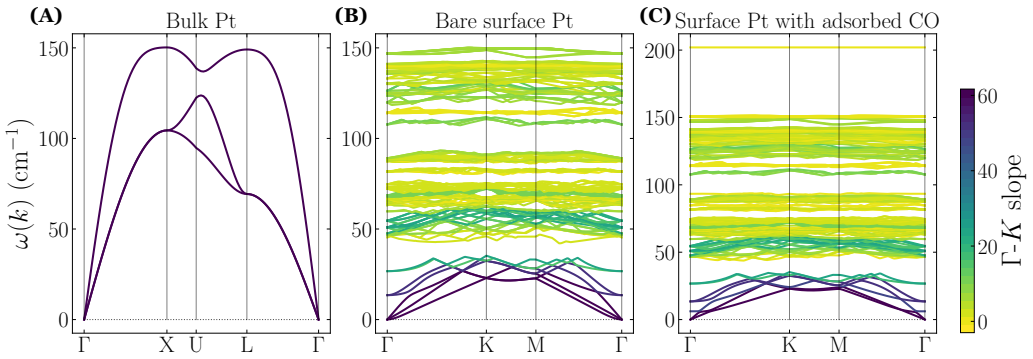

Figure 3.6: Phonon dispersion curves. (A) Bulk Pt dispersion curves calculated using an EMT forcefield and a 10x10x10 atom supercell. (B) Dispersion curves for a 4x4x8 atom surface slab replicated in a 6x6 surface cell (C) Same as middle but with an adsorbed CO molecule corresponding to $\omega_{\text{as}} = 480$ cm^{-1}.

the phonon spectral density in Figure 3.1 arises from an acoustic phonon. The remaining modes are nearly dispersionless — especially the highest frequency mode in Figure 3.6C, which corresponds to the surface site local vibration. The dispersionless nature of this mode supports why it was not seen to be dependent on surface slab size in Figure 3.2.

Using these dispersion relations, we can average the memory kernel across the first Brillouin zone,

$$K(t) = \frac{\mu^2}{mM}\omega_{\text{as}}^4 \sum_j {\sum_{\mathbf{k}}}' \frac{|U_{sj}(\mathbf{k})|^2}{\omega_j(\mathbf{k})^2} \cos(\omega_j(\mathbf{k})t), \tag{3.19}$$

$$\bar{K}(\omega) = \frac{\mu^2}{mM}\omega_{\text{as}}^4 \sum_j {\sum_{\mathbf{k}}}' \frac{|U_{sj}(\mathbf{k})|^2}{\omega_j(\mathbf{k})^2} \delta(\omega - \omega_j(\mathbf{k})), \tag{3.20}$$

where the primed summation is taken over all wavevectors in the first Brillouin zone and the sum over j is taken over all surface unit cell phonon modes. Equations 3.19 and 3.20 simply separate the sum in Eqs. 3.1 and 3.2 into two parts: the outer sum varying intra-cell displacements and the inner sum varying inter-cell displacements. In principle, this procedure eliminates the dependence on boundary conditions and generalizes results from a surface unit cell to an infinite periodic surface. However, the depth of the surface (corresponding to the non-periodic dimension) is still limited to the depth of the surface unit cell.

Figure 3.7 demonstrates the results of Eq. 3.19 and Eq. 3.20 and compares them

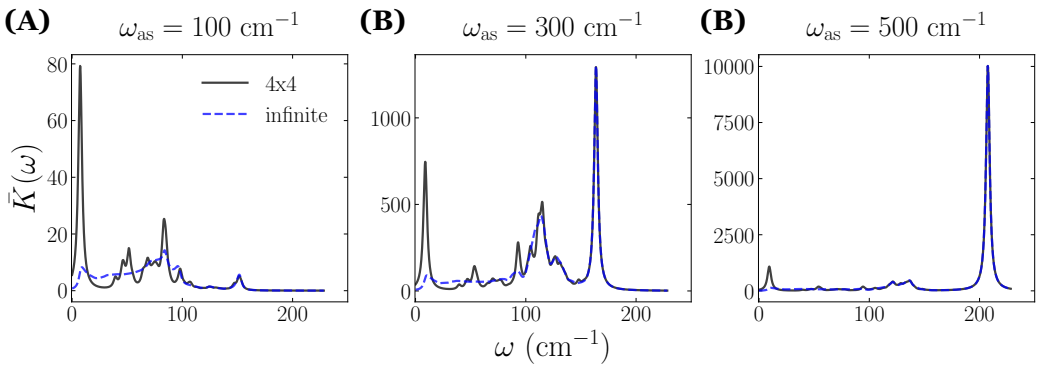

Figure 3.7: Memory kernel for a model adsorbate on Pt(111) in a finite system (black) and in the infinite system size limit (blue). The finite system is equivalent to the 4x4x8 results presented in Figure 1.

to previous results for a single unit cell slab. Naturally, the strong coupling of the adsorbate to a single acoustic mode is broadened, resulting in a much flatter spectral density in the low frequency ranges. Such a flat spectral density is characteristic of Markovian (white) noise. The high-frequency range of the spectral density is largely unaltered, due to the dispersionless nature of the high frequency modes. In the Appendix C, we discuss how a flat spectral density for acoustic phonons is consistent with predictions from continuum elastic theory.

3.5 Anharmonic effects

The memory kernels depicted in Figure 3.1 have finite correlations even as $t \to \infty$. Such persistent correlations are a hallmark of non-ergodicity, suggesting that the adsorbate will not relax to thermal equilibrium. Indeed, all finite harmonic baths are non-ergodic. However, even a very small amount of anharmonicity can break the symmetry of the harmonic motion and cause the memory kernel to decay as $t \to \infty$. Physically, such a decay is equivalent to heat-dissipation in the lattice due to phonon-phonon interactions. In this section, we explore how anharmonicity affects the memory by calculating the memory kernel using time-correlation functions computed with molecular dynamics (MD) simulation. Methods for calculating memory kernels using MD simulation are discussed in Appendix B. Here we use the same EMT model

as in Section 3.2 for the surface motion, and model the surface-adsorbate interaction as a Morse potential,

$$V(z) = D(1 - e^{-\alpha(z-z_0)^2}) \tag{3.21}$$

where the well-depth D is 0.77 eV, and the equilibrium distance z_0 is 2.5 Å, and the well-width α is treated as a tunable parameter that determines the adsorbate-surface frequency ω_{as}.

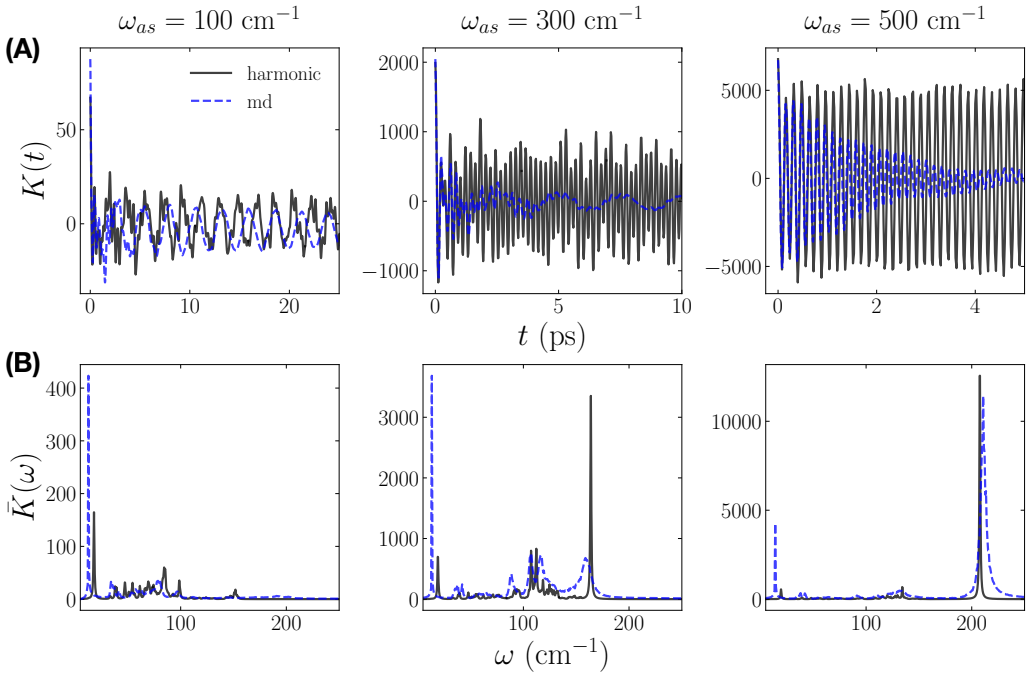

Figure 3.8: (A) Memory kernel and (B) spectral density for a 4x4x4 Pt surface calculated via Eqs. 3.1 and 3.2 versus with time-correlation functions calculated with MD simulations (Eq. B.1).

Figure 3.8 compares memory kernels and spectral densities calculated from diagonalizing the Hessian (harmonic approximation) versus those directly from MD, which account for anharmonicities. We see that the anharmonicities in the potential, though very small in magnitude, broaden the spectral density and cause the memory to decay asymptotically, as expected. We also note that anharmonicity causes a small shift in the frequency of the peaks in the spectral density as well, however, such phonon-softening effects are also expected and a consequence of anharmonicity affecting the dispersion relations [89].

Besides memory decay and dissipation, anharmonicity (specifically in the adsorbate-surface interactions) is also responsible for allowing adsorbates to diffuse on the surface of the metal. The harmonic approximation made in arriving at Eq. 3.1 and Eq. 3.2 forces the adsorbate to remain bound to its adsorption site for eternity. However, in reality, adsorbates diffuse along the surface, and even chemisorbed species will hop between surface sites at a characteristic rate[90]. In order to model the influence of surface diffusion on the memory kernel, we stochastically varied the surface site the adsorbate was bound to over the course of an MD simulation. Waiting times were sampled from an exponential distribution, such that the adsorbate would hop at a certain average rate k_{h}. Results from these simulations are shown in Figure 3.9.

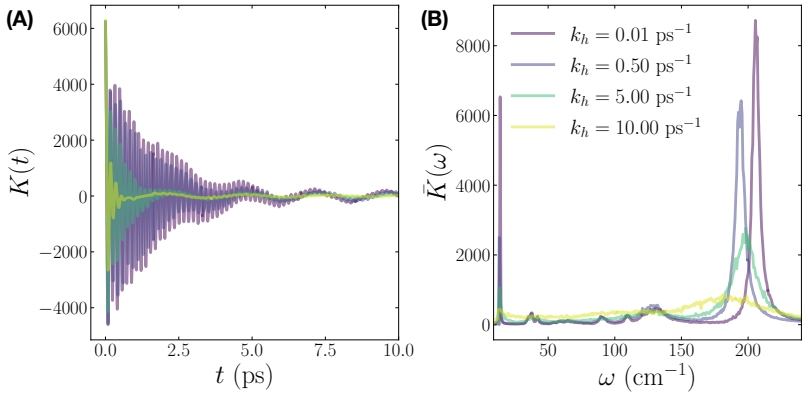

Figure 3.9: Dependence of (A) memory kernel and (B) spectral density on hopping rate with $\omega_{\mathrm{as}} = 480$ cm^{-1}.

In Figure 3.9, we see that as the hopping rate increases, the spectral density flattens, becoming more structureless and Markovian. Essentially, as the rate of hopping becomes faster than the frequency of the phonon modes, the modes do not have time to respond/relax before the adsorbate hops to another site. This mixing of timescales causes the frequency information of the phonon bath to be lost. On the other hand, when k_{h} is much larger than the phonon frequencies, the spectral density looks nearly identical to that of a particle fixed to a single site. For reference, k_{h} for chemisorbed species like CO k_{h} is typically on the order of nanoseconds to microseconds[90], which is much larger than the frequency of the local phonon mode that dominates the contribution to the memory kernel. These observations imply

that chemisorbed species form localized vibrational distortions that travel with them as they move along the surface. Such behavior is highly reminiscent of a *polaron*, more specifically, a surface polaron.

3.6 Ramifications for reaction rates

The most salient motivation for studying the effects of phonons on molecular adsorbates is to derive how phonons affect rate constants of catalytic reactions. Here we derive phononic corrections to rate constants of reaction near equilibrium using harmonic transition-state theory (Eq. 2.35), noting once again that such a method is essentially equivalent to Kramers-Grote-Hynes (KGH) theory, however in KGH it is assumed that the coupling between the system and adsorbate remains the same in the reactant and transition states, while in harmonic transition state we may relax that assumption. We begin by deriving a formula for the phonon corrections in general, and then specialize to the case of desorption rates and compare the results of our model to experimental temperature dependent rate constants for CO and Xe desorption from Pt(111).

Following the formalism Section 2.3, using harmonic transition state theory we can express the rate constant for a reaction at a solid surface as in terms of the total molecular + solid Hessian,

$$k = \frac{\lambda^{\ddagger}}{2\pi}\sqrt{\frac{\det(\mathbf{H}_{\mathrm{R}})}{\det(\mathbf{H}^{\ddagger})}}e^{-\beta(E^{\ddagger}-E_R)}, \tag{3.22}$$

where $\mathbf{H}_{\mathrm{R}}$ is the mass-weighted Hessian expanded around the reactant state, $\mathbf{H}^{\ddagger}$ is the mass-weighted Hessian expanded around the transition state, $\lambda^{\ddagger}$ is the frequency of the unstable mode of $\mathbf{H}^{\ddagger}$, and "det" denotes the matrix determinant. These Hessians can be organized into a block structure corresponding to the molecular/adsorbate

degrees of freedom, the solid degrees of freedom, and the coupling between them,

$$\mathbf{H} = \left(\begin{array}{c|c} \mathbf{H}_\mathrm{A} & \mathbf{C} \\ \hline \mathbf{C}^T & \mathbf{H}_\mathrm{B} \end{array}\right). \tag{3.23}$$

The determinant of such a block matrix may be evaluated as,

$$\det(\mathbf{H}) = \det(\mathbf{H}_\mathrm{B}) \times \det(\mathbf{H}_\mathrm{A} - \mathbf{C}^T\mathbf{H}_A^{-1}\mathbf{C}). \tag{3.24}$$

Using this determinant identity together with Eq. 3.22 yields,

$$k = \frac{1}{2\pi} \times \prod_{i=0}^{N_S-1} \frac{\omega_i}{\omega_i^\ddagger} \times \frac{\prod_{i=0}^{N_A-1} \tilde{f}_i}{\prod_{i=1}^{N_A-1} \tilde{f}_i^\ddagger} \times e^{-\beta\Delta E^\ddagger}, \tag{3.25}$$

where N_S and N_A are the number of solid and adsorbate degrees of freedom respectively ($N_S + N_A = N$)), ω_i and $\omega_i^\ddagger$ are the phonon frequencies in the reactant and transition state respectively, and $\tilde{f}_i$ and $\tilde{f}_i^\ddagger$ are effective molecular frequencies in the reactant and transition state respectively. Let $\mathbf{H}_A$ be the mass-weighted Hessian of the adsorbate degrees of freedom, then $\tilde{f}_i$ and $\tilde{f}_i^\ddagger$ are the eigenfrequencies of an effective molecular Hessian,

$$\tilde{\mathbf{H}}_A = \mathbf{H}_A - \mathbf{C}\boldsymbol{\Omega}^{-2}\mathbf{C}^T, \tag{3.26}$$

where the shift term is equal to the instantaneous ($t = 0$) component of the memory kernel, $\mathbf{K}(t = 0) = \mathbf{C}\boldsymbol{\Omega}^{-2}\mathbf{C}^T$.

From Eq. 3.25, we see that phonons induce two correction factors to the rate constant: the ratio of solid phonon frequencies between the reactant and transition state, and the shift in molecular frequencies both in the reactant and transition state. The ratio of phonon frequencies is largest when the reactant molecules are strongly coupled

to the surface and the transition state is not. In Figure 3.1C we demonstrated even a molecule strongly coupled to a surface affects only the highest frequency phonon mode, leaving the bulk of the phonon density of states unchanged. Thus, we can approximate the ratio of phonon frequencies as,

$$\prod_{i=0}^{N_S-1} \frac{\omega_i}{\omega_i^\ddagger} \approx \frac{\tilde{\omega}_\mathrm{D}}{\omega_\mathrm{D}}, \tag{3.27}$$

where $\tilde{\omega}_\mathrm{D}$ is the highest frequency phonon mode when the reactants are bound to the surface and ω_D is the bare solid Debye frequency. For physisorbed species $\tilde{\omega}_\mathrm{D}$ and ω_D are the same, however, for chemisorbed species the in the adsorbate frequencies is a thermodynamic correction arising from phonons altering the free energy surface along the reaction coordinate. Indeed, if we denote $\tilde{f}_0$ as the adsorbate normal mode along the reaction coordinate Eq. 3.25 can be simplified to,

$$k = \frac{\tilde{f}_0}{2\pi} \times \prod_{i=0}^{N_S-1} \frac{\omega_i}{\omega_i^\ddagger} \times e^{-\beta(\Delta E^\ddagger + T\Delta \tilde{S}^\ddagger)}, \tag{3.28}$$

where $\Delta \tilde{S}^\ddagger$ is the effective barrier entropy,

$$\Delta \tilde{S}^\ddagger = k_\mathrm{B} \sum_{i=0}^{N_A-1} \ln\left(\frac{\tilde{f}_i^\ddagger}{\tilde{f}_i}\right). \tag{3.29}$$

In the harmonic approximation this barrier entropy is independent of temperature, but in a more general context it can be shown to be temperature dependent.

3.6.1 Surface desorption

The formalism described in the previous subsection can be applied to derive a rate constant for surface desorption. For gas-phase desorption, the reaction coordinate can be defined as the distance of the adsorbate center of mass from the surface. Furthermore, in many cases, the desorption process is *barrierless*, meaning that at the transition state the molecule and surface do not interact [48]. Thus, we can adapt

Eq. 3.27 to express the gas-phase desorption rate constant,

$$k_d = \frac{\tilde{\omega}_{\text{as}}}{2\pi} \times \frac{\tilde{\omega}_{\text{D}}}{\omega_{\text{D}}} \times e^{-\beta \Delta E^{\ddagger}}, \tag{3.30}$$

where $\tilde{\omega}_{\text{as}}$ is the effective adsorbate-surface interaction frequency satisfying,

$$\tilde{\omega}_{\text{as}} = \sqrt{\frac{\mu}{m}\omega_{\text{as}}^2 - K(t=0)}. \tag{3.31}$$

While transition state theory is primarily a classical theory [63, 91–93], studies have demonstrated improved agreement with experiment when introducing quantum corrections, such as accounting for the rotational motion of the molecule or using quantum harmonic oscillator partition functions instead of classical oscillator partition functions [48, 90]. We will thus compare four different models for the desorption rate constant to experimental results: (1) a fixed-surface model using classical harmonic oscillator partition functions and a rotation correction,

$$k_{d1} = \frac{\omega_{\text{as}}}{2\pi} \times \frac{2I}{\hbar^2 \beta} \times e^{-\beta \Delta E^{\ddagger}}, \tag{3.32}$$

where I denotes the moment of inertia of the adsorbate. (2) a phonon-corrected model using classical harmonic oscillator partition functions and a rotational correction,

$$k_{d2} = \frac{\tilde{\omega}_{\text{as}}}{2\pi} \times \frac{\tilde{\omega}_{\text{D}}}{\omega_{\text{D}}} \times \frac{2I}{\hbar^2 \beta} \times e^{-\beta \Delta E^{\ddagger}}, \tag{3.33}$$

(3) a fixed-surface model using quantum harmonic oscillator partition functions and a rotational correction,

$$k_{d3} = \frac{1 - e^{-\beta \hbar \omega_{\text{as}}}}{2\pi \beta \hbar} \times \frac{2I}{\hbar^2 \beta} \times e^{-\beta \Delta E^{\ddagger}}, \tag{3.34}$$

and (4) a phonon-corrected model using quantum harmonic oscillator partition functions and a rotational correction,

$$k_{d4} = \frac{1 - e^{-\beta \hbar \tilde{\omega}_{\text{as}}}}{2\pi \beta \hbar} \times \frac{1 - e^{-\beta \hbar \tilde{\omega}_{\text{D}}}}{1 - e^{-\beta \hbar \omega_{D}}} \times \frac{2I}{\hbar^2 \beta} \times e^{-\beta \Delta E^{\ddagger}}. \tag{3.35}$$

Note that by "fixed-surface" we do not mean a surface at absolute zero, but rather a surface that acts as an ideal, structureless, thermal environment.

3.6.2 Comparison to experiment

We have compared results from Eqs. 3.32 to 3.35 to experimental temperature-dependent desorption rate constants for CO and Xe from a Pt(111) surface. Parameters used in computing Eqs. 3.32 to 3.35 are presented in Table 3.1. Note that because Xe is an atomic species, its moment of inertia is 0, and thus we ignore the rotational partition function factors for Xe calculations. The phonon corrections in Eq. 3.33 and Eq. 3.35 were computed using a 4x4x8 EMT surface slab and subsequently averaged across the first Brillouin zone to approximate an infinite surface, as described in Sec. 3.4. In the SI, we illustrate that the rate constant corrections we present here are not sensitive to the size of the surface slab used, or whether one accounts for phonon dispersion.

Table 3.1: Parameters used for computing desorption rate constants shown in Figure 3.11.

	$E^{\ddagger}$ (eV)	ω_{as} (cm^{-1})	$\tilde{\omega}_{\text{as}}$ (cm^{-1})	$\tilde{\omega}_{\text{D}}$ (cm^{-1})
CO	1.47[90]	480[70]	164	203
Xe	0.245[94]	28[79]	21	156

The CO desorption rate constants were taken from Ref. 90, and the Xe desorption rates were taken from Ref. 94. In Ref. 90, desorption rate constants were calculated by fitting the time-dependent flux from a beam scattering experiment to two models: a single exponential model and a bi-exponential model. The single exponential model fit the entire flux signal, mixing contributions from terrace and steps. Meanwhile the bi-exponential model separated the flux into a fast component, arising from terrace desorption, and a slow component, arising from step to terrace diffusion followed by terrace desorption. While our TST calculations do not include the role of steps, we compare the results of our models to data both from the single exponential model and the fast component of the bi-exponential for thoroughness and transparency.

In Figure 3.10A, we see that the phonon-corrected models (Eqs. 3.33 and 3.35)

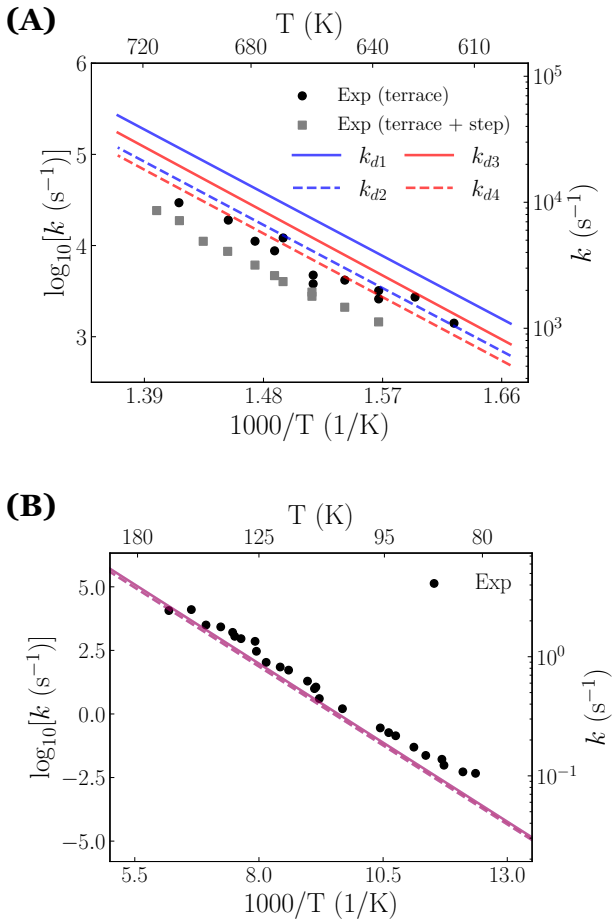

Figure 3.10: Rate constants for desorption from a Pt(111) surface. (A) CO desorption. Grey squares are experimental data which mixed contributions from both steps and terraces. Black circles refer experimental data where the kinetics of terrace desorption was isolated. (B) Xe desorption.

give improved agreement with experimental results for CO desorption. In particular, k_{d4} the quantum, flexible surface model and the terrace desorption rate constants yield the best agreement. The improved agreement when using Eqs. 3.33 to 3.35 versus Eqs. 3.32 to 3.34 arises from the reduced adsorbate-surface frequency $\tilde{\omega}_{\mathrm{as}}$. Physically, the flexible surface reduces the stiffness of the adsorbate-surface bond, leading to a lower frequency of attempts over the barrier and a lower rate prefactor.

In Figure 3.10B, we demonstrate results for Xe. Here, all the TST models lie essentially on top of each other and are lower in value than the experimental rate constants, although by a small margin. The smaller phonon corrections for Xe versus CO are a natural result of the weaker interactions with the surface. Weaker coupling

means a small $K(t=0)$, leading to $\tilde{\omega}_{\text{as}} \approx \omega_{\text{as}}$. Furthermore, weak coupling also results in a phonon density of states that is unchanged from a bare lattice, implying $\tilde{\omega}_{\text{D}} = \omega_{\text{D}}$. In general, the theory presented in this section suggests the phonon corrections to the rate constant are much larger for chemisorbed species than physisorbed species.

It is worth emphasizing that using slightly different values for the surface binding energy, $\Delta E^{\ddagger}$, can substantially shift the quality of agreement of theoretical calculations with experiment. The major impediment to the first principles calculation of chemical rates is still the calculation of the barrier energy, and the phonon corrections to the rate constants seem to be a comparatively minor factor, even for chemisorbed species. Indeed, the purpose of this section was not to demonstrate that the magnitude of phonon corrections to reaction rates is large, but rather to illustrate that the theoretical models we developed in Sections 2 and 3, when combined with transition state theory, produce physically interpretable results which correspond well with existing experimental measurements.

3.7 Phonon response functions

Many important experiments, such as the surface acoustic wave experiments outlined in Chapter 1, apply external driving forces to the surface and/or adsorbate, observing large increases in the reaction rates when such applied forces are at particular polarizations and/or frequencies[16, 17, 39]. The transition state theories we previously developed cannot be used to understand such experiments as they are only valid in ensembles near thermal equilibrium. In fact, understanding how reactions respond to time-dependent external forces remains one of the grand challenges of modern theoretical chemistry and non-equilibrium statistical mechanics[95, 96]. In this section, we analyze the effect that external driving applied to a thin platinum film has on the motion of the molecular adsorbate.

We begin with the explicit solid/bath equation of motion derived in Chapter 2

(Eq. 2.9), but now include an additional driving force term,

$$\ddot{\mathbf{x}}_{\mathrm{B}}(t) = -\mathbf{D}^2_{\mathrm{BA}}\mathbf{x}_{\mathrm{A}}(t) - \mathbf{D}^2_{\mathrm{BB}}\mathbf{x}_{\mathrm{B}}(t) + \mathbf{F}(t), \tag{3.36}$$

where $\mathbf{F}(t)$ is the external force acting on each surface atom. This equation may be integrated analytically and inserted back into the equation of motion for adsorbates yielding a GLE with an additional driving force term,

$$\ddot{\mathbf{x}}_{\mathrm{A}}(t) = -\frac{\partial V_{\mathrm{A}}}{\partial \mathbf{x}_{\mathrm{A}}}(t) - \left[\mathbf{D}^2_{\mathrm{AA}} - \mathbf{K}(t=0)\right]\mathbf{x}_{\mathrm{A}}(t) - \int_0^t d\tau \mathbf{K}(t-\tau)\dot{\mathbf{x}}_{\mathrm{A}}(t) + \mathbf{R}(t) + \int_0^t \boldsymbol{\chi}(t-\tau)f(t), \tag{3.37}$$

where $f(t)$ is the time-dependent part of the external force vector, $\mathbf{F}(t) = \mathbf{F}f(t)$, and the response function vector $\boldsymbol{\chi}$ is given by,

$$\boldsymbol{\chi} = \mathbf{C}\frac{\sin(\boldsymbol{\Omega}t)}{\boldsymbol{\Omega}}\mathbf{U}^T\mathbf{F}. \tag{3.38}$$

The definitions of $\mathbf{C}$, $\boldsymbol{\Omega}$ and $\mathbf{U}$ were given in Section 2.1. The response function $\boldsymbol{\chi}$ is nearly equivalent to the position memory kernel $\boldsymbol{\Theta}$ defined in Eq. 2.13, and indeed the two can be shown to be equal (up to a multiplicative constant) in the case where the external force couples to the surface atoms in the same manner as the adsorbate, $\mathbf{F} = \mathbf{D}_{BA}$. In most experimentally relevant cases however, external forces are not solely applied to the adsorption sites, but to the boundary atoms of the surface.

We study three different cases for applied force $\mathbf{F}$ and their concomitant effects on the response function for the surface-normal motion of an adsorbate with $\omega_{\mathrm{as}} = 480$ cm^{-1}. The adsorbate is bound atop a surface site in the center of a 25x25x10 platinum film. Schematics of cases studied are presented in Figure 3.11A. In the first case, an extensional (surface-normal) force is applied to the bottom-layer of the thin Pt film. In the second case a shear force is applied to the bottom-layer. In the final case, a spherically symmetric force of radius 0.5 nm is applied at the edge of the top-layer of the Pt film. The final case mimics many surface shockwave experiments [97]. Note that in all these cases the time-envelope $f(t)$ may be different, but this does not affect χ.

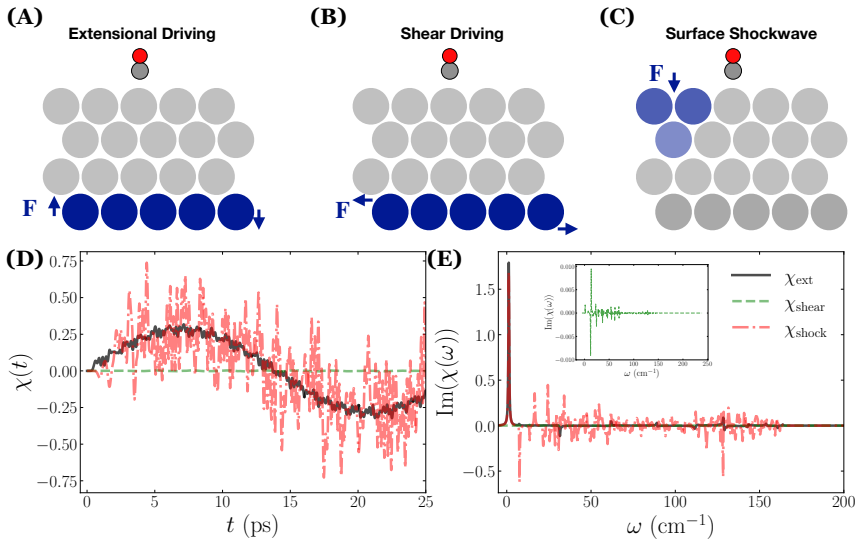

Figure 3.11: Schematics for (A) extensional forces, (B) shear forces, and (C) shockwave forces used for computing χ. (D) Time-domain response functions for extensional, shear, and shockwave forces. (F) Imaginary part of frequency-domain response function. Response functions were calculated using a 25x25x10 Pt film with an EMT forcefield[71].

Figure 3.11D and E analyze time-domain response function and the imaginary part of its Fourier transform, respectively. We see that the response of the adsorbate to the extensional and surface-shock forces is significantly larger (by about two orders of magnitude) than the response to the shear forces. This discrepancy in the response is because extensional and surface-shock forces couple songly to a z-polarized acoustic phonon that also couples strongly to the adsorbate, while the shear forces do not. Our results seem to correspond to observations made in Ref. [16] wherein similar extensional forces were observed to produce a rate enhancement for the oxidation of ethanol to acetaldehyde on a palladium catalyst, but shear forces were not. The strong coherence of the $\chi_{\mathrm{ext}}(t)$ suggests that a resonant force applied at the same frequency to the catalyst would cause $\int_0^t \chi(t-\tau)f(t)$ to grow over time, which might allow the adsorbate to cross barriers more easily. Interestingly, while the magnitude of χ_{shock} is similar χ_{ext}, the oscillations in χ_{shock} are much less coherent, negating any possible resonance effects.

3.8 Conclusions

In this chapter, we have developed a theory for how surface phonons couple to molecular adsorbates based on Mori-Zwanzig theory and the generalized Langevin equation. By integrating out the solid degrees of freedom (assuming they could be described harmonically), we derived a GLE for the adsorbate, wherein the memory kernel is merely a sum of the phonon frequency of the solid weighted by their coupling to the adsorbate. We demonstrated that this memory kernel depends sensitively on the frequency of the adsorbate-surface bond. When the frequency of this bond is smaller than the Debye frequency of the solid, adsorbates couple primarily to the acoustic phonons of the solid. When the frequency of the bond is much larger than the Debye frequency of the solid, adsorbates couple primarily to the dispersionless local vibrations of the adsorption site. Anharmonicity causes the memory to decay exponentially with time, but does not change the qualitative correlation between the frequency of the surface-adsorbate bond and the dominant frequency in the memory kernel. Subsequently, we used harmonic transition state theory to derive phononic corrections to reaction rate constants. We show that these corrections improved agreement between theory and experiment for CO desorption rates from Pt(111). Finally, we studied the phonon response functions, demonstrating that applied extensional forces couple far more strongly to adsorbates than applied shear forces.

Chapter 4

Phonon-induced memory on surface sites [1]

4.1 Introduction

Molecular beam scattering experiments are an invaluable tool for understanding the properties and mechanisms of surface reactions [19, 21–24, 26, 28, 29, 99–106]. In this chapter, we develop a generalized Langevin equation (GLE) based model suitable for the simulation of the effects of surface vibrations in such experiments. Rather than projecting out the entire solid to derive a GLE for the adsorbates degrees of freedom, our model projects out the bulk lattice degrees of freedom to derive a GLE for surface sites which are explicitly and anharmonically coupled to adsorbates. Much like the memory kernels for adsorbates, the memory kernels for surface sites also exhibit a strong coupling to particular acoustic phonon modes. We demonstrate that nanoscale boundary conditions, such as those of atomistic simulation, cause the acoustic modes to oscillate on chemical timescales. We show that this effect can lead to a systematic decrease in sticking coefficients in surface scattering simulations from the macroscopic limit. A GLE-based approach, such as we present here, can mitigate these unwanted finite-size effects without necessitating large and computationally expensive simulations.

[1] Based on work published in Ref. 98. Copyright 2023, American Chemical Society.

Using the same GLE model described in Chapter 3 for surface scattering simulations would be inappropriate due to the harmonic approximation for the adsorbate-surface interaction. At sufficiently large distances, the adsorbate and surface no longer interact, suggesting that the memory and random forces in a GLE for adsorbates should be spatially dependent and decay to zero as the adsorbate-surface distance increases. Indeed, the memory kernels presented in Section 3.5, which supposedly account for anharmonicity, only average over the true spatial dependence of the memory kernel. Capturing this spatial dependence of the memory is essential for calculating the energy dissipated due to vibrations over the course of a surface scattering trajectory. Unfortunately, the numerical parametrization and simulation of GLEs with spatially-dependent memory are quite expensive. A simpler and more elegant approach is to only project out the solid atoms that interact with adsorbates and treat all of the adsorbate and adsorption site degrees of freedom explicitly. Under such a scheme, the equations of motion become,

$$\ddot{\mathbf{x}}_{\mathrm{A}}(t) = -\frac{\partial V_{\mathrm{AS}}}{\partial \mathbf{x}_{\mathrm{A}}}, \tag{4.1}$$

$$\ddot{\mathbf{x}}_{\mathrm{S}}(t) = -\frac{\partial V_{\mathrm{AS}}}{\partial \mathbf{x}_{\mathrm{S}}} - \left[\omega_{\mathrm{S}}^2 - \mathbf{K}_{\mathrm{S}}(t=0)\right]\mathbf{x}_{\mathrm{S}}(t) - \int_0^t d\tau \mathbf{K}_{\mathrm{S}}(t-\tau)\dot{\mathbf{x}}_{\mathrm{S}}(t) + \mathbf{R}_{\mathrm{S}}(t), \tag{4.2}$$

where $\mathbf{x}_{\mathrm{A}}$ and $\mathbf{x}_{\mathrm{S}}$ are the mass-weighted displacements of adsorbates and adsorption sites respectively, V_{AS} is the adsorbate-surface interaction, ω_{S} is the local oscillation frequency of adsorption sites, and $\mathbf{K}_{\mathrm{S}}$ and $\mathbf{R}_{\mathrm{S}}$ are the memory and random forces respectively arising from the influence of bulk lattice motion on the adsorption sites. We use the subscript roman S to differentiate this memory kernel from the memory kernel acting on adsorbates discussed in Chapter 3.

Adelman and Doll[45] were the first to discuss in depth how the GLE could be used to model the effect of substrate phonons on a site within a harmonic lattice and the concomitant effects on gas-surface scattering. Later, Tully developed the generalized Langevin oscillator (GLO) method, wherein the motion of a surface site is described via a GLE with a memory kernel that is given by a single exponentially

damped sinusoid [46]. As we discuss in Section 4.2.1, such a memory kernel is equivalent to coupling the surface atom to a single dissipative (ghost) oscillator. Tully's GLO method has seen much success as a highly computationally efficient way of modeling surface dynamics, particularly in application to molecular beam scattering experiments [27, 83, 84, 107, 108]. However, modeling the dynamics of the surface with a single mode is a limiting approximation; in principle, surface atoms should couple to each normal mode of the lattice.

In this chapter, we extend Tully's GLO model to allow for a memory kernel of arbitrary complexity, and examine when and how the properties of the memory kernel affect molecular adsorption and surface scattering. We call this model the lattice generalized Langevin equation (LGLE). In order to parameterize the memory kernel, we use data taken from atomistic simulations. Crucially, in Section 4.2 we show that the qualitative properties of the memory kernel are independent of the atomistic model details, such as forcefield parameters.

The remainder of this chapter is organized as follows. In Section 4.2, we detail our simulation methods and review the extended variable transformation used to map the non-Markovian dynamics to a bath of dissipative harmonic oscillators. In Section 4.3, we analyze memory kernels derived from these simulations, varying the choice of force fields, elemental composition, solvation state, and surface-slab size. Finally, in Section 4.4 we discuss how the properties of the memory kernel generated via our approach affect surface scattering, highlighting systematic errors that occur in sticking coefficients when basing the memory kernel on data from small periodic systems. Much of the content of this chapter is adapted from Ref. 98.

4.2 Methods

4.2.1 Extended variable transformation

Computing the integral over the memory kernel in Eq. 4.2 is computationally intensive, especially for systems with long correlation times. A common strategy for

simplifying this calculation is to expand the GLE back out into a set of Markovian equations[109, 110]. These equations describe a system bilinearly coupled to a bath of dissipative, stochastic harmonic oscillators. The advantage of using a dissipative bath, as opposed to an energy conserving bath, is that an extended bath can often be represented with one or two dissipative oscillators and that doing so dramatically reduces the dimensionality of the equations of motion. Here we briefly summarize the method, a more thorough derivation may be found in Appendix D.

Given a GLE with a memory kernel that is a finite sum of exponentially damped sinusoids,

$$K(t) = \sum_{i=1}^{N} e^{-\gamma_i t} \left(C_i \cos(\omega_i t) + D_i \sin(\omega_i t) \right), \tag{4.3}$$

the original non-Markovian equation of motion can be replaced with,

$$\frac{d}{dt} \begin{pmatrix} \dot{\mathbf{x}}_{\mathrm{A}} \\ \mathbf{b} \end{pmatrix} = \begin{pmatrix} -\frac{\partial W}{\partial \mathbf{x}_{\mathrm{A}}} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & \mathbf{\Lambda}_{\mathrm{AB}} \\ \mathbf{\Lambda}_{\mathrm{BA}} & \mathbf{\Lambda}_{\mathrm{BB}} \end{pmatrix} \begin{pmatrix} \dot{\mathbf{x}}_{\mathrm{A}} \\ \mathbf{b} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{\Sigma}_{\mathrm{BB}} \end{pmatrix} \begin{pmatrix} \\ d\mathbf{W} \end{pmatrix}. \tag{4.4}$$

Here $\mathbf{b}$ is a set of bath variables we must involve in time with our system. $d\mathbf{W}$ is an array of uncorrelated Gaussian random variables satisfying $\langle dW_i(t) dW_j(0) \rangle = \delta_{ij}\delta(t)$, where δ_{ij} is the Kronecker delta and $\delta(t)$ the Dirac delta. The matrix $\mathbf{\Lambda}_{\mathrm{BB}}$ is block diagonal with entries,

$$\mathbf{\Lambda}_{\mathrm{BB}} = \begin{pmatrix} 2\gamma_i & \sqrt{\gamma_i^2 + \omega_i^2} \\ -\sqrt{\gamma_i^2 + \omega_i^2} & 0 \end{pmatrix}, \tag{4.5}$$

and $\mathbf{\Lambda}_{\mathrm{AB}}$ and $\mathbf{\Lambda}_{\mathrm{BA}}$ are arrays of form,

$$\mathbf{\Lambda}_{\mathrm{AB}} = \begin{pmatrix} \sqrt{\frac{C_i}{2} - 2\frac{D_i\omega_i^2}{\gamma_i}} & \sqrt{\frac{C_i}{2} + 2\frac{D_i\omega_i^2}{\gamma_i}} \end{pmatrix}, \qquad \mathbf{\Lambda}_{\mathrm{BA}} = \begin{pmatrix} \sqrt{\frac{C_i}{2} - 2\frac{D_i\omega_i^2}{\gamma_i}} \\ \sqrt{\frac{C_i}{2} + 2\frac{D_i\omega_i^2}{\gamma_i}} \end{pmatrix}. \tag{4.6}$$

The matrix $\mathbf{\Sigma}_{\mathrm{BB}}$ is related to $\mathbf{\Lambda}_{\mathrm{BB}}$ by the equation,

$$\mathbf{\Sigma}_{\mathrm{BB}} \mathbf{\Sigma}_{\mathrm{BB}}^T = k_{\mathrm{B}} T \left(\mathbf{\Lambda}_{\mathrm{BB}} + \mathbf{\Lambda}_{\mathrm{BB}}^T \right), \tag{4.7}$$

which ensures that the ensuing dynamics obey the fluctuation-dissipation theorem.

Tully's GLO model is based on the same approach beginning from Eq. 4.3 and setting the number of terms $N = 1$. In the work we present in this manuscript, we determine the optimal values of N, as determined based on the analysis of $K(t)$. We drop the sin terms in Eq. 4.3, thus casting the memory kernel as a set of exponentially damped cosines and yielding a Lorentzian power spectral density of the form,

$$\bar{K}(\omega) = \sum_{i=1}^{N} C_i \left(\frac{\gamma_i}{\gamma_i^2 + (\omega - \omega_i)^2} \right). \tag{4.8}$$

4.2.2 Simulation details

Simulations were performed using the Atomic Simulation Environment (ASE) [73–78] and LAMMPS[111] packages. Simulations using Effective Medium Theory (EMT) [71] or Embedded Atom Method (EAM) [112] forcefields for metal dynamics were performed using ASE. The parameters for these forcefields were taken from Ref. 113 and Ref. 114 respectively. Simulations using Lennard-Jones forcefield, both solvated and in vacuum state, were performed using LAMMPS. Lennard-Jones forcefield parameters were taken from Ref. 115. The solvent used was SPC/E [116].

All simulations were performed in two steps. First, a temperature equilibration step was run for 50 picoseconds at 300K using a Langevin thermostat. Afterwards simulations were run in a constant energy ensemble using the velocity Verlet algorithm for 4 nanoseconds. Only data from the NVE step was used in subsequent analysis and calculations. All simulations were performed using periodic boundary conditions in the X/Y directions (parallel to the surface). Four atoms of the bottom row of the unit cell were rigidly constrained in order to remove center of mass motion.

For the surface scattering simulations used to generate data for Section 4.4, 5000 independent trajectories were averaged per value of the incident velocity to obtain sticking coefficients for GLE simulations, while 2000 independent trajectories were averaged for EMT simulations. Each trajectory was twenty picoseconds in length, which we found to be adequate for the convergence of results.

4.3 Memory kernels for surface sites in metal lattices

We begin by analyzing the memory kernel for the fluctuations of a single atom site in the surface of a 4x4x4 cell of Pt(111) with periodic boundary conditions. Results were calculated for each surface site individually and subsequently averaged. All memory kernels and spectral densities presented in the main text are calculated via analysis of time-correlation functions from molecular dynamics (MD) simulation (see Appendix B).

Figure 4.1A presents the orthogonal x, y, and z components of the memory kernel calculated from all-atom simulations using an Effective Medium Theory (EMT) forcefield. The x and y components arise from fluctuations in the plane of the lattice and the z component arises from fluctuations normal to it. Note the anisotropy between the x and y components and the z component, a simple consequence of the difference in the coordination number, and anisotropy between longitudinal and transverse modes. In the analysis that follows, we focus only on K_z—the component of the memory kernel acting in the z coordinate—because fluctuations in this direction (perpendicular to the metal surface) are most relevant to surface adsorption/desorption.

In Figure 4.1C we present the noise power spectrum (Fourier transform of the memory kernel) of the z component $K_z(\omega)$ specifically. As elaborated upon in Section 2.1, each peak in the power spectrum gives information about the lattice phonon modes and how they couple to the motion of a surface site. The value of ω at each peak corresponds to the frequency of the mode, the width γ corresponds to the timescale of energy exchange or dissipation between the mode and the surface site, and the coefficient C indicates the coupling strength of the system to that particular mode.

The spectral densities in Figure 4.1C are essentially bimodal. The red peak—henceforth referred to as the acoustic peak—is centered at a low frequency ($\omega = 18.6\text{cm}^{-1}$) and thus exchanges energy relatively slowly ($\gamma = 0.03\text{ps}^{-1}$), while the blue peak—henceforth referred to as the Debye peak—is centered near the Debye fre-

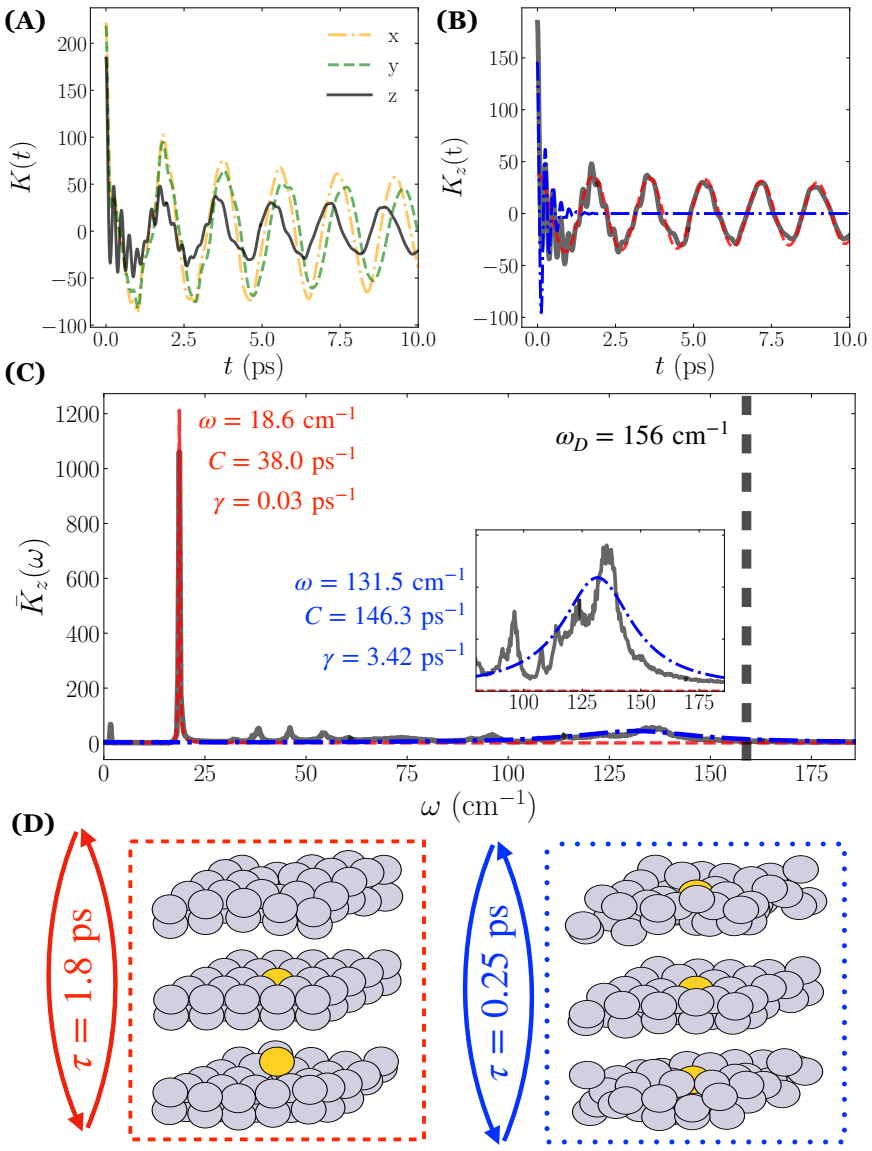

Figure 4.1: Memory kernel and random force power spectrum for surface sites of a Pt(111) lattice computed using an EMT forcefield. (A) Memory kernel for fluctuations in x/y (in-surface plane) and z (out of plane) directions. (B) z component of the memory kernel. Red and blue lines are two exponential sinusoids optimized to fit the computed memory kernel (grey line). (C) Power spectrum of z component of the memory kernel. The grey dashed vertical line corresponds to the experimental Debye frequency. (D) Depiction of lattice normal modes most associated with red and blue lines. τ is the period of the respective normal mode. The yellow-highlighted sphere in the middle of the lattice represents the chosen adsorption site, which is fixed in space, while the surrounding grey atoms represent the fluctuating bulk lattice.

quency of Pt ($\omega = 131\text{cm}^{-1}$) and exchanges energy relatively rapidly ($\gamma = 3.42\text{ps}^{-1}$). By comparing the spectral densities computed from MD simulation to the spectral

densities computed using the harmonic approximation, it is possible to determine precisely which normal modes of the lattice are primarily responsible for these two peaks. These normal modes are illustrated in Figure 4.1D. The acoustic peak arises from a longitudinal acoustic oscillations normal to the surface plane. Meanwhile, the Debye peak arises from many closely spaced normal modes near the Debye frequency, which consist of atomic scale local oscillations.

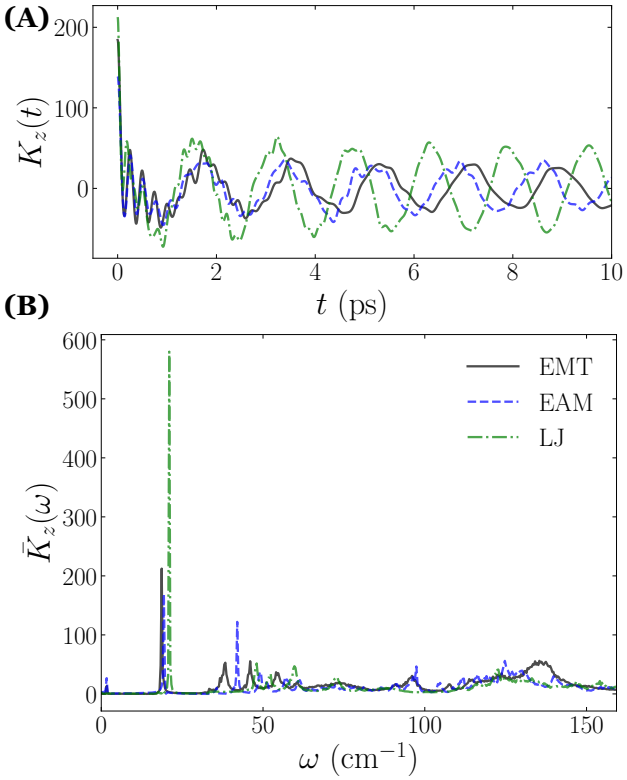

Figure 4.2: (A) Memory kernel and (B) spectral densities for surface site fluctuations of Pt(111) simulated using three different atomistic models: Effective Medium Theory, Embedded-Atom Method, and a Lennard-Jones model. The acoustic peak at 20 cm^{-1} of the spectral density has been scaled by a factor of 0.3 for the purposes of visual clarity.

In order to ensure the validity and transferability of our results, we tested the forcefields other than EMT. These results are illustrated in Figure 4.2. The Lennard-Jones (LJ) model is based on a very different underlying physics than the EMT/EAM models (LJ model uses only pairwise interactions, while EMT/EAM are both many-body potentials based on the local atom density). Despite this fact, all three models produce the same qualitative bimodal form. Much of the quantitative difference

between the EMT/EAM and LJ models can be explained by the fact that the LJ model results in a stiffer lattice than both EMT and EAM. The lattice stiffness can be roughly quantified in terms of the average value of the mass-weighted Hessian. For EMT the stiffness of the 4x4x4 Pt(111) lattice is 19.5 kJ/(mol nm^2), for EAM it is 19.8 kJ/(mol nm^2), and for LJ it is 37.1 kJ/(mol nm^2).

We also tested lattices of different elemental composition and surface facets. Once again, although variations were observed in the location, widths, and heights of the primary peaks of the spectral density, all of the lattices exhibited the same qualitative bimodal response. It is worth noting that this universality is not trivial. The bimodal behavior is not recovered in simple 1D systems with nearest-neighbor interactions (see Ref. 117 and 118 and Appendix E for more details), and therefore is an emergent property of the 3D metal lattice (see Appendix E).

4.3.1 Finite-size effects

Periodic boundary conditions are used in MD simulations to mimic an infinite system, but can still introduce finite-size effects into phonon-mediated processes. Specifically, the size of the crystalline solid sets the maximum phonon wavelength. As such, increasing the size of the simulation should shift the frequencies of the acoustic modes, which in turn affects the shape of the memory kernel. In Figure 4.3 we confirm that the acoustic peak shifts to lower frequency as the size of the lattice increases, while the Debye peak remains unchanged. Even when increasing the lattice size of ∼45 nm (8000 atoms), the memory kernel and power spectrum do not converge. In fact, the frequency ratio between the acoustic peak of different size lattices roughly agrees with the results of an isotropic wave equation, suggesting that the acoustic peak will decrease like as $1/L$ where L is the side-length of the lattice.

The demonstration that the memory kernel exhibits significant finite size effects has implications for the accuracy of simulation studies of surface phenomena. Perhaps most notably, this demonstration suggests that in the macroscopic limit, the frequencies of acoustic modes will be too low to affect any chemical dynamics at the surface. In other words, the acoustic modes are effectively frozen. Therefore, for observables

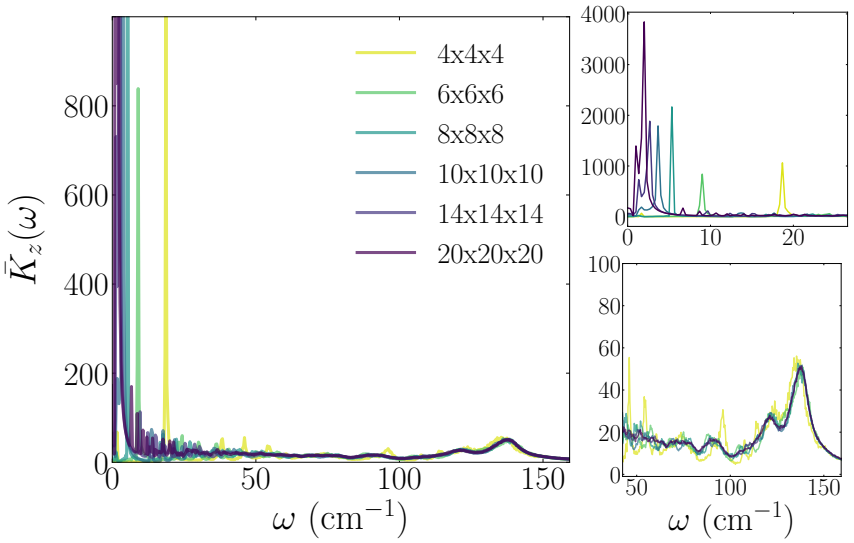

Figure 4.3: Memory kernel and spectral densities for surface site fluctuations of Pt(111) lattices of different sizes. The plots on the right zoom-in on the regions near to acoustic and Debye peaks respectively

that depend on memory, all finite-size simulations contain an intrinsic error which is purely kinetic in nature. We demonstrate this explicitly in Section 4.4.

The size-dependence of the memory kernel may also have ramifications for nanoparticle catalyst design, because it demonstrates how nanoparticle vibrational modes behave quite differently than their macroscopic counterparts, much as their electronic modes do. Indeed, experimental studies of electron relaxation in metal-supported nanoparticles have already shown that the phonon-mediated dissipation of electron energy depends strongly on the nanoparticle size [119].

4.3.2 Solvation effects

We now consider the effects of an adjacent solvent on the memory kernel of surface lattice vibrations. Thus far, we have assumed an idealized lattice in the limit of low pressure and substrate surface coverage. Most catalysts operate under conditions where there is significant surface coverage by solvent and/or reactive species. Here we will explore how solvation affects surface site fluctuations by computing memory kernels for Pt(111) surfaces solvated in SPC/E water using the CF approach. We will save the more difficult, yet still very important question, of surface coverage by

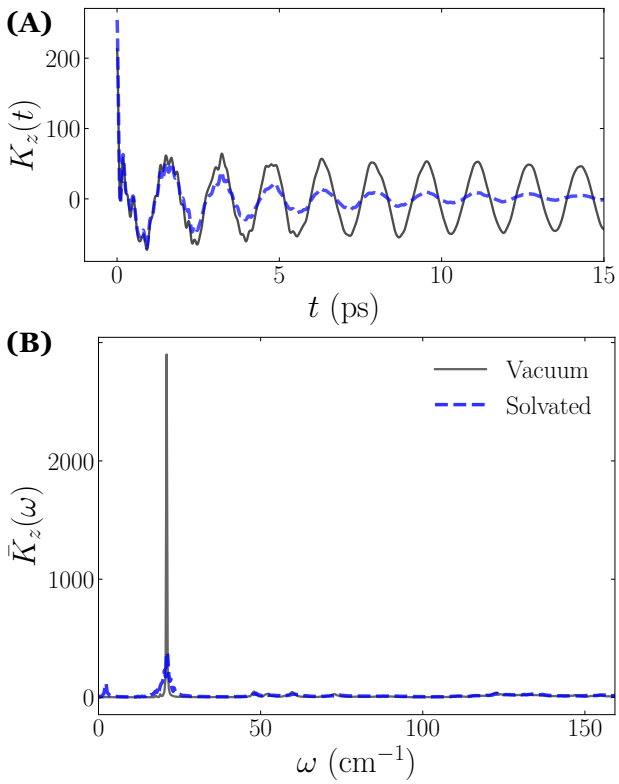

Figure 4.4: (A) Memory kernel and (B) spectral densities for surface site fluctuations of Pt(111) simulated using a LJ model with and without SPC/E solvent.

physisorbed and chemisorbed species for future study.

Figure 4.4 demonstrates the difference between Pt(111) surfaces in vacuum versus in solvent. The primary difference is in the damping of the acoustic mode, whose coupling to the surface site motion is much smaller when the surface is solvated. This effect is likely attributable to the additional pressure exerted by the solvent, making large fluctuations in the direction normal to the surface plane more energetically costly. The damping of the acoustic mode suggests that the finite-size effects discussed previously are likely far less important for solvated surfaces than they are for surfaces in the gas phase.

4.4 How memory affects scattering

Tully's GLO model is often used in simulations of surface scattering (either reactive or non-reactive) as an efficient computational method for describing energy loss

to the lattice during the scattering process [27, 83, 84, 107, 108]. In this section, we employ the LGLE for the same purpose, specifically studying the differences between the finite-size limit and the macroscopic limit (when the low-energy acoustic modes are held fixed).

We demonstrate our approach on the simple case of the non-reactive scattering of Argon on Pt(111), in the direction normal to the surface. The PES describing the interaction between the Argon atom and the Pt surface is taken to be of Morse form,

$$U(\Delta z) = D\left[1 - e^{-a(\Delta z - z_0)}\right]^2, \tag{4.9}$$

where Δz is the vertical distance between the Ar and the nearest Pt atom, D is a parameter which controls the depth of the PES well, a controls the width of the well, and z_0 is the location of the well's minimum. The values of these parameters were fit from DFT calculations presented in Ref. 79 using the vdW-DF2 density functional. Each scattering trajectory was initialized outside the Morse well at a distance of $z = 15$ Angstroms, with a fixed initial velocity directed into the well and randomized lattice atom positions/velocities drawn from a thermal distribution at 300K. From these trajectories, the sticking probability, S, was calculated as the ratio of trajectories that remain trapped within the well ($z \leq 15$) Angstroms after a collision with the surface.

Figure 4.5A illustrates variations in the sticking probability for using four different simulations. The blue curve uses an all-atom simulation using an EMT forcefield to treat the metal degrees of freedom. The red and orange curves use the LGLE parameterized from a 4x4x4 EMT simulation to treat the metal. The red curve uses only two damped sinusoids to fit $K(t)$, while the orange curve uses a five term fit to give a more accurate estimation of the memory kernel and power spectrum. The black curve corresponds to the extrapolated macroscopic limit of the LGLE, wherein the surface site motion is coupled to only to the Debye mode (see 4.5C).

The blue, red, and orange curves of Figure 4.5A largely agree with one another, illustrating that the LGLE accurately captures the dynamics of the forcefield it is

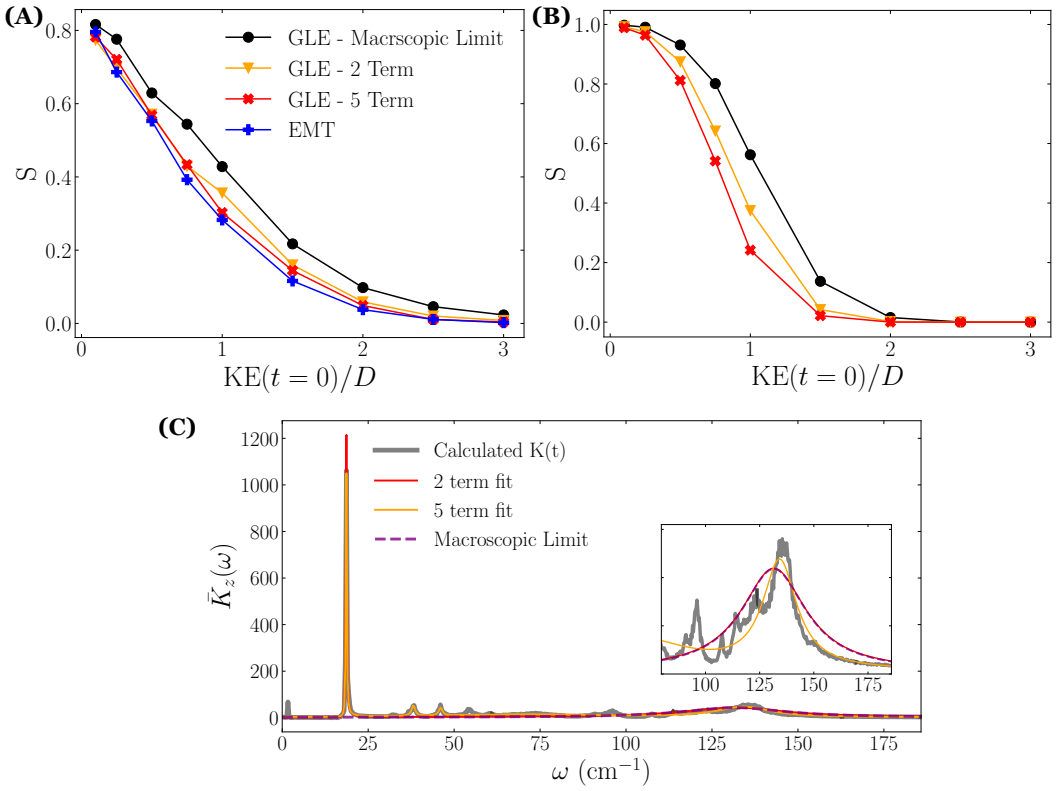

Figure 4.5: Sticking probabilities S as a function of the ratio of the incident kinetic energy to the well depth KE$(t=0)/D$. (A) Results for Morse PES with $D = 6.62$eV. (B) Results for Morse PES with an increased well-depth, $D = 30.62$eV. (C) Power spectra for a 4x4x4 Pt(111) lattice calculated using the CF method overlayed with the fits used to calculate results in (A) and (B).

parameterized from. More interesting however, is the consistent increase in the sticking probability between the nanoscale lattices (either modeled with EMT or GLE) and the macroscopic limit. This discrepancy can be qualitatively explained by the relative dissipation rates of the acoustic and Debye modes. Since nanoscale lattices couple the motion of surface atoms to the acoustic modes, and these acoustic modes dissipate energy much slower than the Debye mode, collisions with nanoscale lattices are more elastic.

In Figure 4.5B we study the scattering probability of a stickier particle with a well-depth that is nearly 5 times greater. Increasing D increases the effective coupling between the adsorbate and the metal phonon bath, exacerbating the finite-size effects seen in Figure 4.5B. The results in Figure 4.5 highlight an error that is inherent to nanoscale simulations of surface-molecule interactions. The implication of Figure 4.5

is that all nanoscale atomistic simulations of surface scattering contain an intrinsic error. This error is due purely to the phonon confinement effects imposed by the boundary conditions, and can be exacerbated by errors in the adsorbate's binding energy to the metal.

4.5 Conclusions

In this paper, we presented the lattice generalized Langevin equation, a model for simulating the effects of lattice phonons on surface atoms. The most important parameter in this model is the memory kernel. We parameterize the memory kernel using data from MD simulations, showing that it has a universal bimodal form due to coupling to both coherent acoustic oscillations as well as modes near the Debye frequency. This bimodal form is non-trivial, as it is not recovered in exactly solvable 1D systems with nearest neighbor interactions. Since the frequency of the acoustic oscillations depends on the size of the lattice, and nanoscale MD simulations impose unphysical phonon confinement effects, observables which depend on surface phonons will also contain artifacts. We showed that this was indeed the case for the surface trapping probability for a simple system of Argon on Pt(111).

The advantages of the LGLE model are, first, its computational efficiency, as it reduces the N degrees of freedom of the lattice to only a small handful of terms needed to describe the motion of a surface site. This dimensionality reduction makes the LGLE much faster than all-atom forcefields and particularly well-suited for the simulation of trajectory ensembles. Second, the insight that can be gained from studying the memory kernel, as we illustrated throughout this paper. Third, the transferability of the model. Once the LGLE is parameterized for a given type of lattice, any surface reaction with that lattice can use the same LGLE, given that the thermodynamic conditions (temperature/pressure/surface coverage/solvation) are roughly the same.

Part II

Solvent fluctuations at surfaces

Chapter 5

The influence of solvent on surface adsorption and desorption

5.1 Introduction

In Part I we focused on the dynamics and behavior of gas-phase adsorbates interacting with solid surfaces. However, many important technologies, including electrolyzers, batteries, fuel cells, hydrothermal reforming, and some capacitors, require controlling and exploiting chemical reactions that occur at solid-liquid interfaces [120–123]. In such reactions, the solvent degrees of freedom may also play crucial role in the reaction mechanism, thermodynamics and rate. [124–129]

Computational tools, such as classical and first-principles molecular dynamics simulations, can be utilized to gain insight into the mechanisms of surface chemical processes. However, the application of these tools to solvated interfaces is often impeded by the significant computational burden of simulating an extended explicit solvent due to system size and sampling requirements. Simulation cells with dimensions larger than 1nm^3 (and ideally much larger) are required to develop a solid-liquid interface that also includes a region of bulk liquid and an interfacial liquid surface that is free from finite size effects (e.g., due to orientational correlations that span the periodic boundaries). In addition, these large simulations must be sampled in order to represent the effects of thermal fluctuations in local solvent structure and

composition, e.g., due to the presence or absence of a dilute electrolyte species. Due to these challenges, molecular dynamics simulations of solvated reactions at metal interfaces often employ implicit solvent or small ice-like water layers [124–126, 130]. Such approaches have provided valuable insight, but they can neglect longer spatiotemporal correlations, focusing only on how water affects the energetics of local complexes. In principle, the computational burden of sampling the solvent could be reduced using enhanced sampling methods like metadynamics and/or the projection operator methods discussed in Part I. However, it is difficult to determine important and physically meaningfully collective coordinates of the solvent a priori.

In order to address these issues in this chapter, we use transition path sampling (TPS) to test the importance of different collective coordinates in solvent phase adsorption and desorption. We specifically study a model system of CO on a Pt surface, very similar to the model introduced in Chapter 3, due to its simplicity as well as the ubiquitous importance of CO adsorption in many catalytic reactions and processes, such as carbon dioxide (CO_2) electroreduction[131], Fischer-Tropsch synthesis [132], automotive catalysts [133], and supercritical-water biomass gasification [134].

The potential energy surface (PES) for CO desorption at a gas-phase Pt interface can be represented with only a few coordinates (e.g., CO-surface distance, CO orientation). In contrast, the PES for CO desorption at a liquid water-Pt interface is rugged and inherently high dimensional due to the influence of solvent configuration on the process. On this rugged landscape, there are numerous possible paths, corresponding to different molecular mechanisms. Determining which of these paths are relevant and which collective variables unite them can present a significant challenge. The method of transition path sampling (TPS), first proposed by Dellago et al.,[135, 136] was developed to address this challenge. TPS combined with tools of statistical analysis and enhanced sampling can be used to identify and validate possible reaction coordinates that incorporate the collective solvent degrees of freedom, and derive reaction free energy surfaces based on these coordinates. Using TPS we determine that solvent reorganization, as quantified by the solvent coordination number, is an important component of the reaction coordinate, thereby demonstrating solvent's crucial

role in the CO desorption pathway.

The remainder of this chapter is organized as follows. In Chapter 5.2, we provide an introduction to TPS and describe how it can be used to determine the importance of putative collective coordinates. In Chapter 5.2, we outline procedure and simulation methods we used to generate our results, and in Chapter 5.3 we provide and analyze said results. Specifically, using TPS we examine three potential collective variables relevant to CO adsorption/desorption: the distance from the surface, the solvent coordination number, and the CO orientation. We show that both the distance to the surface and the solvent coordination number are crucial to understanding the transition states of the desorption process. Thereafter, we use metadynamics to sample the potential of mean force along these collective coordinates and demonstrate that while in gas-phase the desorption process has no barrier, in solution phase the potential of mean force has a broad transition state arising from solvent reorganization.

5.2 Transition path sampling

Here we provide an overview of TPS theory noting that excellent reviews of TPS can be found in Ref. 137 and Ref. 138. TPS is a method for the Monte-Carlo sampling of trajectories connecting reactants and products, which here we define to be adsorbed (A) and desorbed (D) states, respectively. A trajectory is a time-ordered set of states, $\mathbf{\Gamma}(\tau) = \{\Gamma_0, \Gamma_{dt}, \Gamma_{2dt}, \ldots, \Gamma_\tau\}$, where $\Gamma_t = \{\mathbf{q}_t, \mathbf{p}_t\}$ is a vector of both positions and momenta at time t, and τ denotes the total length of the trajectory. Let $\mathbf{\Gamma}^{(o)}$ denote an initial trajectory in this transition path ensemble. In order to sample a new trajectory $\mathbf{\Gamma}^{(n)}$ from $\mathbf{\Gamma}^{(o)}$, we first select a time index t', sample a new state $\Gamma_{t'}^{(n)}$, and accept that state based on the Metropolis criterion,

$$P(\Gamma_{t'}^{(n)}|\Gamma_{t'}^{(o)}) = \min\left[1, \frac{p(\Gamma_{t'}^{(n)})}{p(\Gamma_{t'}^{(o)})}\right]. \tag{5.1}$$

Subsequently, we integrate the equations of motion forward and backward in time from $t = t'$ to $t = \tau$ and $t = 0$. In order to ensure that the entire trajectory connects reactants and products we accept the trajectory based on the criterion,

$$P(\mathbf{\Gamma}^{(n)}|\mathbf{\Gamma}^{(o)}) = h_{\mathrm{A}}(\Gamma_0^{(n)})h_{\mathrm{D}}(\Gamma_\tau^{(n)})P(\Gamma_{t'}^{(n)}|\Gamma_{t'}^{(o)}) \tag{5.2}$$

where the indicator function $h_{\mathrm{A}}(\Gamma_0^{(n)})$ returns 0 if the trajectory hasn't reached the reactant/adsorption (A) basin by $t = 0$ and otherwise returns one, and $h_{\mathrm{D}}(\Gamma adfs_\tau^{(n)})$ works similarly for the product/desorption D basin. The use of h_{A} and h_{D} ensures all paths sampled are legitimate transition paths connecting reactants and products.

The application of a TPS algorithm yields an ensemble of trajectories that connect the reactant and product states, *i.e.*, the A and D states in the case of this study. Once this ensemble has been constructed, it can be analyzed to determine statistics of the transition process. One of the most important statistics is the *committor*, p_D, the probability that a system which begins is in neither the A nor D basin, will relax to state D. The committor is a statistical indicator of the reaction progress. The transition state ensemble (TSE) can be identified as an isocomittor surface where $p_{\mathrm{D}} = p_{\mathrm{A}} = 0.5$, as such states are, by definition, equally likely to relax to reactants or products. We note that in general the value of p_{D} for a given configuration, and therefore the properties of the TSE, depend on how A and D are defined and on how trajectories are constrained during the implementation of TPS.

Analysis of the TSE allows one to discern which collective coordinates are necessary to characterize the transition state and are therefore good reaction coordinates. The statistics of committor values, quantified via the committor distribution function, $P(p_{\mathrm{D}})$, can be analyzed to evaluate whether a proposed reaction coordinate is a good descriptor of the reaction mechanism. More specifically, the so-called *histogram test* considered the shape of $P(p_{\mathrm{D}})$ evaluated at the value of the reaction coordinate for which $\langle p_{\mathrm{D}} \rangle = 0.5$.[139] A reaction coordinate that accurately captures the reaction mechanism will present a histogram ($P(p_{\mathrm{D}})$) that is narrowly peaked at $p_{\mathrm{D}} = 0.5$. As the shape of the committor histogram is highly dependent on the number of bins and

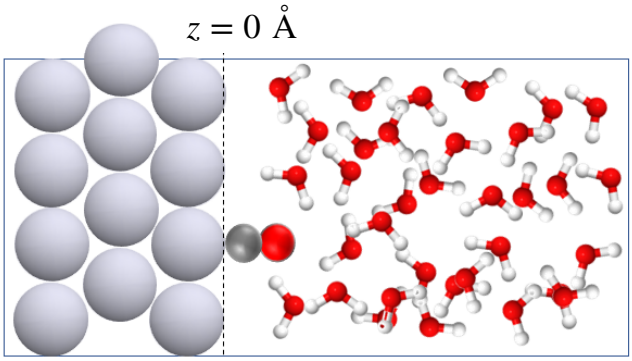

Figure 5.1: Snapshot of simulation system consisting of a three-layer Pt(100) surface with CO molecule adsorbed atop a central site and 57 water molecules in 1650 Å^3 volume above surface to achieve a water density of roughly 1 gm/cm^3.

the size of the dataset, it is common to fit the mean and variance to a beta-distribution model, giving a less biased estimate of the distribution [140].

5.3 Simulation methods

Our basic model, illustrated in Figure 5.1, consists of a 4×4×3 (atoms) Pt(100) surface with a single CO molecule adsorbed atop a central surface site. The surface was solvated by a layer of 57 water molecules roughly 1.5nm thick to yield a water density of 1 g/cm^3. The system was periodically replicated in the directions parallel to the Pt surface. Fixed, reflecting boundary conditions were used in the direction normal to the surface. A SPC-pol3 forcefield[141] was used to model the water–water interactions. For metal–water interactions, we used universal force-field parameters with Lorentz-Berthelot[142] mixing rules. To accurately capture the interaction between the carbon atom of the CO molecule (C_{CO}) and the platinum surface, we fitted density functional theory calculations to a Morse potential,

$$U_{C-Pt}(z) = D(1 - e^{-\alpha(z-z_0)^2}) \tag{5.3}$$

where z is the displacement between the carbon atom of CO and a Pt surface atom, the well-depth D is 346 meV, the well width α is 3 Å^{-1}, and the equilibrium C–Pt

distance is 2.5 angstroms. These parameters were fit to DFT calculations of a single CO molecule adsorbed atop a Pt(100) site using BEEF-vdW functional. A plane-wave basis set was used with a 600 eV energy cutoff. Electronic structure calculations were performed with the GPAW program [143, 144]. All simulations were performed in LAMMPS [111] using a constant-temperature (NVT thermostat) at 300K, with Pt atoms held fixed.

For TPS simulations, we defined the states A and D based on the distance of the CO carbon from the plane of the Pt surface, which we denote as z ($z = 0$ defines the position of the Pt surface). Specifically, we define the adsorbed state, A, as all configurations with $z < 2.5$ angstroms and the desorbed state, D, as all configurations with $z > 5.0$ angstroms. We used the open source package *openpathsampling* to generate 2000 different transition paths connecting the two basins.[145] We employed the spring algorithm [146] for our calculations, which is similar in concept the the shooting algorithm but randomizes the time index at which new trajectories are generated, picking a new random velocity at a given timestep. Note that spring shooting algorithms always use a stochastic thermostat. We assigned a configuration to the TSE if it was one of the 1000 configurations closest to the $p_{\mathrm{D}} = 0.5$ configuration.

After using TPS to identify a collective reaction coordinate, we utilized metadynamics to compute the associated free energy surface. Metadynamics, developed by Laio and Parrinello,[147] samples rare events by adding an artificial Gaussian bias potential to the existing potential energy surface. The bias potential is updated at regular intervals in an MD simulation and pushes the system away from regions that have been heavily sampled. Metadynamics has been applied to numerous problems in chemistry and materials science [148–154]. We performed well-tempered metadynamics[155, 156] by adding a bias potential every 100 fs with a bias-factor of 5, Gaussian height 0.5 kcal/mol and width 0.2 in an NVT MD simulation at 300 K. We ran these MD simulations for 10 ns each to obtain a two-dimensional free-energy surface (FES) for the chosen collective variables through TPS - z-coordinate of CO molecule (z), and water coordination around CO molecule (η_w). Metadynamics simulations were performed using the PLUMED plugin in LAMMPS [157].

We note that our TPS simulations utilized an additional Morse potential set at $z = 13.5$, with parameters $D = 0.43$ eV, $\alpha = 3$ Å^{-1}, and $r_0 = 0.5$ Å. The purpose of this additional potential is to introduce an artificial basin of attraction to the desorbed state, which otherwise presents a flat free-energy profile. By introducing this basin of attraction, the TPS trajectories can be shorter in length, leading to improved computational efficiency. This basin is displaced from the transition state and thus does not influence the characteristics of the transition state. The basin is not present in the calculation of the free energy surfaces.

5.4 Results and discussion

5.4.1 TPS

For any adsorption/desorption process, the direction normal to the metal surface (the z-axis in Figure 5.1) is an intuitive choice of reaction coordinate. However, especially in the presence of a solvent, there may be other collective variables, for example the adsorbate's surroundings and its orientation, that are crucial for modeling the transition states of the desorption process. To determine the significance of a given collective variable, we use TPS and committor analysis, as described in the Section 2.1.

Figure 5.2(A)–(C) shows the average committor probability p_{D} for a configuration to reach the desorbed state as a function of three collective variables: z_{C}, η_{W}, and ϕ_{CO}. z_{C} is the distance between $\mathrm{C_{CO}}$ and the Pt surface. η_{W}, which we henceforth term the *solvent coordination number*, is defined as a function of the distance of the nearest-neighboring water molecules to the CO molecule,

$$\eta_{\mathrm{W}} = \sum_i \begin{cases} 0.5\left[\cos\left(\dfrac{\pi d_i}{r_c}\right) + 1\right] & d_i < r_c \\ 0 & d_i \geq r_c \end{cases} \tag{5.4}$$

where the index i goes over all water molecules, d_i is the distance between $\mathrm{C_{CO}}$ and

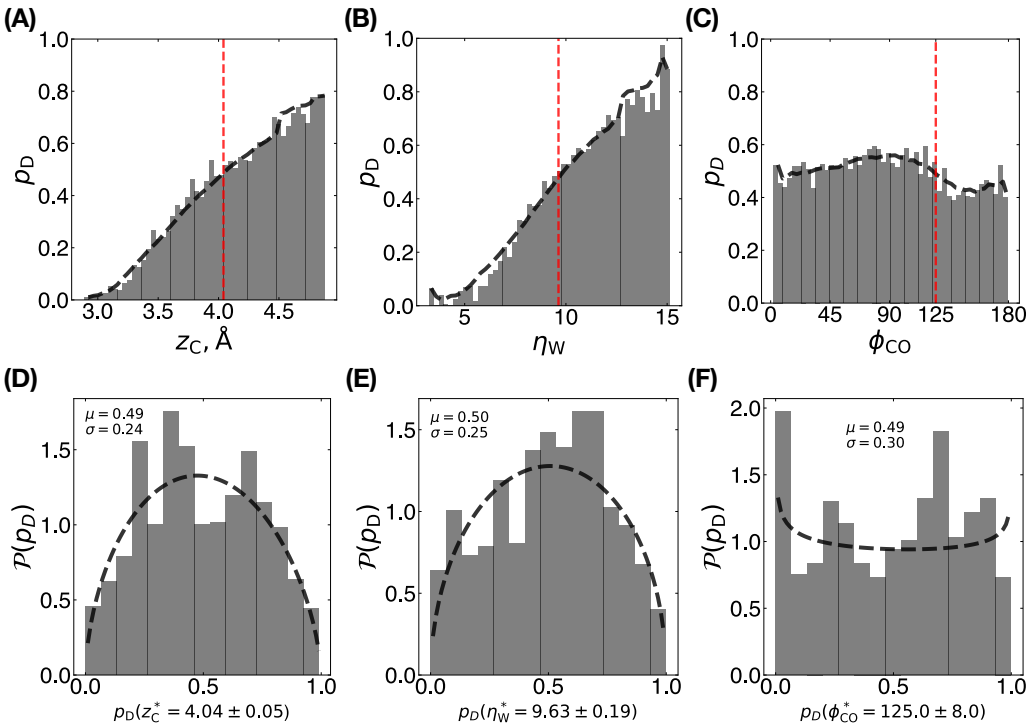

Figure 5.2: (A) Committor as a function of distance to surface. The black dashed line is a windowed average over the bar heights with a window of 5. The red vertical dashed line corresponds to the configuration closest to $p_D = 0.5$. (B) Committor as a function of water coordination number. (C) Committor as a function of CO orientation. (D) Committor histogram (grey bars) and beta distribution model (black dashed line) for configurations near $z_C^\ddagger$. (E) Committor histogram and beta distribution model for configurations near $\eta_W^\ddagger$. (F) Committor histogram and beta distribution model for configurations near $\phi_{CO}^\ddagger$.

the center of mass of water molecule i, and r_c is a cutoff radius. We set r_c to 6.5 angstroms. Eq. 5.4 allows us to quantify the CO molecule's local solvation shell using a differentiable quantity suitable for the application of enhanced sampling methods. ϕ_{CO} is the CO molecule's orientation relative to the surface, with 0° corresponding to the carbon-down configuration.

As expected, the committor probability increased monotonically with z_C, reinforcing the trivial notion that z_C, is an important collective variable. We also observe that $p_D(\eta_W)$ increases monotonically, indicating that solvent coordination is also important. The Pearson correlation coefficient between z_C and η_W taken over all configurations sampled is about 0.7, indicating that both collective variables, while correlated, contain some independent information. In contrast to z_C and η_W, the ϕ_{CO}

shows no clear correlation to the committor, indicating that it is not an important collective variable determining the outcome of CO adsorption/desorption. Previous experimental and theoretical studies have suggested that, in gas phase, *CO* orientation can play an important factor in kinetics of adsorption and desorption when at sufficiently high temperature, due to the large loss in rotational entropy upon binding to the surface [158–160]. However, the same studies found that such effects were negligible at room temperature. Since all our simulations were carried out at 300K, our finding that ϕ_{CO} is not correlated with the committor is consistent with the existing literature for gas-phase desorption. However, we note that in general the committor is a temperature-dependent variable, and the conclusions here may not necessarily hold true at higher temperatures.

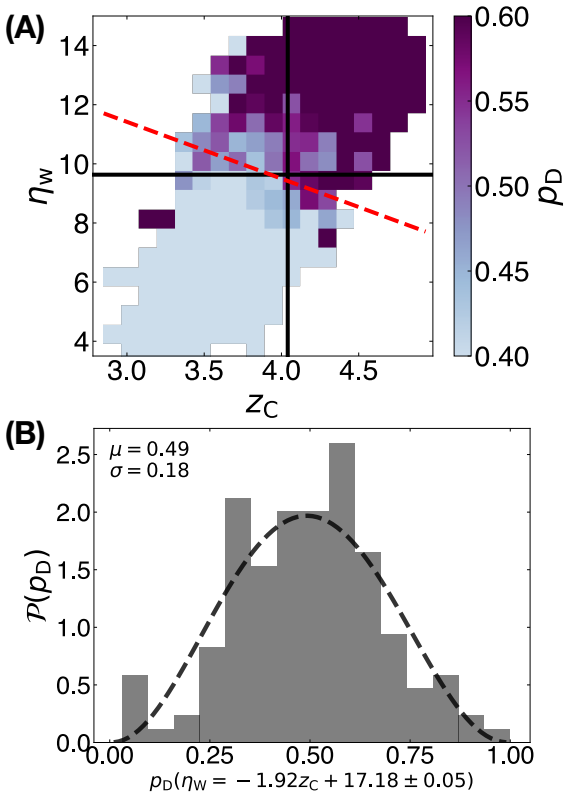

Figure 5.3: (A) 2D committor probability as a function of the distance to surface and the water coordination number. The black vertical line corresponds to $z_{\mathrm{C}} = z_{\mathrm{C}}^{\ddagger}$, the brown horizontal line corresponds to $\eta_{\mathrm{W}} = \eta_{\mathrm{W}}^{\ddagger}$, and the red dashed line to states along $\eta_{\mathrm{W}} = 1.92z_{\mathrm{C}} + 17.18$. (B) Committor histogram and beta distribution model taken for states near the red dashed line in 3A.

The red, vertical, dashed lines in Figure 5.2(A)–(C) represent the value of each coordinate for which the mean committor has the value $\langle p_{\mathrm{D}} \rangle = 0.5$. We henceforth denote the value of the collective variables at these transition states with a superscript double dagger, e.g., $z^{\ddagger}$. Figure 5.2(D)–(F) presents the committor distribution function (gray bars) and corresponding beta-distribution estimates (black dashed lines), corresponding to 1000 configurations taken near these transition states. A good coordinate should display a histogram narrowly peaked around $p_{\mathrm{D}} = 0.5$, meaning typical configurations at the nominal transition state also have $p_{\mathrm{D}} = 0.5$. As expected, all of these distributions have means at ($p_D \approx 0.5$). Unsurprisingly, the histogram/distribution for ϕ_{CO} is particularly wide, suggesting that $\phi^{\ddagger}_{\mathrm{CO}} = 125$ is not a good description of the transition state.

The committor distributions for both $z^{\ddagger}_{\mathrm{C}}$ and $\eta^{\ddagger}_{\mathrm{W}}$ transition states are centered near the mean; however, they are still much wider than what is considered gold-standard for a good collective coordinate ($\sigma = 0.15$ [140]), illustrating that there are still many configurations with $z^{\ddagger}_{\mathrm{C}} \approx 4.04$ or $\eta^{\ddagger}_{\mathrm{W}} \approx 9.63$ which do not necessarily have $p_D \approx 0.5$. Therefore, while both z_{C} and η_{W} correlate strongly with the adsorption/desorption process, alone they do not seem to capture all the information necessary to capture the transition states, and therefore the kinetics.

To further expand on these results, in Figure 5.3(A) we present a 2-dimensional histogram of the committer probability as a function of z_{C} and η_{W} simultaneously. The previous transition states $z^{\ddagger}_{\mathrm{C}}$ and $\eta^{\ddagger}_{\mathrm{W}}$ are indicated by the black vertical and horizontal lines respectively. One can clearly see that many configurations along these lines have committor values either much higher or much lower than $p_{\mathrm{D}} = 0.5$. For example, a CO molecule at $z_{\mathrm{C}} = z^{\ddagger}_{\mathrm{C}}$ is very likely to desorb if it also begins with a solvent coordination number $\eta_{\mathrm{W}} > 12$. Instead of the one-dimensional transition state surfaces defined by $z^{\ddagger}_{\mathrm{C}}$ and $\eta^{\ddagger}_{\mathrm{W}}$, a better transition state surface can be defined along the line $\eta^{\ddagger}_{\mathrm{W}} = 1.92 z^{\ddagger}_{\mathrm{C}} + 17.18$, shown as the red dashed line in Figure 5.3(A). The committor histogram along this 2D transition-state surface, Figure 5.3(B), has a significantly smaller standard deviation ($\sigma = 0.18$) than considering z_{C} and η_{W} alone. However, we acknowledge that there may yet be some other subtle solvent

coordinates playing a role in the desorption process.

5.4.2 Free-energy surface

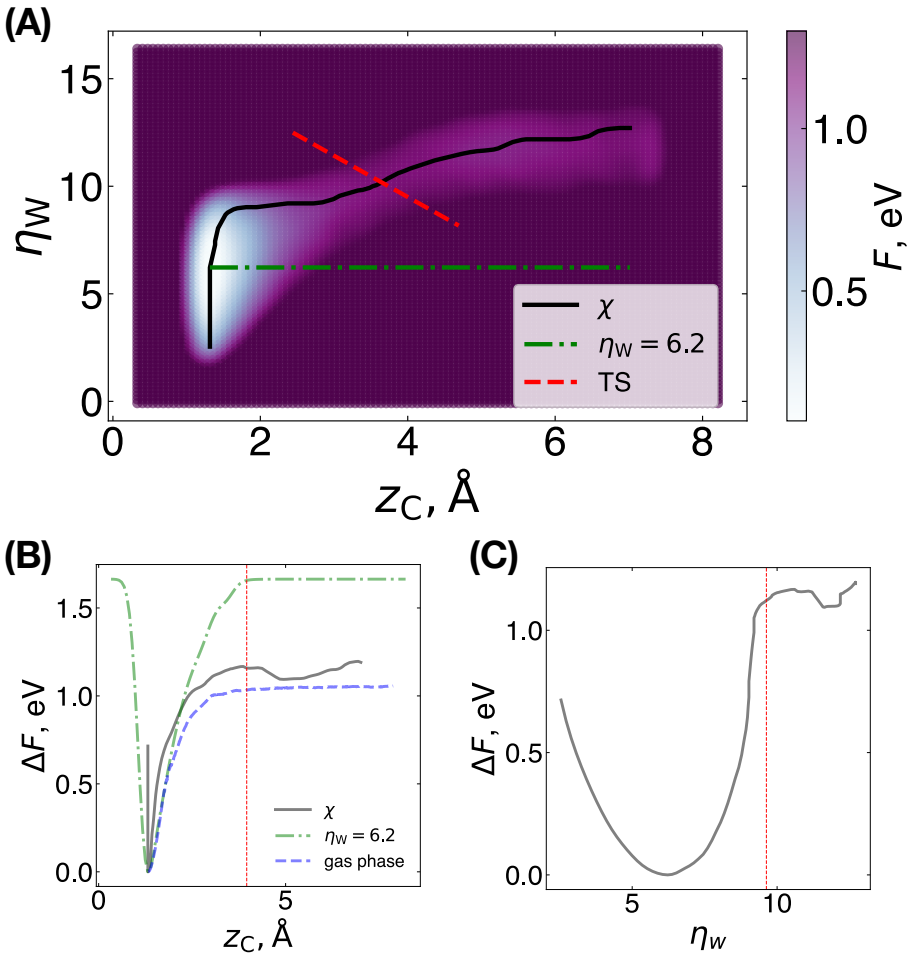

Figure 5.4: (A) 2D FES along z_C and η_W calculated using metadynamics. The black contour corresponds to the path of minimum local free-energy stage (steepest descent) between reactants and products. A minor windowed average smoothing, with a window of 3 was applied to this contour to remove noise. The green contour represents a path of constant η_W. The red dashed line represents the transition state surface shown in Figure 2A found using TPS. (B) 1D cross-section of free-energy as a function of z_C along the χ and $\eta_W = 6.2$ contours, as well as the gas phase desorption free-energy surface. (C) 1D cross-section of free-energy as a function of η_W along the χ contour. The red vertical lines show the $p_D = 0.5$ states found with TPS.

In Figure 5.4, we present the 2-dimensional FES - F - for CO desorption along the relevant coordinates identified by committor analysis, z_C and η_W. This FES was calculated using metadynamics. Note that while the adsorption basin is clearly

evident, with a minimum around $z_C = 1.3$ and $\eta_W = 6.4$, and the desorption basin is quite wide.

We considered two desorption pathways. The green line corresponds to a process where the water coordination number remains constant as the CO leaves the surface. The black line corresponds to the pathway along the gradient of the free-energy surface (henceforth denoted as χ) between the adsorbed and desorbed states. This is the path a CO molecule is most likely to travel. Analyzing the changes in slope of χ allows one to make conclusions about the steps and therefore mechanism of the desorption pathway. In particular, the results of Figure 5.4A suggest a multistep mechanism where, first, the solvent coordination number increases at roughly constant z_C, second, the molecule increases its distance to the surface at roughly constant η_W, and third, both z_C and η_W increase together. The initial increase in the coordination number occurs due to a constriction of the first solvation shell around the CO molecule before it desorbs.

In Figure 5.4B and Figure 5.4C, we present one-dimensional cross sections of F for both the constant-η_W and χ pathways, and contrast them to the gas-phase Pt-CO potential of mean force for a single CO molecule desorbing in vacuum. Comparing the dashed-dotted green line and solid black line reveals that the reaction barrier is much higher (≈ 0.5eV) if CO does increase its coordination number while desorbing. Also notable is the presence of a barrier for the desorption of CO in solvent (along χ), which does not exist for the corresponding system in vacuum. This barrier/transition state is not a result of a single, high-energy, bottleneck conformation, as in gas phase reactions, but a collective effect arising from solvent reorganization.[161]. The collective nature of this transition state, similar to transition states observed in classical nucleation theory, supports why it is naturally broad.

The red-dotted lines in Figure 5.4 correspond to the transition states ($p_D = 0.5$) we found using TPS. We see excellent agreement with the location of these lines and the location of the barriers on the free-energy surface. The agreement between TPS and metadynamics emphasizes the robustness of our results.

With regards to kinetics, the increased free-energy barrier to desorption between

the gas and liquid phase (∼0.2 eV), suggests a much slower desorption rate in solution. More interestingly, the fact that solvent reorganization must occur in order for the reagent to desorb highlights the importance of understanding surface–solvent interactions. A surface which interacts with solvent very strongly may hinder solute desorption not only directly, but through hindering solvent reorganization. In this study, we used a simple Lennard Jones model for the H_2O–Pt interaction, which leads to a somewhat weak interaction. We believe that a fruitful area for further research may be to study the effects of solvent-reorganization under different, and in particular stronger, surface-solvent potentials.

5.5 Conclusions

In this paper, we have used enhanced sampling methods to analyze the role of solvent in the desorption of a model system of a CO molecule on a Pt(100) surface. We first employed transition path sampling to determine the relative importance of different collective coordinates. Using the committor analysis we illustrated that both the distance to the surface and the solvent coordination number are needed to give a good description of transition states. Conversely, we found that the orientation of the CO molecule played a negligible role in the desorption process. We subsequently sampled the free-energy surface along our multi-dimensional reaction coordinate using metadynamics. Both methods showed excellent correspondence in the locations of transition states. Our metadynamics analysis indicated that desorption happens in a multi-step mechanism, where the initial step involves solvent reorganization and nearly constant distance to the surface. Not allowing for solvent reorganization leads to a free-energy barrier that is significantly higher. Our results highlight the importance of the solvent for kinetics of reactions at interfaces.

Part III

Exciton dynamics

Chapter 6

Machine learning Frenkel Hamiltonians [1]

6.1 Introduction

The performance of organic-based semiconductors depends not only on their molecular composition but also on how molecules are arranged within the material. The inter- and intramolecular interactions that underlie this dependence are well understood, but predicting how many such interactions combine to produce macroscopic material properties remains a significant scientific problem. Because these materials are often amorphous, lacking in microscopic periodicity, this problem cannot generally be solved with analytical theory. Computational modeling has played and will continue to play a vital role in advancing our understanding of these materials and guiding their development. Unfortunately, current methods for modeling materials with specific molecular structures include computational bottlenecks that limit their scope to system sizes that are too small to infer bulk properties. Here, we demonstrate that these bottlenecks can be effectively eliminated through the application of machine learning (ML), thereby enabling existing frameworks for computational modeling to be extended to larger and more experimentally relevant system sizes.

Organic molecular semiconductors (OMSs) provide an attractive alternative to

[1] Based on work published in Ref. 162. Copyright 2020, American Institute of Physics.

traditional inorganic semiconductors and have been increasingly utilized in the development of devices with notable applications to lighting, solar energy conversion, and even nanoscle electronic circuitry.[163–167] The performance of these materials is determined in large part by the properties of Coulombically bound excited electron-hole pairs, known as Frenkel excitons. When molecules within the material are weakly coupled, the static and dynamic properties of these excitons can be efficiently simulated with the Frenkel-Davydov model [168, 169], in which the molecular system is described in terms of a Hamiltonian for a coupled network of singly excitable chromophores. When appropriately parameterized, this efficient model has been shown to accurately reproduce experimental measurements, [170, 171] and large scale electronic structure calculations.[171–174]

The Frenkel Hamiltonian can be parameterized directly from first-principles based on all-atom configurations of representative systems of many organic conjugated molecules [169, 171, 173, 174]. Unfortunately, the procedure for this parameterization requires performing, at a minimum, one excited-state electronic-structure (ES) calculation per fragment. This cost is further exacerbated if one incorporates time dependence into the parameters, for example by analyzing a time series of configurations generated from molecular dynamics (MD) simulation. Despite the computational cost, including explicitly derived time dependence in the Frenkel Hamiltonian provides several key advantages over static models or models with phenomenological time dependence. Most significantly, trajectory-based parameterizations allow one to systematically evaluate the consequences of molecular disorder and correlated environmental noise on exciton transport.[171, 175] However, even for a system containing $\sim$10 chromophores (far below any reasonable approximation of the bulk), a 1ns simulation can require $\sim 10^6$ ES calculations. Even with the least demanding ES methods, this requirement represents an extremely large computationally expense.

Notably, however, many of the systems of interest, such as organic semiconductors [171, 175–178], DNA nanowires [179, 180], and biological light-harvesting complexes [174, 181–183] are comprised of many *identical* chromophore units with very similar electronic structure properties, *e.g.*, differing in nuclear or environmental con-

figuration. Here, we circumvent the costly model parameterization by developing a ML model to accelerate the computations associated with the ES calculations. This model utilizes the similarities across different configurations to rapidly predict molecular excitation energies and intermolecular electronic couplings. In this manuscript, we present this ML model and examine its application to two different OMSs morphologies made up of sexithiophene (T6) molecules.

The first system is a monolayer phase, in which molecules are aligned longitudinally in a manner analogous to experimental structures obtained via vapor deposition [184–186]. The second system is a completely amorphous bulk phase. We choose to study T6 because it has been widely characterized experimentally [184, 186–197], and its optoelectronic properties can be generalized to polythiophenes (such as P3HT). Mechanistic insight into the interplay between spatial disorder and exciton transport in these systems and other OMSs is key to making them more efficient and tunable, and our ML approach is an important step for building computer models capable of giving such insight.

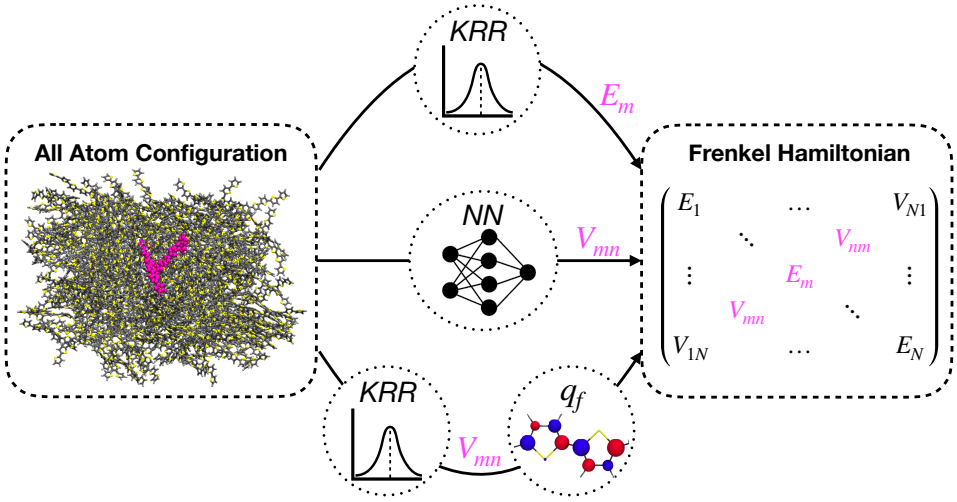

Figure 6.1: Depiction of machine learning schemes described by this work. Three machine learning models are tested, one for the excitation energy and two for the excitonic couplings. The deep neural network (DNN) model for direct prediction of couplings is shown to perform worse than the transition charge based kernel ridge regression (KRR) model.

Our approach to building a ML model Frenkel Hamiltonian, as illustrated in Figure 6.1, is to train separate models for the diagonal and off-diagonal elements, corre-

sponding to the excitation energies, E_m, and intermolecular electronic couplings, V_{mn}, using information only from the nuclear conformations of the constituent molecules. We consider two distinct approaches to building a ML model for V_{mn}. The first is to predict V_{mn} directly from bimolecular features. The second, is to predict a set of atom-centered transition charges using unimolecular features, from which the coupling may be obtained. We show that this transition charge approach is far superior to the direct approach, even when taking into account the additional error from the transition charge embedding.

In all cases, we evaluate the performance of our ML model through a set of four rigorous criteria. The first criteria is that the model apply to a broad range of ES methods, and therefore, we study its performance on both semi-empirical Pariser-Parr-Pople/Configurations Interaction Singles (PPP/CIS) [198] and time-dependent density functional theory (TDDFT) calculations. The second criteria is that the model require no more than 10,000 ES calculations as training data to yield error levels below that of the intrinsic uncertainty in the ES method. Third, the model should produce significantly less error than the intrinsic fluctuations in E_m and V_{mn} given by the interactions within the MD bath. The final criteria is that the model be *transferable*, meaning it should still perform well when encountering molecular conformations far outside what it was trained on. This criterion not only gives us added confidence in the performance of our model, but also allows us to apply our model beyond the morphology it was originally trained on. We show that both our E_m ML model and our transition charge based V_{mn} ML model satisfy the aforementioned criteria.

The organization of this chapter is as follows. We give a derivation the Frenkel Hamiltonian and associated matrix elements from first principles in Section 6.2. We then give a description of methods used for data generation, organization, and machine-learning in Section 6.3. In Secs. 6.4 through 6.6, we development and evaluate the ML models for the excitation energies and intermolecular electronic couplings. Finally, in Sec. 6.7, we provide concluding remarks.

6.2 The Frenkel-Davydov Model

The Frenkel-Davydov model supposes that the excited-state properties of an N molecule system may be expressed in terms of a reduced excitonic Hamiltonian

$$\hat{H}_F = \sum_m E_m \left|m\right\rangle \left\langle m\right| + \sum_{m \neq n} V_{mn} \left|n\right\rangle \left\langle m\right|, \tag{6.1}$$

where $|m\rangle$ represents the basis state where molecule m is in the excited state (all other molecules are assumed to be in the ground state), E_m is the excitation energy of molecule m and V_{mn} is the coupling between excitations on molecules m and n. When using CIS or TDDFT with the Tamm-Dancoff approximation (TDDFT/TDA), the basis states in Eq. 6.1 may be expressed as

$$|m\rangle = \sum_{ia} D_{ia}^m \left| \psi_a^m (\prod_{j \neq i} \psi_j^m) \psi_1^n \psi_2^n \dots \psi_N^n \right|, \tag{6.2}$$

where ψ_i^m refers to the ith orbital of molecule m and D_{ia}^m is the CIS coefficient for an excitation from occupied orbital i to virtual orbital a on molecule m. This basis can readily be expanded to the full N molecule case by appending ground state MO's to the determinant. Assuming the molecules interact weakly, we may optimize the CIS and MO coefficients for each fragment independently, and derive the Frenkel Hamiltonian matrix elements by projecting the supersystem electronic Hamiltonian into this basis. The couplings (off-diagonal elements) that result from this procedure may be expressed as,

$$V_{mn} = \sum_{iajb} D_{ia}^m D_{jb}^n \Big[2(\psi_i^m \psi_a^m | \psi_j^n \psi_b^n) - (\psi_i^m \psi_j^n | \psi_a^n \psi_b^m) \Big], \tag{6.3}$$

where $(ia|jb) = \int dr_1 dr_2 \psi_i(r_1) \psi_a^*(r_1) \frac{1}{r_{12}} \psi_j(r_2) \psi_b^*(r_2)$. It is common to make several additional approximations to make calculating this coupling more tractable. Many studies neglect the exchange terms in Eq. 6.3 under the assumption that the MO's on different molecules do not have significant spatial overlap [169, 199–204]. We

examined the contribution of exchange to the coupling in section 6.6 of this study, however, it is also worth noting that it is not clear if the Frenkel exciton model itself is appropriate in strong coupling regimes where exchange is significant.

The Coulomb couplings in Eq. 6.3, may also be simplified in terms of an interaction between transition charge monopoles q_f, on the atoms of each molecule [199].

$$V_{mn} \approx \sum_{f,g} \frac{q_f^m q_g^n}{|\mathbf{r}_f - \mathbf{r}_g|}, \tag{6.4}$$

where the summation is taken over pairs of atoms belong to separate molecules, $\mathbf{r}_f$ is the position of atom f, and q_f^m is the charge of atom f of molecule m. The transition charges may be determined via a variety of methods. One popular approach is to use Mulliken population analysis on the transition density matrix,

$$q_f = \sum_{ia} \sum_{\mu}^{N_{ao,f}} \sum_{\nu}^{N_{ao}} D_{ia} C_{i\mu} C_{a\nu} S_{\mu\nu}. \tag{6.5}$$

Here, the summation over μ and ν is taken over the atomic orbitals. $C_{i\mu}$ are atomic orbital coefficients, and $S_{\mu\nu}$ is the overlap between orbitals μ and ν. $N_{ao,f}$ is the number of atomic orbitals on atom f. In PPP/CIS theory, Eq. 6.4 is essentially equivalent to the exact coupling integral, due to the zero-differential ovelap approximation and highly localized orbitals[171]. However, for TDDFT and other ab-initio methods Eq. 6.4 is an approximation which should hold if the two molecules are sufficiently far apart. The particular charge embedding scheme in Eq. 5 is equivalent to Mulliken population analysis performed on the transition density matrix. This approach has been adopted in several studies [199, 200, 202, 203, 205] due to how it is significantly more computationally efficient than directly computing the two-electron integrals in Eq. 6.3. In Section 6.6A., we benchmark differences when calculating couplings using transitions charges derived from Mulliken, Löwdin, and Natural population analysis to the exact Coulomb coupling from TDDFT.

6.3 Computational Details

6.3.1 Data Generation

We consider two materials, each with identical chemical composition but differing molecular morphology. Specifically, we consider a monolayer morphology, of 150 molecules, where molecules are in rough mutual alignment, as illustrated in Figure 6.2A, and an amorphous bulk morphology of 343 molecules, as illustrated in Figure 6.2B. For each morphology, we generate a 10 ps trajectory of the nuclear dynamics using molecular dynamics (MD) simulations with the OPLS/2005 force field [206, 207]. Periodic boundary conditions were employed with the monolayer system periodically replicated in the plane parallel to the film, while the amorphous system was periodically replicated in 3 dimensions. All simulations were done in an NVT ensemble at 300K with a Nosé-Hoover thermostat [208, 209] with a coupling constant of 2.0 ps after equilibration. Simulations were done with the DESMOND package [210].

We map the MD snapshots to the Frenkel Hamiltonians using a set of independent one-molecule excited-state electronic structure (ES) calculations (see Section 6.2). We use two different ES methods to calculate Frenkel Hamiltonians: PPP/CIS and TDDFT with the Tamm-Dancoff Approximation (TDDFT/TDA). For our TDDFT/TDA calculations we employ the CAM-B3LYP functional [211], along with 6-31+G(d) basis set, due to the functional's demonstrated performance at reproducing the coupled-cluster excitation energies of organic chromophores [212–214]. Despite the functionals incomplete use of long-range corrections, Lu et al. demonstrated that the accuracy of lowest singlet excitation energy increases with the size of the number of monomers in an oligothiophene [212]. TDDFT/TDA calculations were done with the PySCF package [215].

Using PPP/CIS we mapped MD snapshots to Frenkel Hamiltonians calculations at 10 fs intervals, assuming we may treat the Frenkel Hamiltonians to be roughly constant within 10fs. Due to significant increased cost of TDDFT/TDA, a longer

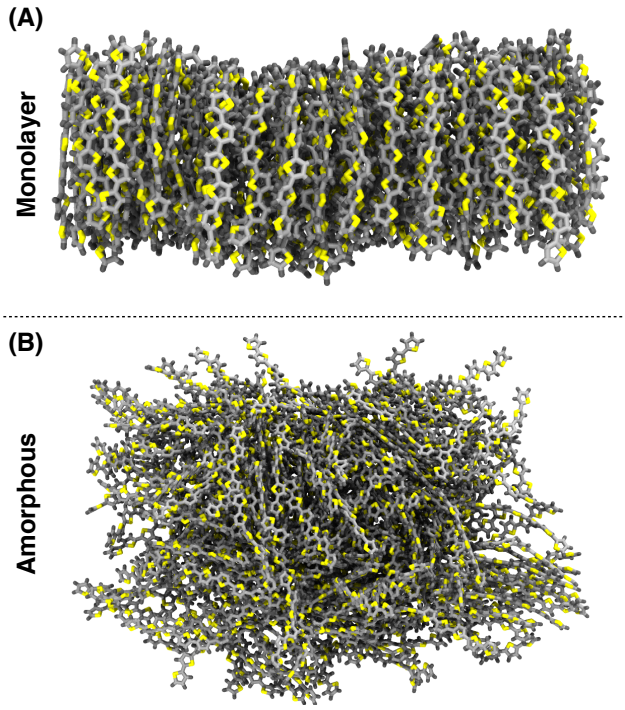

Figure 6.2: Snapshots of different molecular morphologies used in this study. Panels (A) and (B) show the monolayer and amorphous morphologies, respectively.

interval of 100 fs was used when calculating Frenkel Hamiltonians. We note however, that one of the advantage of our proposed ML model is that it may be used to calculate Frenkel Hamiltonians at intervals much smaller than even 10fs with little added cost. Couplings for molecules at intermolecular separations greater than 20Å are neglected due to the fact that they are extremely small (1meV on average for the monolayer system). While the 20Å cutoff has not been explicitly shown to have a negligible effect on exciton dynamics, we hope to use the efficiency of the ML model described herein to further elucidate the effects of coupling cutoffs in future studies.

6.3.2 Data Preparation and Machine Learning

To train and evaluate our unimolecular ML models (for predicting E_m and transition charges using KRR), the data (features and associated values) were separated into training and testing sets. The term "features" refers to the inputs X of the regression algorithm: information about the molecular conformations. The term "values" refers the quantities, $Y = f(X)$ we are trying to predict with the ML model. The

training sets ranged in size from 100 to 10,000 datapoints and the testing sets contained 5,000 datapoints. We found that 5000 testing examples was sufficient to get an unbiased estimate of the error. The training and testing data are always independent and randomly sampled, meaning that no two datapoints are shared between training and testing sets, and both are randomly chosen from the master dataset containing all of the available data.

To train and evaluate our deep neural network (DNN) for the direct prediction of V_{mn}, data was separated into independent randomly-sampled training, validation, and testing sets. The training sets ranged in size 130,000 to 520,000 datapoints, the validation set contained 200,000 datapoints, and the testing set contained 300,0000 datapoints. The validation set was used to pick the optimal number of epochs for early-stopping during the training phase of the DNN. Only PPP/CIS data was used in the training of our direct V_{mn} model. In Section 6.5 we report the sizes of the training sets in terms of the equivalent number of ES calculations needed to produce the couplings, as we believe number of ES calculations to be a more informative measure when trying to ascertain the cost vs accuracy tradeoff of the machine learning approach.

KRR models were trained using the scikit-learn python package[216]. The KRR hyperparameters were selected via 5-fold cross-validation. The DNN described in Section 6.5 was trained with Tensorflow 2.0[217]. The network consisted of three layers with 1200,1000,800 neurons respectively. The first and third layers used sigmoid activation functions, while the second layer used an exponential activation function to avoid the vanishing gradients problem. ADAM was used as the optimizer and Early Stopping with the validation dataset was used to avoid overfitting.

6.4 ML Model For Excitation Energies

There are numerous ML models and featurizations that are very successful at predicting various molecular energies from unimolecular conformational data [218–227]. Recently, Lu et al. [212] compared the performance of a series of molecular DNN's

at predicting excitation energies of high-temperature gas-phased T6 and showed a mean absolute error (MAE) of around 50 millielectronvolts (meV) when compared to the actual TDDFT results. Here, we adopt a simpler approach, using a kernel ridge regression (KRR) model with a Gaussian kernel. KRR offers several advantages to neural network models since it is generally much less expensive to train, has fewer hyperparameters, and is guaranteed to give the optimal solution (lowest training error) for a given set of training data and hyperparameters. Additionally, KRR is a well-understood, well-tested ML model that has already been shown to perform excellently as a way of accelerating non-adiabatic excited state dynamics[227–230].

For features, we employ the Coulomb matrix (CM) [218],

$$C_{ij} = \frac{Z_i Z_j}{r_{ij}}, \tag{6.6}$$

where Z_i are a set of atomic partial charges taken from semi-empirical model, and r_{ij} is the distance between atoms i and j. While traditionally the diagonal elements of the CM are given by $0.5Z_i^{2.4}$, here we neglect the diagonal elements since they give no additional information about the conformation of the molecule. We tested the performance of a series of four different charge models, as well as just using pairwise atomic distances without any charges, and found that while performances were comparable, the Gasteiger charge model[231] gave the lowest error (Figure 6.3A). Thus, in subsequent analysis the Gasteiger charge model was used whenever employing Coulomb matrix features.

We evaluate the performance of our ML model subject to the criteria outlined in the introduction through a series of numerical experiments. We begin by evaluating the dependence of the testing set MAE on the size of the training set for both the monolayer and amorphous systems (Figure6.3B). We see that in both systems the error converges well, with a difference of only 10meV or less in the error between training with 5000 examples versus 10,000 examples. For both systems, the converged errors are significantly lower than the intrinsic uncertainty of a TDDFT calculation with CAM-B3LYP or similar range-separated hybrid functional, which based on previous

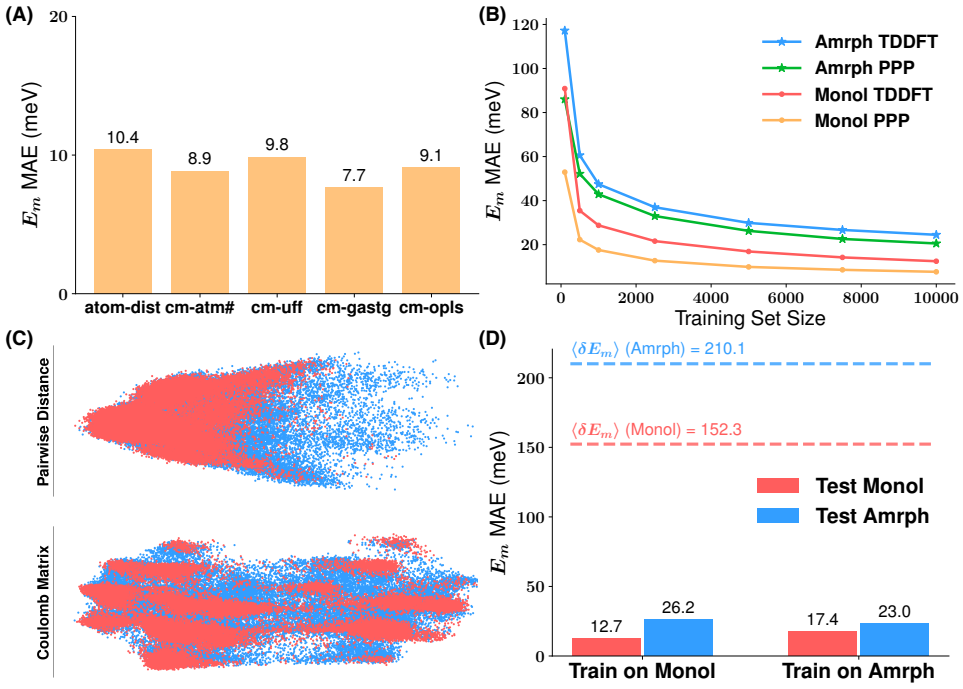

Figure 6.3: (A) Differential performance of using several different charge models with the Coulomb Matrix. Performances reported here are using the 10,000 example monolayer PPP/CIS training set. atom-dist = pairwise atomic distances (no charges), atm# = atomic numbers, uff = Universal Forcefield [232], gastg = Gasteiger charges[231], opls = OPLS/2005 Forcefield [206]. (B) Mean absolute error in E_m predictions versus training set size used. (C) 2D principal component analysis of pairwise atomic distances and coulomb matrices taken from T6 conformations. (D) Cross-morphology error analysis on predictions of E_m from TDDFT. Dashed lines indicate fluctuations (standard deviation) in E_m from MD bath.

studies we estimate to be around 200 meV [171, 212]. When training with 10,000 examples, the errors are only about 0.4% of the mean E_m in the monolayer system and 0.7% of the mean E_m in the amorphous system. However, comparing ML errors to the mean value of the excitation energy is not a very informative metric, as one can subtract any constant value out from diagonal elements of the Frenkel Hamiltonian without changing the dynamics. It is more appropriate to compare the ML errors to the fluctuations (standard deviation) in E_m, as one would expect that ML errors that are of the same order of magnitude or greater than the fluctuations would lead to significant errors in the exciton dynamics [175]. The standard deviation in E_m from the MD bath is about 150meV and 210meV for the monolayer and amorphous systems respectively if calculated with TDDFT, and about 103meV and 181meV if

calculated with PPP/CIS. These fluctuations are significantly higher than the MAE in Figure 6.3B so long as 1000 or more training examples are used. Thus, our model satisfies our first three performance criteria. The strong performance of our model matches results from other studies using kernel based regression methods to predict unimolecular quantities [220, 221, 223, 233].

Interestingly, the errors for the amorphous system tend to be higher than the MAE for the monolayer system. We attribute this increase to the greater range of intermolecular forces and larger conformational space sampled by the amorphous system. In Figure 6.3C, we demonstrate this larger sampling by performing principal component analysis (PCA) in two dimensions on the internal coordinates, showing that the area covered by the amorphous conformations is far larger than the area covered by the monolayer conformations. The wider range of intermolecular forces is also reflected in the larger standard deviation of E_m in the amorphous system.

Also of note is the fact that the error for the TDDFT data tends to be slightly higher than the error for the corresponding PPP/CIS data. The reduced performance on TDDFT data can be attributed to the increased complexity of the excited-state energy landscape from TDDFT as compared to PPP/CIS, which is a natural result of the more complex equations and much larger basis set.

To evaluate the final performance criteria: transferability across morphologies, we take a model trained with 10,000 examples on a given morphology, and test it against conformations from the other morphology. Despite the wider sampled space, the error of the model trained on monolayer conformations does not increase significantly when evaluated against conformations from the amorphous system (Figure6.3D). Indeed, while the KRR model's predictive performance is best on the morphology it was trained on, the the errors from cross-morphology predictions are still small fractions of the fluctuations in E_m. We take this result to indicate that our model can generalize to morphologies of T6 beyond what it was originally trained (if at roughly the same temperature), satisfying our fourth performance criterion.

6.5 Direct ML Model For Couplings

In principle, the increased scaling between the number of ES calculations and the number of couplings values merits using a different ML model than KRR. Each unique pair of molecules equates to a single coupling value, giving a total of $\frac{N(N-1)}{2}$ coupling datapoints per N molecules in a single timestep. Roughly, this scaling equates to 10,000 ES calculations per 260,000 couplings in both the monolayer and amorphous systems. The $O(N^3)$ cost of inverting the kernel matrix makes KRR unsuitable for datasets much larger than 50,000 examples in size, but DNNs have been shown to be excellent at handing "big-data". Thus, we chose to use a 3-layer neural network model for the direct ML of the couplings. For features we employ an intermolecular Coulomb matrix, with pairwise distances between atoms of separate molecules. A very similar model for electron-transfer coupling using a much smaller dataset with KRR was recently reported by Wang et al[234]. We decided to test this model only on the PPP/CIS data, due to the poor convergence of the model and need for larger training sets.

The predictive performance of the direct V_{mn} model is significantly worse than performance of the E_m ML model. Figure 6.4A demonstrates that the errors do not converge as quickly as they did in the E_m model. Although the errors here appear small, especially in comparison to the corresponding errors for E_m, one must keep in mind that the magnitude for V_{mn} is about two order of magnitude smaller than the magnitude of E_m in these systems. The mean magnitude for V_{mn} is only 14.7meV and 8.8meV for the monolayer and amorphous systems respectively, while the mean magnitude of E_m is about 2960meV and 3120meV respectively. Thus, the % error is much larger than what we had for E_m even when training with only 1,000 examples. Perhaps more importantly, the errors are large proportion of the fluctuations in V_{mn}, and the cross-morphology error is very significant. The latter suggests that even if we accept the lower accuracy to training size trade-off of this coupling ML model, for every different morphology that one might want to simulate, one would need to do tens of thousands of additional ES calculations, significantly hampering the efficiency

bonus of using ML.

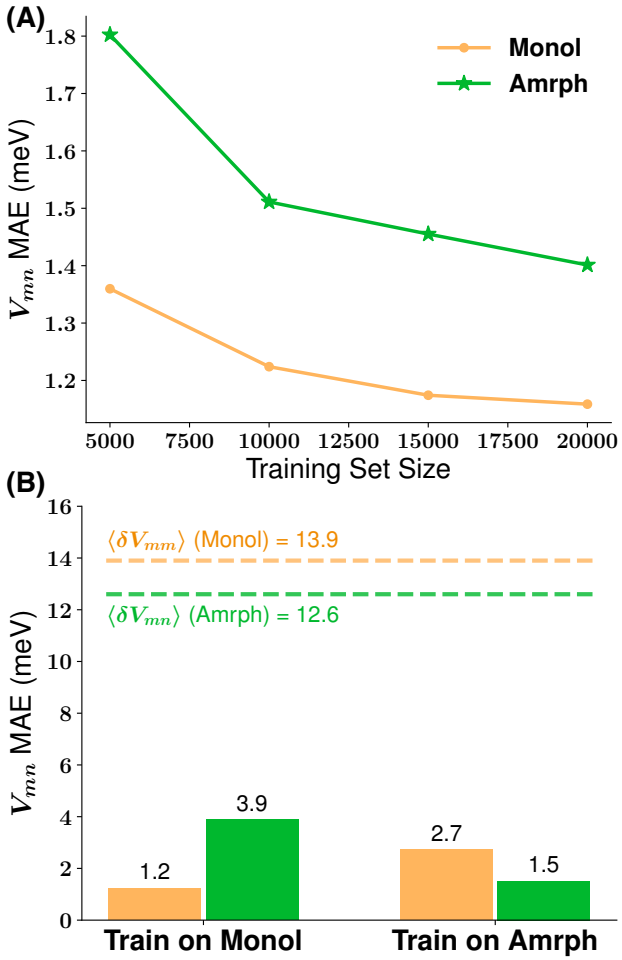

Figure 6.4: V_{mn} prediction performance using direct model. (A) Mean absolute error in PPP/CIS V_{mn} predictions versus the number of one-molecule PPP/CIS excited-state calculations used to generate the training set. (B) Cross-morphology error analysis on predictions of V_{mn} using 10,000 ES calculations = 260,000 coupling training sets.

We believe the poor performance of the direct V_{mn} model can be attributed to the very large bimolecular conformational space, which requires significantly more data to properly sample than that of the unimolecular space for E_m. The unimolecular conformational space only includes small perturbations of the nuclear positions due the vibrations and intermolecular forces, whereas bimolecular conformational space contains the same small perturbations as well as the myriad of different ways the two molecules (and constituent segments/rings) may be oriented with respect to each other. Clearly, sampling such a space thoroughly would require much more data.

Even with the naturally higher ES calculation to number of datapoints ratio (1:1 for E_m vs 1:26 for V_{mn}), the amount of ES calculations required to train a V_{mn} model from bimolecular data far exceeds the number of calculations required to train an equivalent accuracy E_m model from unimolecular data. Motivated by this difficulty, we sought to find a ML approach whereby we could predict V_{mn} from unimolecular features, which lead to our transition charge model.

6.6 Indirect ML Model For Couplings via Transition Densities

6.6.1 Coupling Calculation Benchmarking

In this section we benchmark several methods of calculating the excitonic coupling. We begin by comparing couplings calculated via PPP/CIS to exact couplings calculated via TDDFT (Figure 6.5A). The PPP/CIS couplings tend to significantly underestimate the TDDFT couplings for larger values. The decreased couplings of PPP/CIS would result in lower exciton transport rates when compared to TDDFT, highlighting the importance of the ML approach for allowing us to use higher levels of theory when parametrizing V_{mn}. However, given the very low computational cost and severe simplifications of PPP/CIS, we are encouraged by how well the PPP/CIS and TDDFT couplings are correlated.

There appears to be no clear consensus in the literature about whether exchange (Dexter transfer) is physically significant when considering the excitonic couplings in small-molecule organic semiconductors, so here we benchmark the error when neglecting exchange. In Figure 6.5B we illustrate the exchange contribution to coupling is completely negligible in the monolayer system, with an average difference of 0.03meV from between using exchange or not. This result is somewhat surprising as T6 molecules come in very close proximity to each other during the MD simulation, with the minimum center of geometry distance between two T6 molecules being as small as 3.5 Å.

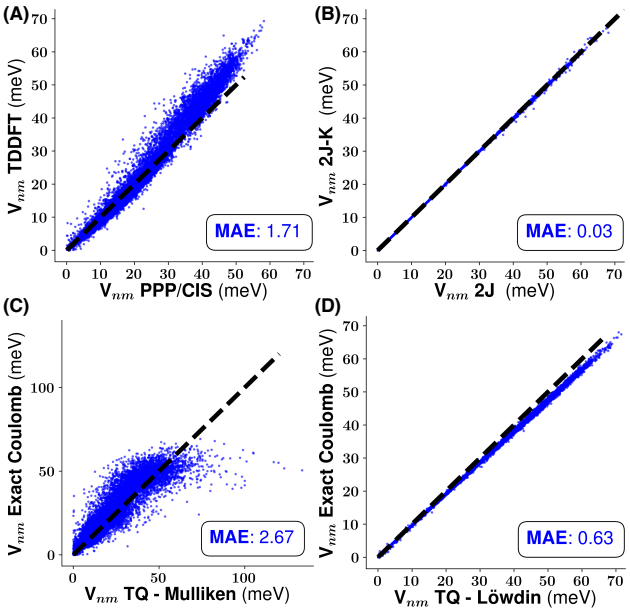

Figure 6.5: Parity plots for benchmarking coupling calculations. All values are from monolayer simulation. (A) Exact PPP/CIS versus TDDFT couplings. (B) TDDFT couplings with and without exchange. (C) TDDFT Coulomb couplings from Mulliken transition charges. (D) TDDFT Coulomb couplings from Löwdin transition charges.

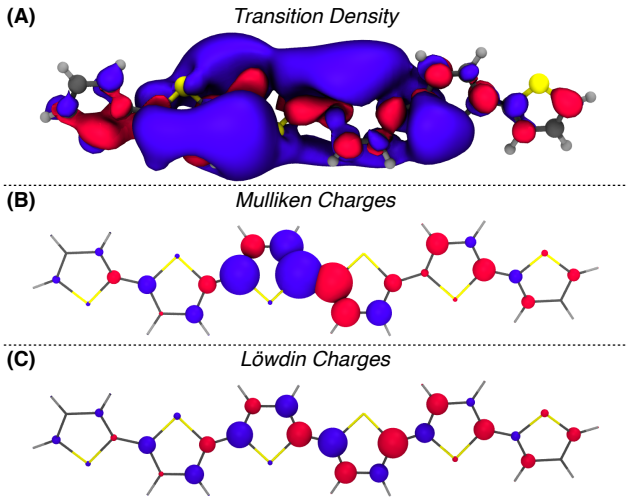

Figure 6.6: Charges and density visualizations for a typical molecule taken from monolayer simulation. (A) Isosurface of transition density using a value of 0.07, violet lobes are positive, red lobes are negative. (B) Mulliken charges, the size of each sphere indicates the magnitude of the assigned charge. (C) Löwdin charges.

Now that we have established that Coulomb couplings are sufficient when computing couplings in T6, we test whether they can be reproduced using a set of transition charges localized on each atom as in Eq. 6.4 (Figure 6.6). Both Mulliken and

Löwdin population analysis were tested. The Mulliken-charge based couplings show significant deviation from the exact Coulomb couplings, due to the significant overestimation of the charge on the central atoms of T6 (Figure 6.6B). Mulliken population analysis is well known to be highly basis set dependent [235]; in the limit where a complete basis set is localized on a single atom, Mulliken analysis will assign the entire transition density to that atom. In T6, the high overlap between the atomic basis functions of the carbons on the central rings with the surrounding atomic orbitals, can lead to an artificial concentration of charge on a few carbon atoms. This concentration of charge leads to couplings which are highly dependent on the positions of these carbons; i.e. if the central carbons are too close together, Mulliken charges can give up to double the actual coupling. Although it is well-known that Löwdin analysis may also show significant basis set dependence, in T6 the Löwdin orbital orthogonalization procedure leads to charges are more evenly spread across the atoms, giving a better representation of the actual charge density, and thus better couplings. The Löwdin charges do an excellent job at reproducing the Coulomb couplings, with an average difference of only 0.6 meV from the exact Coulomb integral calculations.

One may think that, given the high predictive performance of the Löwdin transition charges, it may be possible to calculate the couplings using a single-set of transition charges, similar to how interactions between molecules are calculated in molecular dynamics. This approach would side-step the need for high dimensional regression algorithms like NN and KRR, however, in Figure 6.7 we show that such an approach leads to an enormous increase in error, with an MAE of 1.74meV.

We also explored charge embedding based on the natural atomic orbitals as opposed to the Löwdin orbitals (so-called Natural population analysis), and found that the results were essentially equivalent to the Löwdin analysis Figure 6.8, and therefore proceed with using the Löwdin charges for the remainder of the study.

We emphasize that the transition charge method is effective not only because it gives us another avenue for machine learning the couplings, but also because it is significantly less computationally demanding than calculating the couplings exactly. In Table 6.1, we demonstrate that computing the exact Coulomb coupling calculation

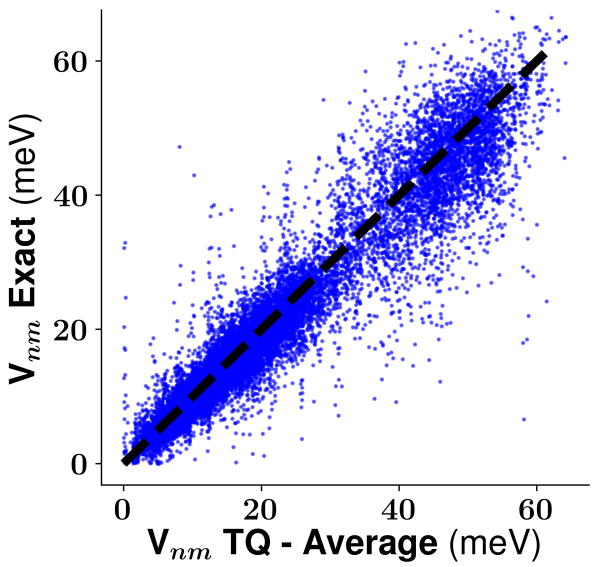

Figure 6.7: Parity plot comparing exact Coulomb couplings to couplings from using a single, average value of the Löwdin transition charges. The MAE here is 1.74 meV

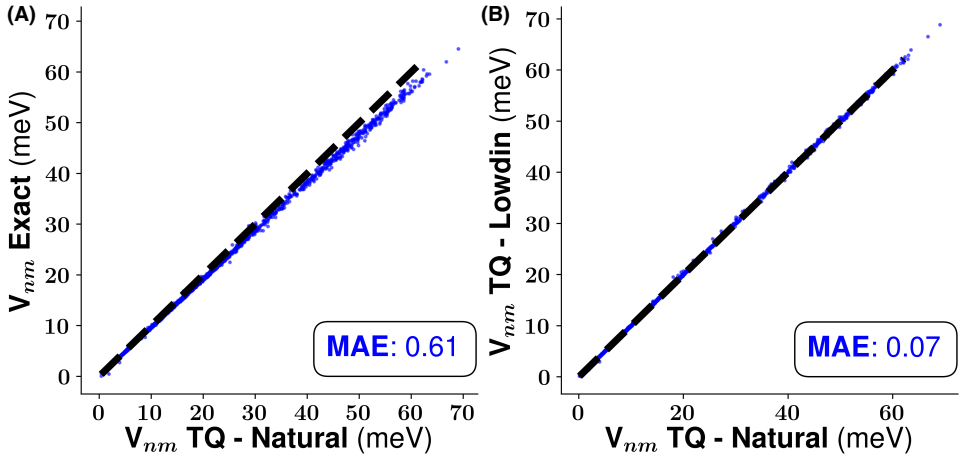

Figure 6.8: (A) Parity plot comparing exact Coulomb couplings to transition charge coupling from natural population analysis. All values are from monolayer simulation. (B) Parity plot comparing NAO and Lowdin transition charge couplings.

takes 8300 CPU hours for the monolayer system. Meanwhile, the cost of the Löwdin transition charge couplings is only 22 CPU hours. While, as shown in Figure5C, the Löwdin transition charge method produces some systematic error for molecules very close together, we believe this degree of error is acceptable in light of the enormous improvement in computational efficiency.

Table 6.1: Relative computational cost for different methods of computing coupling and TDDFT calculations. All CPU times were measured on an Intel(R) Xeon(R) CPU E5-2620 @ 2.10GHz. The costs shown are for the entire 100ps trajectory, but are estimated using the cost of a single calculation, and account for the 20Åcoupling cutoff.

Method	**CPU Hours - Monolayer**	**CPU Hours - Amorphous**
Lowdin TQ	22	51
Coulomb	8300	17800
Coulomb + Exchange	15800	34000

6.6.2 Machine Learning Model

In the preceding section we described how Löwdin transition charges can effectively reproduce the excitonic coupling V_{mn}, in this section we proceed with describing a ML model for V_{mn} by predicting the Löwdin transition charge on each atom from molecular conformations. Our strategy is generally the same as our model for E_m: a KRR model with Coulomb matrix features. However, here we train an individual KRR model per transition charge. This approach is similar to the one described by Brockherde et al.[233] as a way of predicting the ground-state one-electron density.

Even using the same model and features, the percent mean absolute error (%MAE) when predicting individual transition charges is higher than the %MAE when predicting the excitation energies (Table SII). This increase is to be expected however, both the transition density and transition charges vary greatly with small perturbations in the positions of the nuclei. Fortunately, the prediction quality when calculating the couplings from the transition charges improves dramatically, due to how this approach captures the $\frac{1}{r}$ dependence of the Coulomb interaction between charge densities exactly, as opposed to relying on the ML algorithm to determine this dependence on its own. The performance of this transition charge ML model for predicting V_{mn} is significantly better than the performance of the direct model. In Figure 6.9A we see that the error convergence is much improved, almost on par with the energy calculations, and the increase in error from cross-morphology predictions is negligible. Overall, when training with 10,000 examples, the MAE is only a small fraction of

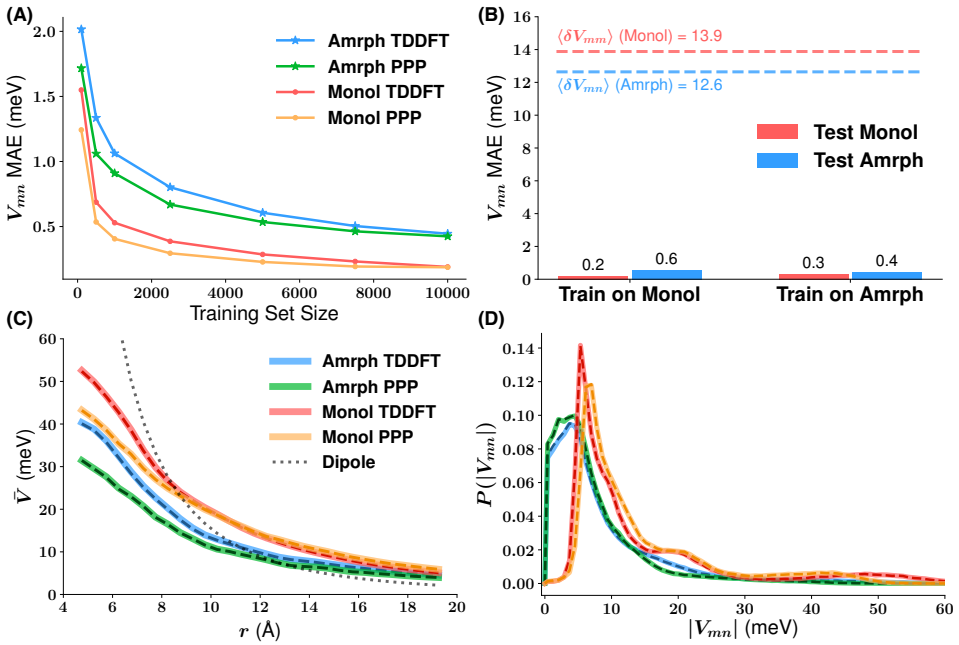

Figure 6.9: Comparing the error between actual transition charge couplings and machine-learned transition charge couplings. (A) Mean absolute error in V_{mn} predictions versus training set size used. (B) Cross-morphology error analysis with TDDFT data. (C) Average coupling versus center of geometry distance, solid lines are from exact transition charges, dashed lines are from ML predicted charges. (D) Couplings distributions.

the fluctuations in V_{mn} and roughly the same as the degree of intrinsic error from the Löwdin charge embedding. Thus, in stark contrast to the direct V_{mn} model, the transition charge V_{mn} model satisfies all of our ML performance criteria.

To further illustrate the high performance of this transition charge model, we analyzed the performance of the model at predicting the average distance dependence of V_{mn} as well as the probability distribution of (Figure 6.9C/D). In both cases, the agreement between ML results and actual values is almost exact, due to how the ML errors are largely random and these random errors are effectively averaged out. Such high accuracy is a very desirable property, as these distributions are an important visual tool for giving mechanistic insight into exciton transport and assessing why exciton diffusion rates differ between systems.

It is worth noting that the error analysis shown in Figure6.9 focuses on differences between the ML predictions and the actual transition charge coupling from Eq. 6.4, not the exact couplings from Eq. 6.3. Thus, in Figure 6.10), we compare the MAE of

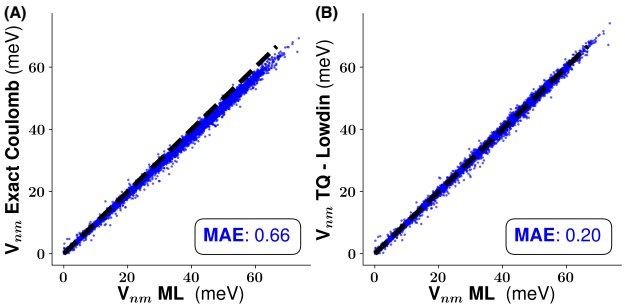

Figure 6.10: Parity plots for comparing accuracy of ML predictions to (A) the exact values, and (B) the values computed with TDDFT Coulomb couplings from Löwdin transition charges. All values are from monolayer simulation. ML model was trained using 10,000 molecular conformations.

the ML predicted couplings to the exact Coulomb couplings in the monolayer morphology. The error between the ML predicted transition charge couplings and exact Coulomb couplings is only 0.03meV greater than the error induced by the charge monopole approximation itself (Figure 6.5D). This result once again indicates the efficacy of the ML approach, as it highlights how the errors induced by the ML (when trained with sufficient data) are insignificant compared to the standard approximations already necessary for computing the electronic structure of extended condensed phase systems.

6.6.3 Trajectory Analysis

Now that we have described ML models capable of reproducing E_m and V_{mn} at high accuracy, we turn to using said models in tandem to predict a time-series of Frenkel Hamiltonians. As proof that our ML models can give accurate dynamics, we compare exciton trajectory statistics from the ML predicted Hamiltonians (using 10,000 training examples) to those directly from PPP/CIS calculations. In order to asses the transferability of the ML models we only include only conformations from the monolayer simulation in the training data. We do not test TDDFT here since the large time interval makes it difficult to accurately assess dynamics.

Exciton dynamics are propagated using a discrete set of time-ordered propagators

calculated at 10fs intervals from MD conformations:

$$|\Psi(n\Delta t)\rangle = \left[\prod_{j=0}^{n} \exp(-iH_j\Delta t/\hbar)\right]_+ |\Psi(0)\rangle\,, \tag{6.7}$$

where $\Psi(t)$ denotes the time dependent exciton wavefunction, Δt is the time-step interval and H_j represents the jth Frenkel Hamiltonian. We study two, experimentally observable, statistics of interest: the exciton mean-square displacement (MSD) and the lowest energy eigenvalue (LEE) of the Frenkel Hamiltonian. The MSD is defined as

$$\mathrm{MSD}(t) = \left\langle \sum_n |\mathbf{r}_n(t) - \mathbf{r}_m(0)|^2 |c_n(t)|^2 \right\rangle_m , \tag{6.8}$$

where $\mathbf{r}_n(t)$ is the center of mass position of molecule n at time t, and $|c_n(t)|^2$ is the probability of finding the the exciton on molecule n at time t. The angled brackets indicate an average over an ensemble of trajectories with the exciton initially localized on each different molecule. Without exciton delocalization, this MSD is equivalent to the usual MSD for particles.

In both cases, the predictions from the ML model match exact PPP/CIS results very well. There is virtually no difference in the MSD, as the ML predictions do not produce systematic errors, and the random errors are washed out when averaging over many trajectories. The agreement for the LEE is not as quantitative as the MSD, particularly for the amorphous system, but still quite strong as the PPP/CIS and ML lines are very well correlated. We believe the close fit between actual MSD/LEE, and the ML predictions illustrates that our model can not only predict individual E_m and V_{mn} elements accurately, but also gives very good predictions for the Frenkel Hamiltonian as a whole.

6.7 Conclusions

In this work we presented a novel ML strategy for generating Frenkel Hamiltonian trajectories from MD conformations. We built separate models for the diagonal exci-

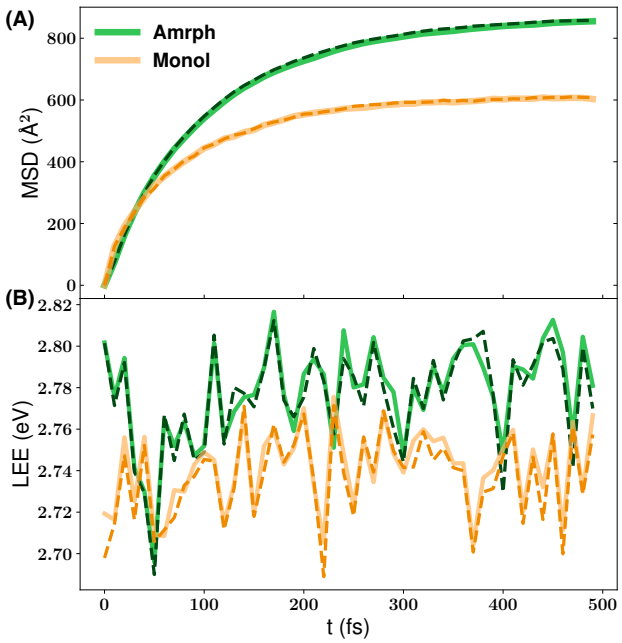

Figure 6.11: PPP/CIS trajectory statistics from monolayer and amorphous simulations. Solid lines are using exact calculations dashed lines are using ML model trained with 10,000 training examples selected only from the monolayer simulation. (A) Mean square displacement. (B) Lowest energy eigenvalue.

f

tation energies, E_m and off-diagonal couplings, V_{mn}, and tested the ML models against a set of four rigorous criteria: (1) applicability across electronic-structure methods, (2) lower errors than intrinsic uncertainty in the electronic-structure method, (3) lower errors than fluctuations from MD bath, and (4) transferability across morphologies. We showed that a simple KRR model with Coulomb matrix features is capable of satisfying all four criteria when predicting E_m. We then attempted to build a ML model to directly predict V_{mn} from two-molecule features, but illustrated that such an approach failed the performance criteria, likely due to the much larger size of the bimolecular conformational space. The failure of this approach motivated us to pursue building an ML model for V_{mn} from unimolecular features by predicting a set of atom-centered transition charges. We demonstrated that these transition charges reproduce the coupling quite poorly when assigned with Mulliken population analysis, but do much better when assigned with Löwdin population analysis. We then built another ML model, completely analogous to the E_m model, but for predicting

each individual transition-charge independently. We demonstrated that like the E_m model, it satisfied the performance criteria. While regression algorithms other than KRR may be proposed and utilized in the future, we hope that the insight into the relative ease of predicting a vectorial unimolecular quantity versus a scalar bimolecular quantity will prove to be generally helpful to the broader community interested in combining machine-learning and quantum chemistry.

The strong performance of both our E_m and transition charge V_{mn} models with 10,000 training examples suggests that an even smaller training set may be employed if necessary. However, 10,000 ES calculations is still orders of magnitude less than the number of calculations necessary to produce a single Frenkel Hamiltonian trajectory in many systems. In the future it would be interesting to expand the models built here using "active learning" methods [236, 237], where the ML algorithm retrains itself when encountering datapoints far outside the training set.

While we studied the application of our model only in T6 systems, we believe it may be applied in across a variety of systems where the Frenkel-Davydov model is appropriate. Of course, care must be taken in systems such as polymers, where individual fragments may have significant wavefunction overlap and exchange coupling is not negligible, as our transition charge V_{mn} model breaks down with significant exchange. However, it may be possible to build a very similar unimolecular V_{mn} model by predicting the entire transition density matrix using features taken from lower level electronic structure data, such as overlap and Fock matrices. We believe this is another exciting area for future development. It is also worth noting that it is dubious whether the assumptions made in the standard Frenkel-Davydov model are appropriate in strong-coupling situations where exchange is significant.

Finally, we note that parametrizing Frenkel Hamiltonians directly from MD conformations does neglect the backaction of the exciton on the nuclear bath. However, we hope that with the improved efficiency of our ML-Frenkel model, we may begin to address this limitation by introducing corrections to the MD forces based on the state of the exciton. We hope to further develop this ML-Frenkel protocol, and after careful validation against experiment, employ it to understand the effects of molec-

ular morphology and correlated environmental noise on exciton diffusion in organic semiconductors and guide organic semiconductor design.

Part IV

Conclusions

Chapter 7

Summary and outlook

In this final chapter, we provide a broad overview of the work done in this thesis, tying together central themes, and subsequently discussing future steps.

The central theme of this thesis has been the role that environmental correlations play in condensed-phase chemical dynamics. This theme is most apparent in Chapters 3 and 4, where we explicitly delineated how the time-correlations of the vibrations of a heterogeneous catalyst surface can affect reaction rates and sticking coefficients. However, the same theme is also present in Chapters 5 and 6. In Chapter 5, we used TPS/metadynamics to study the effects of the spatial correlations of the solvent degrees of freedom on desorption. And while Chapter 6 is primarily a methodological study of ML in application to exciton dynamics, the purpose of developing such a method was to study how spatiotemporal correlations of Frenkel Hamiltonians affect long exciton trajectories.

It is widely appreciated that chemistry in condensed-phase environments differs starkly from chemistry in gas-phase due to how the environment responds and reorganizes itself around the chemical system. The development of theoretical models which capture the effect of the environment on the chemical system is important not only for our fundamental understanding of chemistry, but also for our ability to deliberately design and engineer chemical reactivities.

One of the major difficulties in disentangling the correlations between chemical systems and their environments is the large time and length scales such correlations

can span. Such multiscale dynamics negate the effectiveness of brute-force simulation, but fortunately, there exist many solutions. The projection operator (generalized Langevin equation) approach we took in Chapters 3 and 4 is one way to develop a theory that can naturally span different length scales; as are the enhanced sampling methods we used in Chapter 5. And, of course, machine learning presents a new and very exciting way of making physical simulations much faster and more accurate. One exciting avenue of future research may be to combine machine learning with the Mori-Zwanzig projection operator method. While there already exist many approaches to machine-learning the potential of mean force, far less work has been done on machine-learning the memory integral/random forces.

Many open questions remain about the ways phonons couple to adsorbates. Our studies focused only on elemental solids, and the optical modes of polyatomic solids may add another interesting dimension to our results. We did not develop a far-from-equilibrium reaction rate theory for adsorbates responding to surface acoustic waves. Such a theory is key to definitively resolving the paradox of why and how surface acoustic waves promote reactivity and selectivity. We did not perform an extensive study on how surface-coverage and pressure affect the phonon-induced memory. Finally, our study focused on the effect of phonons for simple adsorption/desorption reactions, and it certainly would be interesting to study the effects on more complex reactions.

We note that much of the same generalized Langevin formalism we used to interrogate how adsorbates respond to phonons can also be used to examine how adsorbates respond to electric fields.

With regards to Frenkel exciton dynamics, we believe it would be very interesting to use the ML formalism described in this thesis to calculate time and spatial correlation functions for Frenkel Hamiltonians. Such correlation functions then could be fit to a Nakajima-Zwanzig generalized quantum master equation, in effect combining the methodologies of Chapters 3 and 4 with that of Chapter 6. Such a quantum master equation approach could yield critical insight into how excitons dissipate energy as they diffuse in organic thin films and how chemical functionalization can be harnessed

to mitigate exciton energy loss.

Appendix A

Non-linear Mori-Zwanzig theory

In Chapter 2.1 a derivation of the generalized Langevin equation was given for a system interacting harmonically with a bath of harmonic oscillators. In such a derivation, both the dynamics of the bath degrees of freedom and the projection operators themselves are linear. However, Mori-Zwanzig theory can be adapted to cases where the dynamics are non-linear and/or the projection operators are non-linear. In this appendix, a more generic derivation of the Mori-Zwanzig equations is given for non-linear equations of motion, with or without non-linear projection operators. The properties of the Mori-Zwanzig generalized Langevin equation under the action of such projection operators is discussed.

Let $\mathbf{\Gamma}$ be some vector of variables following a first-order differential equation,

$$\frac{d}{dt}\mathbf{\Gamma} = \mathcal{L}\mathbf{\Gamma} \tag{A.1}$$

where $\mathcal{L}$ a time-independent operator, though it need not be linear. Let the vector $\mathbf{A} = (A_1(\mathbf{\Gamma}), A_2(\mathbf{\Gamma}), \ldots, A_N(\mathbf{\Gamma}))$ be a vector of collective variables that are differential functions of $\mathbf{\Gamma}$. Suppose that $\mathbf{A} = \mathcal{P}\mathbf{\Gamma}$, where $\mathcal{P}$ is some idempotent operator. Like $\mathcal{L}$, $\mathcal{P}$ need not be linear. Also note that in general $\mathcal{L}$ and $\mathcal{P}$ do not commute. As we are only interested in the time-evolution of $\mathbf{A}$, we introduce $\mathcal{P}$ to into Eq. A.1,

$$\frac{d}{dt}\left(\mathcal{P}\mathbf{\Gamma}\right) = \mathcal{L}\mathcal{P}\mathbf{\Gamma}. \tag{A.2}$$

to arrive at,

$$\frac{d}{dt}\mathbf{A}(t) = e^{t\mathcal{L}}\mathcal{L}\mathbf{A}(t=0), \tag{A.3}$$

where we have explicitly factored the time-propagator, $e^{t\mathcal{L}}$, out of the left-hand side. Let $\mathcal{Q}$ be the orthogonal complement of $\mathcal{P}$, such that $\mathcal{Q} = \mathbb{1} - \mathcal{P}$, where $\mathbb{1}$ is the identity operator. We can insert this identity operator into Eq. A.3 to have,

$$\frac{d}{dt}\mathbf{A}(t) = e^{t\mathcal{L}}(\mathcal{P}+\mathcal{Q})\mathcal{L}\mathbf{A}(0). \tag{A.4}$$

We now introduce the Kawasaki identity:

$$e^{t\mathcal{L}} = e^{t\mathcal{Q}\mathcal{L}} + \int_0^t d\tau e^{(t-\tau)\mathcal{L}}\mathcal{P}\mathcal{L}e^{\tau\mathcal{Q}\mathcal{L}}, \tag{A.5}$$

which can be readily verified using integration by parts. While this identity may seem to arise from some deep mathematical insight, in reality we have done little more than add and subtract one from the original propagator. Using Eq. A.5 we may express Eq. A.4 as,

$$\frac{d}{dt}\mathbf{A}(t) = \mathcal{P}\mathcal{L}\mathbf{A}(t) + \int_0^t d\tau e^{(t-\tau)\mathcal{L}}\mathcal{P}\mathcal{L}e^{\tau\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}\mathbf{A}(0) + e^{t\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}\mathbf{A}(0) \tag{A.6}$$

Eq. A.6 is the fully-general Mori-Zwanzig "generalized-Langevin equation". It is worth noting that it an *exact* identity of the original equation of motion (Eq. A.1), that is, we have made no approximations in arriving at Eq. A.6, we have only re-arranged the terms in the original equation using the projection operators $\mathcal{P}$ and $\mathcal{Q}$. We identify the three terms on the left-hand side of as,

$$\mathbf{F}(t) = \mathcal{P}\mathcal{L}\mathcal{P}\mathbf{A}(t), \tag{A.7}$$

the projected generator of time-translations,

$$\mathbf{R}(t) = e^{t\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}\mathbf{A}(0), \tag{A.8}$$

the orthogonal dynamics and,

$$\mathbf{I}(t) = \int_0^t d\tau e^{(t-\tau)\mathcal{L}} \mathcal{P}\mathcal{L}\mathbf{R}(\tau) \tag{A.9}$$

the memory integral.

The derivation above is completely general, and can be applied to a range of problems in classical mechanics, quantum mechanics, and beyond. However, in order to make Eq. A.6 physically insightful and potential useful, we must first give specific forms for the operators $\mathcal{L}$, $\mathcal{P}$ and $\mathcal{Q}$. In a classical many-body system of N particles, $\mathbf{\Gamma}$ would be $6N$ the coordinates/momenta and $\mathcal{L}$ the classical Liouville operator,

$$\mathcal{L}A = \{H, A\} = \sum_j \frac{\partial H}{\partial p_j}\frac{\partial A}{\partial q_j} - \frac{\partial H}{\partial q_j}\frac{\partial A}{\partial p_j}. \tag{A.10}$$

where $\{,\}$ denotes the classical Poisson bracket, H is the classical Hamiltonian, p_j the momentum coordinates, and q_j the position coordinates. It should be noted that the classical Liouville operator is often defined with an additional imaginary factor, which makes the operator self-adjoint and therefore correspond with the quantum Liouville operator in phase-space formulations of quantum mechanics. However, for our purposes such an imaginary factor is unnecessary.

We now introduce Mori's projection operator and analyze it's concomitant effects on the terms in Eq. A.6. To define Mori's projection operator, once again let $\mathbf{A} = (A_1, A_2, \ldots, A_N)$ be a vector of the collective variables of interest and let $\mathbf{B} = (B_1, B_2, \ldots, B_M)$ be some other vector. Mori's projection operator is defined as,

$$\mathcal{P}_{\text{Mori}} \begin{pmatrix} B_1 \\ \vdots \\ B_M \end{pmatrix} = \frac{\langle \mathbf{B} \otimes \mathbf{A} \rangle}{\langle \mathbf{A} \otimes \mathbf{A} \rangle} \mathbf{A} = \begin{pmatrix} \frac{\langle B_1, A_1 \rangle}{\langle A_1, A_1 \rangle} & \frac{\langle B_1, A_2 \rangle}{\langle A_2, A_2 \rangle} & \cdots & \frac{\langle B_1, A_N \rangle}{\langle A_N, A_N \rangle} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\langle B_M, A_1 \rangle}{\langle A_1, A_1 \rangle} & \frac{\langle B_M, A_2 \rangle}{\langle A_2, A_2 \rangle} & \cdots & \frac{\langle B_M, A_N \rangle}{\langle A_N, A_N \rangle} \end{pmatrix} \begin{pmatrix} A_1 \\ \vdots \\ A_N \end{pmatrix} \tag{A.11}$$

where the angled brackets $\langle , \rangle$ denote an appropriately defined inner product. Usually

this inner-product is defined as the equilibrium correlation function,

$$\langle A_i(\mathbf{\Gamma})B_j(\mathbf{\Gamma})\rangle = \int d\mathbf{\Gamma} A_i(\mathbf{\Gamma})B_j(\mathbf{\Gamma}) \exp\left(-\beta H(\mathbf{\Gamma})\right), \tag{A.12}$$

and it is assumed that the components of $\mathbf{A}$ are orthogonal under this choice of inner product such that

$$\mathcal{P}_{\text{Mori}} \begin{pmatrix} A_1 \\ \vdots \\ A_N \end{pmatrix} = \begin{pmatrix} A_1 \\ \vdots \\ A_N \end{pmatrix}. \tag{A.13}$$

Mori's projection operator is always linear, even if $\mathbf{A}$ or $\mathbf{B}$ are non-linear collective coordinates.

With such a linear projection operator, both the projected "forces" $\mathbf{F}(t)$, and the memory integral become linear functions of $\mathbf{A}$. The projected forces take the form,

$$\mathbf{F}(t) = \frac{\langle \mathcal{L}\mathbf{A} \otimes \mathbf{A}\rangle}{\langle \mathbf{A} \otimes \mathbf{A}\rangle}\mathbf{A}(t) \tag{A.14}$$

which correspond to a harmonic force in $\mathbf{A}$ space, with $\frac{\langle \mathcal{L}\mathbf{A}\otimes\mathbf{A}\rangle}{\langle \mathbf{A}\otimes\mathbf{A}\rangle}$ being the force-constant matrix. Likewise, the memory integral can be written in the form,

$$\mathbf{I}(t) = \int_0^t d\tau e^{(t-\tau)\mathcal{L}} \frac{\langle \mathcal{L}\mathbf{R}(\tau) \otimes \mathbf{A}(0)\rangle}{\langle \mathbf{A} \otimes \mathbf{A}\rangle}\mathbf{A}(0). \tag{A.15}$$

As the matrix $\frac{\langle \mathcal{L}\mathbf{R}(\tau)\otimes\mathbf{A}(0)\rangle}{\langle \mathbf{A}\otimes\mathbf{A}\rangle}$ has no explicit dependence on phase-space variables by virtue of the integral in Eq. A.12, it commutes with the Liouvillian and the time-translation operator. Thus $\mathbf{I}(t)$ may be expressed as,

$$\mathbf{I}(t) = \int_0^t d\tau \frac{\langle \mathcal{L}\mathbf{R}(\tau) \otimes \mathbf{A}\rangle}{\langle \mathbf{A} \otimes \mathbf{A}\rangle}\mathbf{A}(t-\tau). \tag{A.16}$$

Note that since $\mathcal{Q}$ is idempotent, $\mathcal{Q}\mathbf{R} = \mathbf{R}$. Also note that $\mathcal{Q}$ is hermitian under the action of Eq. A.12 and $\mathcal{L}$ is anti-hermitian. With these identities we may express the

memory integral as,

$$\mathbf{I}(t) = -\int_0^t d\tau \frac{\langle \mathbf{R}(\tau) \otimes \mathbf{Q}\mathcal{L}\mathbf{A} \rangle}{\langle \mathbf{A}\mathbf{A} \rangle} \mathbf{A}(t-\tau), \tag{A.17}$$

$$\mathbf{I}(t) = -\int_0^t d\tau \frac{\langle \mathbf{R}(\tau) \otimes \mathbf{R}(0) \rangle}{\langle \mathbf{A} \otimes \mathbf{A} \rangle} \mathbf{A}(t-\tau) = -\int_0^t d\tau \mathbf{K}(t-\tau)\mathbf{A}(\tau) \tag{A.18}$$

where we see that the memory kernel matrix, $\mathbf{K}(t) = \frac{\langle \mathbf{R}(\tau) \otimes \mathbf{R}(0) \rangle}{\langle \mathbf{A}, \mathbf{A} \rangle}$, is an autocorrelation function of the orthogonal dynamics. This fluctuation-dissipation relationship is precisely the same was as derived in Chapter 2 for a system interacting with a harmonic bath. In fact, if we let $\mathbf{A}$ be a simple subset of the phase space variables and linearized $\mathcal{L}$, we could analytically evaluate the orthogonal dynamics and derive exactly the same generalized Langevin equation as in Chapter 2. However, in the case where the potential energy is anharmonic and therefore $\mathcal{L}$ is non-linear, the orthogonal dynamics cannot be solved exactly. Indeed, as both $\mathbf{F}(t)$ and $\mathbf{I}(t)$ are linear under Mori's projection operator, the effect of all non-linearities in $\mathcal{L}$ are pushed into $\mathbf{R}(t)$. For highly anharmonic dynamics, it is not clear to what extent and under what circumstances $\mathbf{R}(t)$ can be interpreted as "random-noise"[51].

While $\mathbf{F}(t)$ is linear under Mori's projection operator, in many practical application we may seek a form of Eq. A.6 such that $\mathbf{F}(t)$ is a non-linear function of the collective coordinates $\mathbf{A}$. For example, it is often convenient for $\mathbf{F}$ to correspond to the equilibrium fluctuations generated by the potential of mean force (Eq. 2.22), so that the memory integral and orthogonal dynamics would correspond to the non-equilibrium dissipation and fluctuations respectively. For a discussion of the derivation of such a generalized Langevin equation using the Mori-Zwanzig formalism, see Ref.54.

Appendix B

Calculating memory kernels from time-correlation functions

As discussed in Chapter 3, while the generalized Langevin equation is most easily derived for a system interacting with a harmonic bath, it can applied to baths which have anharmonicities, so long as they remain near thermal equilibrium. The question arises then, how to compute the memory kernel when the bath exhibits anharmonicities, given that we cannot resolve the bath dynamics analytically. One popular solution is to calculate the memory kernel from time-correlation functions computed using molecular dynamics (MD) simulation. This method begins from recognizing that the random force $R(t)$ must be uncorrelated with the system's momenta: $\langle R(t)p(0)\rangle = 0$. This identity can be considered prerequisite for $R(t)$ to be properly interpreted as a "random" noise, and is indeed consistent with Eq. 2.19.

By taking the equilibrium time correlation function of both sides of the generalized Langevin equation with the initial momentum we find,

$$\langle \dot{p}(t)p(0)\rangle + \left\langle \frac{dW}{dq}(t)p(0) \right\rangle = -\int_0^t K(t-\tau)\,\langle p(\tau)p(0)\rangle\, d\tau. \tag{B.1}$$

Using MD simulation the force-momentum correlation functions (left-hand side) and the momentum autocorrelation function may be computed, and subsequently Eq. B.1 may be solved to find $K(t)$. One may be tempted to use the convolution theorem and

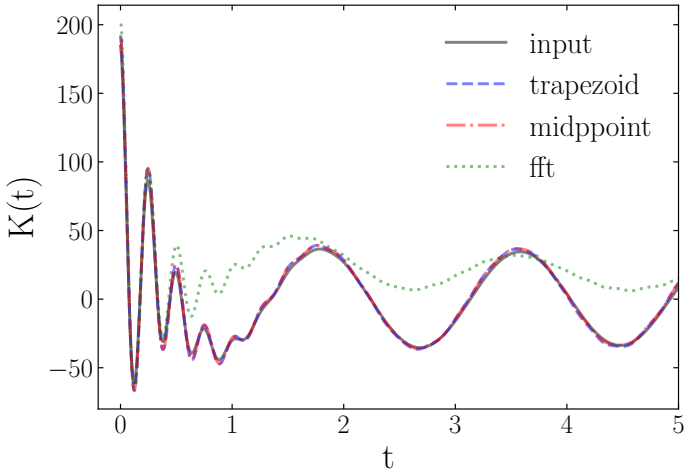

Figure B.1: Memory kernel calculated with by using midpoint quadrature, trapezoidal quadrature, fast Fourier transform methods to solve the Volterra equation versus the input (target) memory kernel used to conduct the simulation. The trapezoidal quadrature used the derivative of Eq. B.1.

fast-Fourier transform to solve Eq. B.1, however such an approach has been shown to lack numerical accuracy [238, 239]. Instead, most approaches to solve Eq. B.1 use real-time iterative methods. Such approaches discretize the memory integral using a numerical quadrature and solving the system of linear equations that results. A comparison of the fast-Fourier transform, midpoint quadrature, and trapezoidal quadrature methods is given in Figure B.1 for a harmonic oscillator interacting with a bath whose memory kernel is a sum of two-exponentially damped sinusoids. We see that both the trapezoidal and midpoint quadratures work well at reproducing the exact/input memory kernel, but that FFT method fails at longer times.

Below we illustrate one approach to calculating the memory kernel using a trapezoidal quadrature. To begin, let us rewrite Eq. B.1 in the form,

$$C_f(t) = -\int_0^t K(t-\tau)C_p(\tau)d\tau, \tag{B.2}$$

where $C_f(t)$ is the force-momentum correlation function and $C_p(t)$ the momentum autocorrelation function. We take the derivative of Eq. B.2, as so gives an equation

which is better defined and more numerically stable for small times t,

$$\dot{C}_f(t) = -K(t)C_p(0) - \int_0^t K(t-\tau)\dot{C}_p(\tau)d\tau. \tag{B.3}$$

We then introduce the trapezoidal quadrature in order to evaluate the memory integral at discrete timesteps Δt,

$$\begin{aligned}
\dot{C}_f(t=0) &= -K(0)C_p(0), \\
\dot{C}_f(t=\Delta t) &= -K(\Delta t)C_p(0) - \frac{\Delta t}{2}\left[K(\Delta t)\dot{C}_p(0) + K(0)\dot{C}_p(\Delta t)\right], \\
\dot{C}_f(t=N\Delta t) &= -K(N\Delta t)C_p(0) - \frac{\Delta t}{2}\left[K(N\Delta t)\dot{C}_p(0) + K(0)\dot{C}_p(N\Delta t)\right] \\
&\quad - \Delta t\sum_{n=1}^{N-1} K((N-n)\Delta t)\dot{C}_p(n\Delta t).
\end{aligned} \tag{B.4}$$

Equation B.4 was used to compute the blue dashed curve in Figure B.1.

Appendix C

Elastic continuum theory

In this appendix, we take an alternative approach to determining the phonon-induced memory in the limit of a macroscopic solid by using continuum elastic theory instead of atomistic models. Such an approach was originally explored by Ref. 240 and Ref. 241. It is theoretically appealing due to the minimal, experimentally accessible free parameters used in elastic theory. However, it also suffers from several limiting assumptions which we explicitly delineate.

Elastic (energy conserving) acoustic waves in a material may be modeled via the Navier-Cauchy equation,

$$\ddot{\mathbf{u}}(\mathbf{r},t) = c_t\vec{\nabla}^2\mathbf{u}(\mathbf{r},t) + (c_l^2 + c_t^2)\nabla\left(\nabla\cdot\mathbf{u}(\mathbf{r},t)\right) + \mathbf{F}(\mathbf{r},t), \tag{C.1}$$

where $\mathbf{u}(\mathbf{r},t)$ is the displacement of the solid at position $\mathbf{r} = (x,y,z)$ and time t, $\vec{\nabla}^2$ is the vector Laplacian, $\nabla\left(\nabla\cdot\right)$ is gradient of the divergence, and $\mathbf{F}$ are external forces. Solutions to this Eq. C.1 can generally be separated into zero divergence and zero curl components corresponding to the transverse and longitudinal modes respectively,

$$\ddot{\mathbf{u}}(\mathbf{x},t) = \mathbf{u}_l(\mathbf{x},t) + \mathbf{u}_t(\mathbf{x},t), \tag{C.2}$$

each of which satisfy a 3D wave equation,

$$\ddot{\mathbf{u}}_{l/t}(\mathbf{x},t) = c_{l/t}\vec{\nabla}^2\mathbf{u}_{l/t}(\mathbf{x},t). \tag{C.3}$$

Consider a single adsorbate degree of freedom whose displacement is denoted by q. If this degree of freedom is harmonically coupled to surface normal (z-axis) displacement the solid at position $\mathbf{r}_0 = (x_0, y_0, L_z)$ then the coupled adsorbate-solid equations are,

$$\ddot{q}(t) = -\frac{1}{m}\frac{dV_A}{dq}(t) - \frac{\mu}{m}\omega_{as}^2(q(t) - u_z(\mathbf{r}_0, t)) \tag{C.4}$$

$$\ddot{\mathbf{u}}(\mathbf{r}, t) - c_t\vec{\nabla}^2\mathbf{u}(\mathbf{r}, t) - (c_l^2 + c_t^2)\nabla\left(\nabla \cdot \mathbf{u}(\mathbf{r}, t)\right) = \frac{\mu\omega_{as}^2}{M}(q(t) - u_z(\mathbf{r}_0, t))a^3\delta\left(\mathbf{r} - \mathbf{r}_0\right)\hat{z}, \tag{C.5}$$

where $\hat{z} = (0, 0, 1)$ is the unit vector in the surface normal, a is spacing between atoms in the crystal, and M is the mass of the solid atom. The a^3 factor arises from taking the continuum limit of a force on a single lattice point and offsets the inverse volume units of the 3D delta function $\delta\left(\mathbf{r} - \mathbf{r}_0\right)$. The forces from the adsorbate on the solid can be be separated in two contributions. A static contribution,

$$\frac{\mu\omega_{as}^2}{M}u_z(\mathbf{r}_0, t)a^3\delta\left(\mathbf{r} - \mathbf{r}_0\right), \tag{C.6}$$

which enforces to the shift in the solid's vibrational spectrum due to the presence of the adsorbate. And, a dynamic contribution,

$$f(\mathbf{r}, t) = \frac{\mu\omega_{as}^2}{M}q(t)a^3\delta\left(\mathbf{r} - \mathbf{r}_0\right), \tag{C.7}$$

representing the time-dependent external force of the adsorbate on the solid. We can rearrange Eq. C.5 as,

$$\left[\frac{d^2}{dt^2} - c_t^2\vec{\nabla}^2 - (c_l^2 + c_t^2)\nabla\nabla\cdot + \frac{\mu\omega_{as}^2}{M}a^3\delta(\mathbf{r} - \mathbf{r}_0)\hat{z}\right]\mathbf{u}(\mathbf{r}, t) = f(\mathbf{r}, t)\hat{z}. \tag{C.8}$$

The operator on the right-hand-side (RHS) of this equation is a linear operator; therefore Eq. C.8 may be solved using the method of Green's functions,

$$\mathbf{u}(\mathbf{r}, t) = \mathbf{u}_0(\mathbf{r}, t) + \int_0^t d\tau \int d\mathbf{r}'\mathbf{G}(\mathbf{r}, t; \mathbf{r}', \tau) \cdot f(\mathbf{r}', \tau)\hat{z} \tag{C.9}$$

where $\mathbf{u}_0(x,t)$ is the solution to the homogeneous case and $\mathbf{G}(\mathbf{r},t;\mathbf{r}',\tau)$ is a 3x3 tensor Green's function corresponding to the operator on the RHS of Eq. C.8. If we substitute this solution into Eq. C.4 we arrive at a GLE where the memory kernel is proportional to the antiderivative of the Green's function,

$$K(t) = \frac{\mu^2\omega_{as}^4}{mM}a^3 \int dt G_{zz}(\mathbf{r}_0,t;\mathbf{r}_0,0). \tag{C.10}$$

Henceforth we will denote $G_{zz}(\mathbf{r}_0,t;\mathbf{r}_0,0)$ as simply $G(t)$ for simplicity. This Green's function may be decomposed in the following form,

$$G(t) = \sum_{\alpha}\sum_{\mathbf{k}} \frac{\sin(c_\alpha |\mathbf{k}| t)}{c_\alpha |\mathbf{k}|} R_{z,\alpha}(\mathbf{r}_0,\mathbf{k}) R^*_{z,\alpha}(\mathbf{r}_0,\mathbf{k}) \tag{C.11}$$

where α denotes phonon polarizations (i.e. transverse or longitudinal), and $R_{z,\alpha}$ are the zth spatial components of the normalized eigenfunctions of the operator on the RHS of Eq. C.8. The spectrum of $\mathbf{k}$ values as well as the specific form of the spatial eigenfunctions depend on the choice of boundary conditions.

Due to the delta function in the adsorbate shift term (Eq. C.6), the allowed $\mathbf{k}$ values and cannot be computed exactly. Indeed, this term makes Eq. C.8 very similar to the Schrödinger Equation with a delta function well, in which the spectrum must be computed numerically as a solution to a system of transcendental equations. However, perturbation theory, physical intuition, and the numerical results presented in Figure 1 of the main text all suggest that the low-frequency acoustic modes of a solid should not depend on the presence of an adsorbate. Therefore, we proceed by ignoring the adsorbate shift term while noting that, by construction, such an approach is only valid for the low-frequency acoustic modes. Setting periodic boundary conditions in the xy plane, fixed boundary conditions at $z = 0$ and Neumann boundary conditions at $z = L_z$ we have,

$$R^*_{z,\alpha}(\mathbf{r},\mathbf{k}) = \frac{2}{\sqrt{L_x L_y L_z}} e^{2\pi i k_x x} e^{2\pi i k_y y} \sin(k_z z) \tag{C.12}$$

where L_x, L_y, and L_z are the size of the solid in the x, y, and z directions respectively.

The allowed are values of k are $k_x = \frac{2\pi n_x}{L_x}$, $k_y = \frac{2\pi n_y}{L_y}$, and $k_z = \frac{(n_z+\frac{1}{2})\pi}{L_z}$, where n_x and n_y are any integers and n_z is any integer greater than or equal to zero. Taking the limit as L_x, L_y, and L_z become very large we find that the Green's function becomes,

$$G(t) = \sum_{\alpha} \frac{1}{8\pi^3} \int d\mathbf{k} \frac{\sin(c_\alpha |\mathbf{k}| t)}{c_\alpha |\mathbf{k}|}. \tag{C.13}$$

It is well-known that the integral in Eq. C.13 diverges if the integral is taken over all k-space due to the contribution wavelengths smaller than the inter-atom spacing. Therefore, the integral in Eq. C.13 should only be taken over first Brillouin zone. Taking inspiration from the Debye model, we may approximate the first Brillouin zone with a radial cutoff $|k_\mathrm{D}|$,

$$G(t) = \frac{1}{2\pi^2} \sum_{\alpha} \int_0^{k_\mathrm{D}} dk \frac{\sin(c_\alpha k t)}{c_\alpha k}. \tag{C.14}$$

Carrying out the integration over k and subsequently integrating over time t, leads to the following formulas for the memory kernel and spectral density,

$$K_{\mathrm{cont}}(t) = \frac{\mu^2 \omega_{\mathrm{as}}^4}{mM} \frac{a^3}{2\pi^2} \left(\frac{2}{c_t^3} + \frac{1}{c_l^3} \right) \frac{\sin(\omega_\mathrm{D} t)}{t}. \tag{C.15}$$

$$\bar{K}_{\mathrm{cont}(\omega)} = \frac{\mu^2 \omega_{\mathrm{as}}^4}{mM} \frac{a^3}{2\pi^2} \left(\frac{2}{c_t^3} + \frac{1}{c_l^3} \right) \Theta(\omega - \omega_\mathrm{D}). \tag{C.16}$$

where Θ is the Heaviside step function. Eq. C.16 illustrates that the memory is flat (Ohmic) with a high frequency cutoff at the Debye frequency.

We can use Eq C.15 to compute the effective adsorbate-surface bond frequency,

$$\tilde{\omega}_{\mathrm{as}} = \sqrt{\frac{\mu}{m}\omega_{\mathrm{as}}^2 - K(t=0)}. \tag{C.17}$$

Figure C.1 illustrates results for $\tilde{\omega}_{\mathrm{as}}^2$ as a function of the Debye frequency and ω_{as}. One can see that for physically reasonable choices of parameters (including those for CO adsorbed on Pt) the effective frequency becomes imaginary, signifying that there is no stable adsorption state. Such a clearly unphysical conclusion is a result

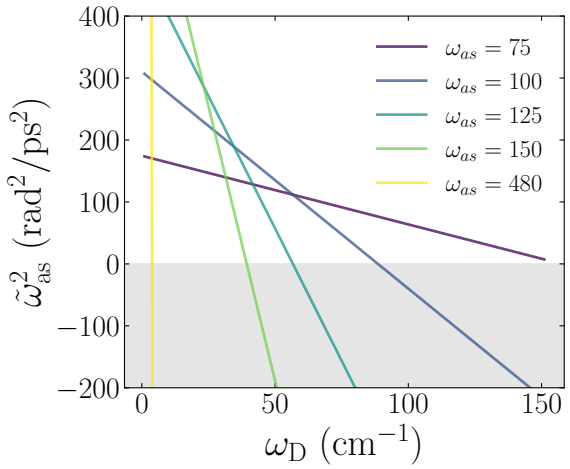

Figure C.1: Square of the effective (phonon-renormalized) surface-adsorbate bond frequency as a function of the solid's Debye frequency.

of the ignoring the adsorbate shift term (Eq. C.6) in the original equations of motion and its concomitant effects on the Green's function/memory kernel. Thus, we again emphasize that the functional forms in Eq. C.15 and Eq. C.16 are only valid for the low-frequency acoustic modes of the solid, and generally integrating all **k** vectors up to the Debye frequency is inappropriate and a lower frequency cutoff should be used.

Appendix D

Derivation of extended variable transformation

In this appendix we give a derivation of the extended-variable transformation, which is used to map the non-Markovian dynamics of a generalization Langevin equation back into a set of Markovian equations describing a system coupled to a multivariate Ornstein-Uhlenbeck (OU) bath. We begin by giving an overview of the OU processes, focusing on the calculation of the two-time autocorrelation function (covariance). Subsequently, we use the OU autocorrelation function to ensure that the dynamics of the extended-variables obey the fluctuation-dissipation theorem.

D.1 Ornstein-Uhlenbeck processes

The Ornstein-Uhlenbeck process is a stationary, Gauss-Markov process, which in 1-dimension, may be described as the solution to the linear stochastic differential equation (SDE),

$$\dot{x}(t) = -\lambda x(t) + \sigma dW(t), \tag{D.1}$$

where λ and σ are parameters of the process and $dW(t)$ is white-noise satisfying, $\langle dW(t)dW(0)\rangle = \delta(t)$. The Langevin equation in the absence of an external potential is an example of an 1-D Ornstein-Uhlenbeck process. Eq. D.1 may be formally solved

to yield,

$$x(t) = e^{-\lambda t}x(0) + \int_0^t d\tau e^{-\lambda(t-\tau)}\sigma dW(\tau). \tag{D.2}$$

Since the noise in Eq. D.1 is additive, the Itô and Stratonovich interpretation of the integral in Eq. D.2 are equivalent, and we may ignore the mathematical subtleties of how to define integrals for stochastic processes. From Eq. D.2 we can immediately see that the average value of x decays exponentially, $\langle x(t)\rangle = e^{-\lambda t}\langle x(0)\rangle$. Taking the two-time correlation yields,

$$\langle x(t+s)x(s)\rangle = e^{-\lambda(t+2s)}\langle x(0)^2\rangle + \sigma^2\int_0^s\int_0^{t+s} d\tau d\tau' e^{-\lambda(t+2s-\tau-\tau')}\langle dW(\tau)dW(\tau')\rangle, \tag{D.3}$$

$$\langle x(t+s)x(s)\rangle = e^{-\lambda(t+2s)}\left(\langle x(0)^2\rangle + \frac{\sigma^2}{2\lambda}\right) + \frac{\sigma^2}{2\lambda}e^{-\lambda t}, \tag{D.4}$$

where we see that that, in order for the time-correlation function to not depend on the absolute time s, and only on the time difference t (in order words in other for the process to be stationary), $x(0)$ must satisfy the Lyapunov stability criterion,

$$\langle x(0)^2\rangle = \frac{\sigma^2}{2\lambda}. \tag{D.5}$$

This equation is essentially a fluctuation-dissipation relationship; for a Langevin equation $\sigma^2 \propto k_{\mathrm{B}}T$ are the thermal fluctuations, and λ is the Markovian friction. Given Eq.D.5, the time correlation function becomes,

$$\langle x(t+s)x(s)\rangle = \frac{\sigma^2}{2\lambda}e^{-\lambda t}. \tag{D.6}$$

Analogously, for a multivariate OU process the SDE may be expressed as,

$$\dot{\mathbf{x}}(t) = -\mathbf{\Lambda}\mathbf{x}(t) + \mathbf{\Sigma}d\mathbf{W}(t), \tag{D.7}$$

where $d\mathbf{W}(t)$ satisfies $\langle dW_i(t)dW_j(0)\rangle = \delta_{ij}\delta(t)$. The underdamped Langevin equation for a harmonic oscillator is an example of a multivariate OU process. Following a similar procedure as the 1D case, we may solve for the time correlation function

yielding,

$$\left\langle \mathbf{x}(t+s)\mathbf{x}^T(s)\right\rangle = e^{-\mathbf{\Lambda}(t+s)}\langle \mathbf{x}(0)\mathbf{x}(0)^T\rangle e^{-\mathbf{\Lambda}^T s} + \int_0^s d\tau e^{-\mathbf{\Lambda}(t+s-\tau)}\mathbf{\Sigma}\mathbf{\Sigma}^T e^{-\mathbf{\Lambda^T}(s-\tau)}. \quad \text{(D.8)}$$

It is important to note that, in general $\mathbf{\Lambda}$, $\mathbf{\Lambda}^T$, and $\mathbf{\Sigma}\mathbf{\Sigma}^T$ do not necessarily commute with one another. Let $\mathbf{c}(t,s) = \left\langle \mathbf{x}(t+s)\mathbf{x}^T(s)\right\rangle$; we can use the following trick in order to simplify Eq. D.8,

$$\mathbf{\Lambda c}(t,s) + \mathbf{c}(t,s)\mathbf{\Lambda}^T = e^{-\mathbf{\Lambda}(t+s)}\left[\mathbf{\Lambda c}(0) + \mathbf{c}(0)\mathbf{\Lambda}^T\right] e^{-\mathbf{\Lambda}^T s} + \int_0^s d\tau \frac{d}{d\tau}\left[e^{-\mathbf{\Lambda}(t+s-\tau)}\mathbf{\Sigma}\mathbf{\Sigma}^T e^{-\mathbf{\Lambda^T}(s-\tau)}\right], \quad \text{(D.9)}$$

$$\mathbf{\Lambda c}(t,s) + \mathbf{c}(t,s)\mathbf{\Lambda}^T = e^{-\mathbf{\Lambda}(t+s)}\left[\mathbf{\Lambda c}(0) + \mathbf{c}(0)\mathbf{\Lambda}^T - \mathbf{\Sigma}\mathbf{\Sigma}^T\right] e^{-\mathbf{\Lambda}^T(s)} + e^{-\mathbf{\Lambda}t}\mathbf{\Sigma}\mathbf{\Sigma}^T. \quad \text{(D.10)}$$

From Eq. D.10 we find that the Lyapunov equation is,

$$\mathbf{\Lambda c}(0) + \mathbf{c}(0)\mathbf{\Lambda}^T = \mathbf{\Sigma}\mathbf{\Sigma}^T, \quad \text{(D.11)}$$

and the corresponding stationary autocorrelation function is given by,

$$\mathbf{\Lambda c}(t,s) + \mathbf{c}(t,s)\mathbf{\Lambda}^T = e^{-\mathbf{\Lambda}t}\mathbf{\Sigma}\mathbf{\Sigma}^T. \quad \text{(D.12)}$$

Under the special case where $\mathbf{\Sigma}\mathbf{\Sigma}^T = k_\mathrm{B}T(\mathbf{\Lambda}+\mathbf{\Lambda}^T)$, Eq. D.12 may be further simplified to,

$$\mathbf{c}(t,s) = k_\mathrm{B}T e^{-\mathbf{\Lambda}t}. \quad \text{(D.13)}$$

In the following section, we will use Eq. D.13 to ensure that the dynamics of the OU bath encode the fluctuation-dissipation relationship for the GLE they are derived from.

D.2 Mapping to a multivariate OU bath

Suppose that the memory kernel can be expanded in the form,

$$\mathbf{K}(t) = \mathbf{\Lambda}_{\mathrm{AB}} \exp\left(-\mathbf{\Lambda}_{\mathrm{BB}} t\right) \mathbf{\Lambda}_{\mathrm{BA}} \tag{D.14}$$

such that the corresponding GLE becomes,

$$\ddot{\mathbf{x}}_{\mathrm{A}}(t) = -\frac{\partial W}{\partial \mathbf{x}_{\mathrm{A}}}(t) - \int_0^\infty \mathbf{\Lambda}_{\mathrm{AB}} \exp\left(-\mathbf{\Lambda}_{\mathrm{BB}}(t-\tau)\right) \mathbf{\Lambda}_{\mathrm{BA}} \dot{\mathbf{x}}_{\mathrm{A}}(\tau) + \mathbf{R}(t), \tag{D.15}$$

$$\ddot{\mathbf{x}}_{\mathrm{A}}(t) = -\frac{\partial W}{\partial \mathbf{x}_{\mathrm{A}}}(t) - \mathbf{\Lambda}_{\mathrm{AB}} \left[\mathbf{Z}_{\mathrm{B}}(t) + \mathbf{R}_{\mathrm{B}}(t)\right], \tag{D.16}$$

where $\mathbf{Z}_{\mathrm{B}}(t)$ and $\mathbf{R}_{\mathrm{B}}(t)$ denote the memory integral and random forces with $\mathbf{\Lambda}_{\mathrm{AB}}$ factored out, respectively. Taking the derivative of $\mathbf{Z}_{\mathrm{B}}(t)$ gives,

$$\dot{\mathbf{Z}}_{\mathrm{B}}(t) = -\mathbf{\Lambda}_{\mathrm{BA}} \dot{\mathbf{x}}_{\mathrm{A}}(t) - \mathbf{\Lambda}_{\mathrm{BB}} \mathbf{Z}_{\mathrm{B}}(t). \tag{D.17}$$

Let $\mathbf{R}_{\mathrm{B}}(t)$ be described as multivariate Ornstein-Uhlenbeck process satisfying the SDE,

$$\dot{\mathbf{R}}_{\mathrm{B}}(t) = -\mathbf{\Lambda}_{\mathrm{BB}} \mathbf{R}_{\mathrm{B}}(t) + \mathbf{\Sigma}_{\mathrm{BB}} d\mathbf{W}(t), \tag{D.18}$$

where $\mathbf{\Sigma}_{\mathrm{BB}}$ is a matrix that must be selected such that $\mathbf{R}(t)$ and $\mathbf{K}(t)$ obey the fluctuation-dissipation theorem (Eq. 2.20). Following Eq. D.12 and Eq. D.13, this condition can be satisfied by selecting $\mathbf{\Sigma}_{\mathrm{BB}}$ such that it satisfies $\mathbf{\Sigma}_{\mathrm{BB}} \mathbf{\Sigma}_{\mathrm{BB}}^T = k_{\mathrm{B}} T (\mathbf{\Lambda}_{\mathrm{BB}} + \mathbf{\Lambda}_{\mathrm{BB}}^T$. Note that numerically, $\mathbf{\Sigma}_{\mathrm{BB}}$ may be calculated via Cholesky decomposition.

We not introduce the extended-variables $\mathbf{b}(t) = \mathbf{Z}_{\mathrm{B}}(t) + \mathbf{R}_{\mathrm{B}}(t)$, which satisfy the equation of motion,

$$\mathbf{b}(t) = -\mathbf{\Lambda}_{\mathrm{BA}} \dot{\mathbf{x}}_{\mathrm{A}}(t) - \mathbf{\Lambda}_{\mathrm{BB}} \mathbf{b}(t) + \mathbf{\Sigma}_{\mathrm{BB}} d\mathbf{W}(t). \tag{D.19}$$

Correspondingly Eq. D.16 becomes,

$$\ddot{\mathbf{x}}_{\mathrm{A}}(t) = -\frac{\partial W}{\partial \mathbf{x}_{\mathrm{A}}}(t) - \mathbf{\Lambda}_{\mathrm{AB}}\mathbf{b}(t). \tag{D.20}$$

The equations above complete the derivation of the extended-variable transformation. This expansion is most useful when the memory kernel behaves like a sum of a few exponentially damped sinusoids. For power-law memory kernels, in principle infinitely large $\mathbf{\Lambda}_{\mathrm{BB}}$ and $\mathbf{\Sigma}_{\mathrm{BB}}$ are needed, negating any computational advantages that arise from avoiding the explicit computation of the memory integral.

Appendix E

Continuum limit of 1D harmonic chain

Consider a 1D chain of harmonic oscillators with spring constant $k = m\omega^2$ and periodic boundary conditions. Every site in the chain is identical, and the dynamical matrix is given by,

$$\mathbf{D}^2 = \omega^2 \begin{pmatrix} -2 & 1 & 0 & 0 & \dots & -1 \\ -1 & 2 & -1 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & \dots & 0 & -1 & 2 & -1 \\ -1 & \dots & 0 & 0 & -1 & 2 \end{pmatrix}. \tag{E.1}$$

Taking our system to be a single site in the lattice, the resulting bath projected matrix is given by

$$\mathbf{D}^2_{\mathrm{BB}} = \omega^2 \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & -1 & 2 & -1 \\ 0 & \dots & 0 & -1 & 2 \end{pmatrix}. \tag{E.2}$$

This matrix may be diagonalized analytically allowing one to find a solution to the memory kernel,

$$K(t) = \frac{4k}{N} \sum_{n=1}^{N} \cos^2(\theta_n) \cos\left(2\omega t \sin(\theta_n)\right), \tag{E.3}$$

where N is the total length of the chain and $\theta_n = \frac{n\pi}{2(N+1)}$. If we take the limit as $N \to \infty$, we see that this sum converges to an integral,

$$K(t) = 8\pi k \int_0^{\pi/2} d\theta \cos^2(\theta) \cos\left(2\omega t \sin(\theta)\right). \tag{E.4}$$

This integral has no closed form solution. However it can be expressed in terms of Bessel functions,

$$K(t) = \frac{4\omega^2}{\pi} \frac{J_1(2\omega t)}{t}, \tag{E.5}$$

where J is a Bessel function of the first kind.

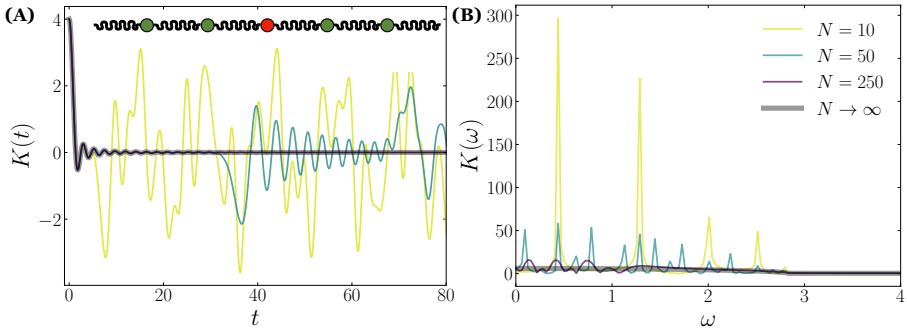

Figure E.1: Comparison of memory kernel (A) and power spectra (B) for single site fluctuations of 1D harmonic chains of various lengths with periodic boundary conditions.

In Figure E.1 we illustrate the size dependence of the memory kernel for a site in a 1D chain. Like the 3D lattices presented in the main text, there is a frequency shift as we move to increase the size of the chain. However, the power spectra is not bimodal, but rather a continuous sum of many modes which decrease in amplitude as we approach the chain's Debye frequency 2ω. Furthermore, the memory kernel also does not decay exponentially, but rather as $\frac{1}{t}$.

Appendix F

Derivation of Frenkel-Davydov model

The Frenkel-Davydov model supposes that the excited-state electronic structure of a supermolecular system may be expressed in terms of a reduced excitonic Hamiltonian,

$$\hat{H}_{\mathrm{F}} = \sum_{m} E_m \left|m\right\rangle \left\langle m\right| + \sum_{m \neq n} V_{mn} \left|m\right\rangle \left\langle n\right|, \tag{F.1}$$

where diagonal elements E_m represent the excitation energy of the exciton localized on molecule m and V_{mn} is the two-molecule excitonic coupling. While this approximation may seem severe, studies have shown that the Frenkel-Davydov model compares well to supermolecular time-dependent density function theory (TDDFT) calculations [182].

In order to relate the matrix elements of Eq. F.1 to molecular electronic structure, one may project the exact many-electron Hamiltonian, $\hat{H}_{\mathrm{el}}$, into a product basis of single excitations on each molecule,

$$\left|m\right\rangle = \left|\Psi_{\mathrm{ex}}^{(m)}\right\rangle \prod_{n \neq m}^{N_{\mathrm{mol}}} \left|\Psi_{\mathrm{gr}}^{(n)}\right\rangle; \ \forall m \in \{1 \ldots N_{\mathrm{mol}}\}, \tag{F.2}$$

where $\left|\Psi_{\mathrm{ex}}^{(m)}\right\rangle$ is the excited-state wavefunction for molecule m, $\left|\Psi_{\mathrm{ex}}^{(n)}\right\rangle$ is the ground-state wavefunction for molecule n, and N_{mol} is the total number of molecular chromophores in the system. While it is true that $\left|m\right\rangle$ must be antisymmetrized in order to be consistent with the indistinguishable of electrons, we may sometimes neglect the

supermolecular antisymmetry of the wavefunction by assuming that there is no spatial overlap (no exchange) between the molecular orbitals on different chromophores, and that therefore electrons can be effectively distinguished based on their localization. This was the ansatz taken by Davydov in his original paper.[169]

Before we proceed with expressing the matrix elements of $\hat{H}_{\text{el}}$ in this Davydov basis, we make two simplifications. First, we work in 2-chromophore basis, $|m\rangle = \left|\Psi_{\text{ex}}^{(m)}\Psi_{\text{gr}}^{(b)}\right\rangle$ and $|n\rangle = \left|\Psi_{\text{ex}}^{(n)}\Psi_{\text{gr}}^{(m)}\right\rangle$ for simplicity. Second, we assume that the single-molecule wavefunctions may be described via configuration interaction single (CIS) or TDDFT with the Tamm-Dancoff approximation (TDDFT-TDA). Under such conditions, our electronic basis states take the form

$$|m\rangle = \sum_{ia} D_{ia}^{(m)} \left|\Phi_{ia}^{(m)}\right\rangle \left|\Phi_{0}^{(n)}\right\rangle, \tag{F.3}$$

where $D_{ia}^{(m)}$ are the CIS coefficients for excitation from occupied orbital i to virtual orbital a on molecule m, $\left|\Phi_{ia}^{(m)}\right\rangle$ is the corresponding excited determinant, and $\left|\Phi_{0}^{(n)}\right\rangle$ is the reference (Hartree-Fock) determinant for molecule n. In terms of molecular orbitals (MOs) Eq. F.3 may be expressed as,

$$|m\rangle = \sum_{ia} D_{ia}^{(m)} \psi_{a}^{(m)} \left(\prod_{j\neq i}^{N_m} \psi_{j}^{(m)}\right) \left(\prod_{j}^{N_n} \psi_{j}^{(n)}\right), \tag{F.4}$$

where $\psi_j^{(m)}$ is the jth MO on molecule m, and N_m and N_n are the total number of electrons in molecule m and n respectively. Note that, by Brillouin's theorem, the overlap matrix for the basis in Eq. F.3 is diagonal. The CIS and MO coefficients for each molecular fragment are optimized via independent self-consistent field calculations.

F.1 Diagonal Elements

We begin by expanding $\hat{H}_{\mathbf{el}}$ into one-molecule and two-molecule contributions,

$$\hat{H}_{\text{el}} = \hat{h}_m + \hat{h}_n + \hat{r}_{mm} + \hat{r}_{nn} + \hat{r}_{mn}, \tag{F.5}$$

where $\hat{h}_m$ represent the sum of all one-molecule-one-electron kinetic + nuclear potential energy terms on molecule m, $\hat{r}_{mm}$ represents the the sum of all one-molecule-two-electron repulsion terms on molecule m, and $\hat{r}_{mn}$ represents the sum the two-molecule-two-electron repulsion terms. Using this notation, the diagonal elements are given by,

$$\langle m|\hat{H}_{\text{el}}|m\rangle = \left\langle \Psi_{\text{ex}}^{(m)} \middle| \hat{h}_m + \hat{r}_{mm} \middle| \Psi_{\text{ex}}^{(m)} \right\rangle + \left\langle \Psi_{\text{gr}}^{(n)} \middle| \hat{h}_n + \hat{r}_{nn} \middle| \Psi_{\text{gr}}^{(n)} \right\rangle + \left\langle \Psi_{\text{ex}}^{(m)}\Psi_{\text{gr}}^{(n)} \middle| \hat{r}_{mn} \middle| \Psi_{\text{ex}}^{(m)}\Psi_{\text{gr}}^{(n)} \right\rangle \quad \text{(F.6)}$$

$$\langle m|\hat{H}_{\text{el}}|m\rangle = E_{\text{ex}}^{(m)} + E_0^{(m)} + \left\langle \Psi_{\text{ex}}^{(m)}\Psi_{\text{gr}}^{(n)} \middle| \hat{r}_{mn} \middle| \Psi_{\text{ex}}^{(m)}\Psi_{\text{gr}}^{(n)} \right\rangle, \quad \text{(F.7)}$$

where $E_{\text{ex}}^{(m)}$ is the excited state energy of molecule m and $E_0^{(n)}$ is the ground state energy of molecule n. We can subtract the sum of ground state energies $E_0^{(m)} + E_0^{(n)}$ from these diagonal matrix elements to obtain,

$$\langle m|\hat{H}_{\text{el}}|m\rangle = E_{\text{ex}}^{(m)} - E_0^{(m)} + \left\langle \Psi_{\text{ex}}^{(m)}\Psi_{\text{gr}}^{(n)} \middle| \hat{r}_{mn} \middle| \Psi_{\text{ex}}^{(m)}\Psi_{\text{gr}}^{(n)} \right\rangle. \quad \text{(F.8)}$$

Most studies will simplify Eq. F.8 such the diagonal matrix elements are just the excitations energies $\langle m|\hat{H}_{el}|m\rangle = E_{\text{ex}}^{(m)} - E_0^{(m)}$, neglecting the two-molecule-two-electron terms. To see when and why we may neglect such terms, let us express them in terms of one-electron densities (see Appendix G),

$$\left\langle \Psi_{\text{ex}}^{(m)}\Psi_{\text{gr}}^{(n)} \middle| \hat{r}_{mn} \middle| \Psi_{\text{ex}}^{(m)}\Psi_{\text{gr}}^{(n)} \right\rangle = \sum_{i_m,j_n}^{N_m,N_n} \int d\mathbf{x}_m d\mathbf{x}_n \Psi_{\text{ex}}^{(m)}(\mathbf{x}_m)^* \Psi_{\text{ex}}^{(m)}(\mathbf{x}_m) \frac{1}{|r_{i_m} - r_{j_n}|} \Psi_{\text{gr}}^{(n)}(\mathbf{x}_m)^* \Psi_{\text{gr}}^{(n)}(\mathbf{x}_n), \quad \text{(F.9)}$$

$$\left\langle \Psi_{\text{ex}}^{(m)}\Psi_{\text{gr}}^{(n)} \middle| \hat{r}_{mn} \middle| \Psi_{\text{ex}}^{(m)}\Psi_{\text{gr}}^{(n)} \right\rangle = \int dr_m dr_n \frac{\rho_{\text{ex}}^{(m)}(r_m)\rho_{\text{gr}}^{(n)}(r_n)}{|r_m - r_n|}, \quad \text{(F.10)}$$

where $\rho_{\text{ex}}^{(m)}$ and $\rho_{\text{gr}}^{(n)}$ are the one-electron densities for the excited-state on molecule m and the ground state on molecule n respectively, $\mathbf{x}$ denotes many-electron spatial + spin coordinates, and r denotes one-electron spatial coordinates. Eq. F.10 represents a Coulomb interaction between molecule m in the excited state and the surrounding

molecules in their ground states. In essence, this term accounts for how the environment around molecule m polarizes in response to an excitation on molecule m. In the case where our molecules are sitting in a homogeneous environment, this term is a constant for all diagonal matrix elements, and therefore can be subtracted out similar to how we subtracted out the ground state energy contributions.

F.2 Off-Diagonal Elements

We begin by expanding the off-diagonal matrix elements in terms of the factors presented in Eq. F.5,

$$\langle m|\hat{H}_{\text{el}}|n\rangle = \left\langle \Psi_{\text{ex}}^{(m)}\middle|\hat{h}_m + \hat{r}_{mm}\middle|\Psi_{\text{gr}}^{(m)}\right\rangle \left\langle \Psi_{\text{gr}}^{(n)}\middle|\Psi_{\text{ex}}^{(n)}\right\rangle + \left\langle \Psi_{\text{ex}}^{(n)}\middle|\hat{h}_n + \hat{r}_{nn}\middle|\Psi_{\text{gr}}^{(n)}\right\rangle \left\langle \Psi_{\text{gr}}^{(m)}\middle|\Psi_{\text{ex}}^{(m)}\right\rangle + \left\langle \Psi_{\text{ex}}^{(m)}\Psi_{\text{gr}}^{(n)}\middle|\hat{r}_{mn}\middle|\Psi_{\text{ex}}^{(n)}\Psi_{\text{gr}}^{(m)}\right\rangle . \tag{F.11}$$

If $\Psi_{\text{ex}}^{(m)}$ is a linear combination singley-excited determinants we can use Brillouin's theorem to reduce the excitonic coupling to,

$$\langle m|\hat{H}_{\text{el}}|n\rangle = \left\langle \Psi_{\text{ex}}^{(m)}\Psi_{\text{gr}}^{(n)}\middle|\hat{r}_{mn}\middle|,\right\rangle \tag{F.12}$$

$$\langle m|\hat{H}_{\text{el}}|n\rangle = 2\sum_{iajb} D_{ia}^{(m)} D_{jb}^{(b)} 2(\psi_i^{(m)}\psi_a^{(m)}|\psi_j^{(n)}\psi_b^{(b)}). \tag{F.13}$$

It is also possible to express the coupling as an interaction between one-electron transition densities defined in Appendix G. Doing so leads to,

$$\langle m|\hat{H}_{\text{el}}|n\rangle = \int dr_m dr_n \frac{\gamma^m(r_m)\gamma^n(r_n)}{|r_m - r_n|}. \tag{F.14}$$

The nature of the coupling term largely depends on the distance between the molecules. In the limit where the molecules are far apart from each other (< 20Å), Förster showed that first order transfer rates can be accurately calculated via the point-dipole approximation [242]. At shorter distances, interaction must be smeared

out over transition densities or higher order multipole terms; and, at very small distances where there may be significant overlap between the wavefunctions of neighboring sites, Dexter showed that exchange interactions play a physically important role in excitation transfer[243].

Appendix G

1-electron density matrices and transition density matrices

The one-electron reduced density matrix (1-DM) is defined as the partial trace of the many-electron density matrix over all except one electron's spatial + spin coordinates. Concretely, in the position + spin representation the 1-DM is given by,

$$P(x, x') = N \int dx_2, \ldots, dx_N \Psi(x, x_2, \ldots, x_N) \Psi^*(x', x_2, \ldots, x_N), \tag{G.1}$$

where x are the spatial + spin coordinates of an electron, Ψ is the many-electron wavefunction, and N is the total number of electrons. In the case where our wavefunction is given by a single reference determinant, 1-DM is simply a sum over molecule spin orbitals (see Ref. 235 pg 139 and 253),

$$P(x, x') = \sum_{i}^{N/2} \psi_i(x) \psi_i^*(x'). \tag{G.2}$$

Often we are only interested in the spatial components of the 1-DM. If we discard the spin components and expand the molecular orbitals (MOs) in terms of atomic

orbitals (AOs), we arrive at,

$$P(r,r') = 2\sum_{i}^{N/2}\sum_{m,n} C_{im}C_{in}^{*}\phi_m(r)\phi_n^{*}(r') = \sum_{m,n} P_{mn}\phi_m(r)\phi_n^{*}(r'), \tag{G.3}$$

where r are the spatial coordinates of an electron, C_{im} is the expansion coefficient of molecular orbital i in atomic orbital m, ϕ denote the atomic orbitals themselves, and P_{mn} is the 1-DM in the atomic-orbital basis. The diagonal elements of $P(r,r')$ correspond to the *one-electron density*, the probability of finding a single electron in some region of space,

$$\rho(r) = P(r,r) = \sum_{\mu,\nu} P_{\mu\nu}\phi_\mu(r)\phi_\nu^{*}(r). \tag{G.4}$$

Similarly, the one-electron transition density matrix (1-TDM) is defined as the partial trace of the the outer product of two different many-electron wavefunctions: the ground state and the excited state. Within the position + spin representation it is given by,

$$\Gamma(x,x') = N\int dx_2,\ldots,dx_N \Psi_{\mathrm{gr}}(x,x_2,\ldots,x_N)\Psi_{\mathrm{ex}}^{*}(x',x_2,\ldots,x_N). \tag{G.5}$$

In the simplest case where the ground state is given by a single determinant and the excited state is a linear combination of single excitations (CIS or TDDFT/TDA) the 1-TDM takes the form,

$$\Gamma(x,x') = \sum_{ia} D_{ia}\psi_i(x)\psi_a^{*}(x'). \tag{G.6}$$

Eq. G.6 elegantly illustrates that the 1-TDM is equivalent to the rectangular CIS matrix (D_{ia}) in the molecular spin-orbital basis. Discarding spin components and expanding the MOs into AOs gives,

$$\Gamma(r,r') = \sum_{ia}\sum_{mn} D_{ia}C_{im}C_{an}^{*}\phi_m(r)\phi_n^{*}(r') = \sum_{mn}\Gamma_{mn}\phi_m(r)\phi_n^{*}(r'), \tag{G.7}$$

Were the occupied i and virtual a orbitals indices range only over spatial orbitals. Setting $x = x'$, we are left with the one-electron transition density,

$$\gamma(x) = \sum_{\mu\nu} \Gamma_{\mu\nu} \phi_\mu(x) \phi^*_\nu(x). \tag{G.8}$$

Often an extra factor of $\sqrt{2}$ is introduced to the definition of $\gamma(x)$, however this is done purely for convenience when calculating the excitation transfer coupling.

Bibliography

(1) *Catalysts for Nitrogen Fixation: Nitrogenases, Relevant Chemical Models and Commercial Processes*; Smith, B. E., Richards, R. L., Newton, W. E., Eds.; Springer Netherlands: Dordrecht, 2004, DOI: 10.1007/978-1-4020-3611-8.

(2) Smith, C.; Hill, A. K.; Torrente-Murciano, L. *Energy & Environmental Science* **2020**, *13*, 331–344, DOI: 10.1039/C9EE02873K.

(3) Christensen, C. H.; Nørskov, J. K. *The Journal of Chemical Physics* **2008**, *128*, 182503, DOI: 10.1063/1.2839299.

(4) Friend, C. M.; Xu, B. *Accounts of Chemical Research* **2017**, *50*, 517–521, DOI: 10.1021/acs.accounts.6b00510.

(5) Hu, X.; Yip, A. C. K. *Frontiers in Catalysis* **2021**, *1*, DOI: 10.3389/fctls.2021.667675.

(6) Kate, A.; Sahu, L. K.; Pandey, J.; Mishra, M.; Sharma, P. K. *Current Research in Green and Sustainable Chemistry* **2022**, *5*, 100248, DOI: 10.1016/j.crgsc.2021.100248.

(7) Li, Y.; Yu, J. *Nature Reviews Materials* **2021**, *6*, 1156–1174, DOI: 10.1038/s41578-021-00347-3.

(8) Sabatier, P., *La Catalyse en chimie organique*; Numérique Premium: 2013, DOI: 10.14375/NP.9782369430186.

(9) Rootsaert, W. J. M.; Sachtler, W. M. H. *Zeitschrift für Physikalische Chemie* **1960**, *26*, 16–26, DOI: 10.1524/zpch.1960.26.1_2.016.

(10) Trasatti, S. *Electrochimica Acta* **1987**, *32*, 369–382, DOI: 10.1016/0013-4686(87)85001-6.

(11) Nørskov, J. K.; Rossmeisl, J.; Logadottir, A.; Lindqvist, L.; Kitchin, J. R.; Bligaard, T.; Jónsson, H. *The Journal of Physical Chemistry B* **2004**, *108*, 17886–17892, DOI: 10.1021/jp047349j.

(12) Nørskov, J. K.; Abild-Pedersen, F.; Studt, F.; Bligaard, T. *Proceedings of the National Academy of Sciences* **2011**, *108*, 937–943, DOI: 10.1073/pnas.1006652108.

(13) Ooka, H.; Nakamura, R. *The Journal of Physical Chemistry Letters* **2019**, *10*, 6706–6713, DOI: 10.1021/acs.jpclett.9b01796.

(14) Ooka, H.; Huang, J.; Exner, K. S. *Frontiers in Energy Research* **2021**, *9*.

(15) Medford, A. J.; Vojvodic, A.; Hummelshøj, J. S.; Voss, J.; Abild-Pedersen, F.; Studt, F.; Bligaard, T.; Nilsson, A.; Nørskov, J. K. *Journal of Catalysis* **2015**, *328*, 36–42, DOI: 10.1016/j.jcat.2014.12.033.

(16) Inoue, Y. *Surface Science Reports* **2007**, *62*, 305–336, DOI: 10.1016/j.surfrep.2007.01.001.

(17) Shetty, M.; Walton, A.; Gathmann, S. R.; Ardagh, M. A.; Gopeesingh, J.; Resasco, J.; Birol, T.; Zhang, Q.; Tsapatsis, M.; Vlachos, D. G.; Christopher, P.; Frisbie, C. D.; Abdelrahman, O. A.; Dauenhauer, P. J. *ACS Catalysis* **2020**, *10*, 12666–12695, DOI: 10.1021/acscatal.0c03336.

(18) Winters, H. F. *The Journal of Chemical Physics* **1975**, *62*, 2454–2460, DOI: 10.1063/1.430722.

(19) Luntz, A. C.; Harris, J. *Surface Science* **1991**, *258*, 397–426, DOI: 10.1016/0039-6028(91)90934-K.

(20) Campbell, V. L.; Chen, N.; Guo, H.; Jackson, B.; Utz, A. L. *The Journal of Physical Chemistry A* **2015**, *119*, 12434–12441, DOI: 10.1021/acs.jpca.5b07873.

(21) Nave, S.; Jackson, B. *Physical Review Letters* **2007**, *98*, 173003, DOI: 10.1103/PhysRevLett.98.173003.

(22) Nave, S.; Jackson, B. *The Journal of Chemical Physics* **2009**, *130*, 054701, DOI: 10.1063/1.3065800.

(23) Tiwari, A. K.; Nave, S.; Jackson, B. *Physical Review Letters* **2009**, *103*, 253201, DOI: 10.1103/PhysRevLett.103.253201.

(24) Tiwari, A. K.; Nave, S.; Jackson, B. *The Journal of Chemical Physics* **2010**, *132*, 134702, DOI: 10.1063/1.3357415.

(25) Guo, H.; Farjamnia, A.; Jackson, B. *The Journal of Physical Chemistry Letters* **2016**, *7*, 4576–4584, DOI: 10.1021/acs.jpclett.6b01948.

(26) Nattino, F.; Costanzo, F.; Kroes, G.-J. *The Journal of Chemical Physics* **2015**, *142*, 104702, DOI: 10.1063/1.4913979.

(27) Nattino, F.; Galparsoro, O.; Costanzo, F.; Díez Muiño, R.; Alducin, M.; Kroes, G.-J. *The Journal of Chemical Physics* **2016**, *144*, 244708, DOI: 10.1063/1.4954773.

(28) Shakouri, K.; Behler, J.; Meyer, J.; Kroes, G.-J. *The Journal of Physical Chemistry Letters* **2017**, *8*, 2131–2136, DOI: 10.1021/acs.jpclett.7b00784.

(29) Shakouri, K.; Behler, J.; Meyer, J.; Kroes, G.-J. *The Journal of Physical Chemistry C* **2018**, *122*, 23470–23480, DOI: 10.1021/acs.jpcc.8b06729.

(30) Stair, P. C.; Weitz, E. *The Journal of the Optical Society of America B* **1987**, *4*, 255–260, DOI: 10.1364/JOSAB.4.000255.

(31) Lucchese, R. R.; Tully, J. C. *The Journal of Chemical Physics* **1998**, *81*, 6313, DOI: 10.1063/1.447540.

(32) Hall, R. B. *The Journal of Physical Chemistry* **1987**, *91*, 1007–1015, DOI: 10.1021/j100289a003.

(33) Zimmermann, F. M.; Ho, W. *Surface Science Reports* **1995**, *22*, 127–247, DOI: 10.1016/0167-5729(96)80001-X.

(34) Gladh, J.; Hansson, T.; Öström, H. *Surface Science* **2013**, *615*, 65–71, DOI: 10.1016/j.susc.2013.05.002.

(35) Tetenoire, A.; Juaristi, J. I.; Alducin, M. *Frontiers in Chemistry* **2023**, *11*.

(36) Alducin, M.; Camillone, N.; Hong, S.-Y.; Juaristi, J. I. *Physical Review Letters* **2019**, *123*, 246802, DOI: 10.1103/PhysRevLett.123.246802.

(37) Serrano Jiménez, A.; Sánchez Muzas, A. P.; Zhang, Y.; Ovčar, J.; Jiang, B.; Lončarić, I.; Juaristi, J. I.; Alducin, M. *Journal of Chemical Theory and Computation* **2021**, *17*, 4648–4659, DOI: 10.1021/acs.jctc.1c00347.

(38) Bonn, M.; Funk, S.; Hess, C.; Denzler, D. N.; Stampfl, C.; Scheffler, M.; Wolf, M.; Ertl, G. *Science* **1999**, *285*, 1042–1045, DOI: 10.1126/science.285.5430.1042.

(39) Kelling, S.; Cerasari, S.; Rotermund, H. H.; Ertl, G.; King, D. A. *Chemical Physics Letters* **1998**, *293*, 325–330, DOI: 10.1016/S0009-2614(98)00811-2.

(40) Kelling, S.; Saito, N.; Inoue, Y.; King, D. A. *Applied Surface Science* **1999**, *150*, 47–57, DOI: 10.1016/S0169-4332(99)00222-6.

(41) Mitrelias, T.; Kelling, S.; Kvon, R. I.; Ostanin, V. P.; King, D. A. *Surface Science* **1998**, *417*, 97–106, DOI: 10.1016/S0039-6028(98)00671-2.

(42) Von Boehn, B.; Foerster, M.; von Boehn, M.; Prat, J.; Macià, F.; Casals, B.; Khaliq, M. W.; Hernández-Mínguez, A.; Aballe, L.; Imbihl, R. *Angewandte Chemie International Edition* **2020**, *59*, 20224–20229, DOI: 10.1002/anie.202005883.

(43) Nishiyama, H.; Inoue, Y. *Surface Science* **2006**, *600*, 2644–2649, DOI: 10.1016/j.susc.2006.03.047.

(44) Deutch, J. M.; Silbey, R. *Physical Review A* **1971**, *3*, 2049–2052, DOI: 10.1103/PhysRevA.3.2049.

(45) Adelman, S. A.; Doll, J. D. *The Journal of Chemical Physics* **1976**, *64*, 2375–2388, DOI: 10.1063/1.432526.

(46) Tully, J. C. *The Journal of Chemical Physics* **1980**, *73*, 1975–1985, DOI: 10.1063/1.440287.

(47) Garrison, B. J.; Adelman, S. A. *Surface Science* **1977**, *66*, 253–271, DOI: 10.1016/0039-6028(77)90411-3.

(48) Tully, J. C. *Surface Science* **1994**, *299-300*, 667–677, DOI: 10.1016/0039-6028(94)90688-2.

(49) Zwanzig, R. *The Journal of Chemical Physics* **1960**, *33*, 1338–1341, DOI: 10.1063/1.1731409.

(50) Mori, H. *Progress of Theoretical Physics* **1965**, *33*, 423–455, DOI: 10.1143/PTP.33.423.

(51) Chorin, A. J.; Hald, O. H.; Kupferman, R. *Proceedings of the National Academy of Sciences* **2000**, *97*, 2968–2973, DOI: 10.1073/pnas.97.7.2968.

(52) Zwanzig, R. *Journal of Statistical Physics* **1973**, *9*, 215–220, DOI: 10.1007/BF01008729.

(53) Kupferman, R.; Stuart, A. M.; Terry, J. R.; Tupper, P. F. *Stochastics and Dynamics* **2011**, DOI: 10.1142/S0219493702000571.

(54) Ayaz, C.; Scalfi, L.; Dalton, B. A.; Netz, R. R. *Physical Review E* **2022**, *105*, 054138, DOI: 10.1103/PhysRevE.105.054138.

(55) Kramers, H. A. *Physica* **1940**, *7*, 284–304, DOI: 10.1016/S0031-8914(40)90098-2.

(56) Grote, R. F.; Hynes, J. T. *The Journal of Chemical Physics* **1980**, *73*, 2715–2732, DOI: 10.1063/1.440485.

(57) Schroeder, J.; Troe, J.; Vöhringer, P. *Zeitschrift für Physikalische Chemie* **1995**, *188*, 287–306, DOI: 10.1524/zpch.1995.188.Part_1_2.287.

(58) Rondin, L.; Gieseler, J.; Ricci, F.; Quidant, R.; Dellago, C.; Novotny, L. *Nature Nanotechnology* **2017**, *12*, 1130–1133, DOI: 10.1038/nnano.2017.198.

(59) Pollak, E.; Grabert, H.; Hänggi, P. *The Journal of Chemical Physics* **1989**, *91*, 4073–4087, DOI: 10.1063/1.456837.

(60) Kappler, J.; Daldrop, J. O.; Brünig, F. N.; Boehle, M. D.; Netz, R. R. *The Journal of Chemical Physics* **2018**, *148*, 014903, DOI: 10.1063/1.4998239.

(61) Eyring, H. *The Journal of Chemical Physics* **1935**, *3*, 107–115, DOI: 10.1063/1.1749604.

(62) Wigner, E. *Zeitschrift für Physikalische Chemie* **1932**, *19B*, 203–216, DOI: 10.1515/zpch-1932-1920.

(63) Pollak, E.; Talkner, P. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **2005**, *15*, 026116, DOI: 10.1063/1.1858782.

(64) Pollak, E. *The Journal of Chemical Physics* **1986**, *85*, 865–867, DOI: 10.1063/1.451294.

(65) Farahvash, A.; Willard, A. P. A theory of phonon induced friction on molecular adsorbates, 2023, DOI: 10.48550/arXiv.2309.08447.

(66) Chen, B. W. J.; Xu, L.; Mavrikakis, M. *Chemical Reviews* **2021**, *121*, 1007–1048, DOI: 10.1021/acs.chemrev.0c01060.

(67) *Journal of the American Chemical Society* **2008**, *130*, 7300–7304, DOI: 10.1021/ja800286k.

(68) Michalsky, R.; Zhang, Y.-J.; Medford, A. J.; Peterson, A. A. *The Journal of Physical Chemistry C* **2014**, *118*, 13026–13034, DOI: 10.1021/jp503756g.

(69) Ertl, G.; Neumann, M.; Streit, K. M. *Surface Science* **1977**, *64*, 393–410, DOI: 10.1016/0039-6028(77)90052-8.

(70) Steininger, H.; Lehwald, S.; Ibach, H. *Surface Science* **1982**, *123*, 264–282, DOI: 10.1016/0039-6028(82)90328-4.

(71) Nørskov, J. K.; Lang, N. D. *Physical Review B* **1980**, *21*, 2131–2136, DOI: 10.1103/physrevb.21.2131.

(72) Jacobsen, K. W.; Stoltze, P.; Nørskov, J. K. *Surface Science* **1996**, *366*, 394–402, DOI: 10.1016/0039-6028(96)00816-3.

(73) Larsen, A. H. et al. *Journal of Physics: Condensed Matter* **2017**, *29*, 273002, DOI: 10.1088/1361-648X/aa680e.

(74) Tadmor, E. B.; Elliott, R. S.; Sethna, J. P.; Miller, R. E.; Becker, C. A. *JOM* **2011**, *63*, 17, DOI: 10.1007/s11837-011-0102-6.

(75) Schiøtz, J. EMT potential for Ni developed by Jacobsen, Stolze, and Norskov (1996) v001, OpenKIM, https://doi.org/10.25950/ef22cbc3, 2019, DOI: 10.25950/ef22cbc3.

(76) Schiøtz, J. Effective Medium Theory (EMT) potential driver v004, OpenKIM, https://doi.org/10.25950/7e5b8be7, 2019, DOI: 10.25950/7e5b8be7.

(77) Tadmor, E. EAM potential (LAMMPS cubic hermite tabulation) for Pt (Universal3) developed by Foiles, Baskes, and Daw (1986) v000, OpenKIM, https://doi.org/10.25950/24de6537, 2018, DOI: 10.25950/24de6537.

(78) Elliott, R. S. EAM Model Driver for tabulated potentials with cubic Hermite spline interpolation as used in LAMMPS v005, OpenKIM, https://doi.org/10.25950/68defa36, 2018, DOI: 10.25950/68defa36.

(79) Chen, D.-L.; Al-Saidi, W. A.; Karl Johnson, J. *Journal of Physics: Condensed Matter* **2012**, *24*, 424211, DOI: 10.1088/0953-8984/24/42/424211.

(80) Lahee, A. M.; Toennies, J. P.; Wöll, C. *Surface Science* **1986**, *177*, 371–388, DOI: 10.1016/0039-6028(86)90146-9.

(81) Gunasooriya, G. T. K. K.; Saeys, M. *ACS Catalysis* **2018**, *8*, 10225–10233, DOI: 10.1021/acscatal.8b02371.

(82) Rittmeyer, S. P.; Bukas, V. J.; Reuter, K. *Advances in Physics: X* **2018**, *3*, 1381574, DOI: 10.1080/23746149.2017.1381574.

(83) Zhou, X.; Jiang, B. *The Journal of Chemical Physics* **2019**, *150*, 024704, DOI: 10.1063/1.5078541.

(84) Zhou, Y.; Zhou, L.; Hu, X.; Xie, D. *The Journal of Physical Chemistry C* **2020**, *124*, 10573–10583, DOI: 10.1021/acs.jpcc.0c01247.

(85) Fortini, A.; Mendelev, M. I.; Buldyrev, S.; Srolovitz, D. *Journal of Applied Physics* **2008**, *104*, 074320, DOI: 10.1063/1.2991301.

(86) Hong, S.; Rahman, T. S.; Heid, R.; Bohnen, K. P. *Physical Review B* **2005**, *72*, 205424, DOI: 10.1103/PhysRevB.72.205424.

(87) Benedek, G.; Bernasconi, M.; Chis, V.; Chulkov, E.; Echenique, P. M.; Hellsing, B.; Toennies, J. P. *Journal of Physics: Condensed Matter* **2010**, *22*, 084020, DOI: 10.1088/0953-8984/22/8/084020.

(88) Bortolani, V.; Franchini, A.; Santoro, G.; Toennies, J. P.; Wöll, C.; Zhang, G. *Physical Review B* **1989**, *40*, 3524–3545, DOI: 10.1103/PhysRevB.40.3524.

(89) Shih, P.; Berkelbach, T. C. *Physical Review B* **2022**, *106*, 144307, DOI: 10.1103/PhysRevB.106.144307.

(90) Golibrzuch, K.; Shirhatti, P. R.; Geweke, J.; Werdecker, J.; Kandratsenka, A.; Auerbach, D. J.; Wodtke, A. M.; Bartels, C. *Journal of the American Chemical Society* **2015**, *137*, 1465–1475, DOI: 10.1021/ja509530k.

(91) Garrett, B. C.; Truhlar, D. G. *The Journal of Physical Chemistry* **1979**, *83*, 1052–1079, DOI: 10.1021/j100471a031.

(92) Chandler, D. *The Journal of Chemical Physics* **1978**, *68*, 2959–2970, DOI: 10.1063/1.436049.

(93) Miller, W. H. *Annual Review of Physical Chemistry* **2014**, *65*, 1–19, DOI: 10.1146/annurev-physchem-040513-103720.

(94) Rettner, C. T.; Bethune, D. S.; Schweizer, E. K. *The Journal of Chemical Physics* **1990**, *92*, 1442–1457, DOI: 10.1063/1.458103.

(95) Kuznets-Speck, B.; Limmer, D. T. *Proceedings of the National Academy of Sciences* **2021**, *118*, e2020863118, DOI: 10.1073/pnas.2020863118.

(96) Kuznets-Speck, B.; Limmer, D. T. *Biophysical Journal* **2023**, *122*, 1659–1664, DOI: 10.1016/j.bpj.2023.03.031.

(97) Shugaev, M. V.; Zhigilei, L. V. *Journal of Applied Physics* **2021**, *130*, 185108, DOI: 10.1063/5.0071170.

(98) Farahvash, A.; Agrawal, M.; Peterson, A. A.; Willard, A. P. *Journal of Chemical Theory and Computation* **2023**, DOI: 10.1021/acs.jctc.3c00473.

(99) Sitz, G. O. *Reports on Progress in Physics* **2002**, *65*, 1165, DOI: 10.1088/0034-4885/65/8/202.

(100) Juurlink, L. B. F.; Killelea, D. R.; Utz, A. L. *Progress in Surface Science* **2009**, *84*, 69–134, DOI: 10.1016/j.progsurf.2009.01.001.

(101) Chadwick, H.; Beck, R. D. *Annual Review of Physical Chemistry* **2017**, *68*, 39–61, DOI: 10.1146/annurev-physchem-052516-044910.

(102) Luntz, A. C.; Bethune, D. S. *The Journal of Chemical Physics* **1989**, *90*, 1274–1280, DOI: 10.1063/1.456132.

(103) Henkelman, G.; Jónsson, H. *Physical Review Letters* **2001**, *86*, 664–667, DOI: 10.1103/PhysRevLett.86.664.

(104) Nave, S.; Tiwari, A. K.; Jackson, B. *The Journal of Chemical Physics* **2010**, *132*, 054705, DOI: 10.1063/1.3297885.

(105) Egeberg, R. C.; Larsen, J. H.; Chorkendorff, I. *Physical Chemistry Chemical Physics* **2001**, *3*, 2007–2011, DOI: 10.1039/B101177O.

(106) Diekhöner, L.; Mortensen, H.; Baurichter, A.; Jensen, E.; Petrunin, V. V.; Luntz, A. C. *The Journal of Chemical Physics* **2001**, *115*, 9028–9035, DOI: 10.1063/1.1413746.

(107) Busnengo, H. F.; Dong, W.; Salin, A. *Physical Review Letters* **2004**, *93*, 236103, DOI: 10.1103/PhysRevLett.93.236103.

(108) Bukas, V. J.; Mitra, S.; Meyer, J.; Reuter, K. *The Journal of Chemical Physics* **2015**, *143*, 034705, DOI: 10.1063/1.4926989.

(109) Ceriotti, M.; Bussi, G.; Parrinello, M. *Journal of Chemical Theory and Computation* **2010**, *6*, 1170–1180, DOI: 10.1021/ct900563s.

(110) Baczewski, A. D.; Bond, S. D. *The Journal of Chemical Physics* **2013**, *139*, 044107, DOI: 10.1063/1.4815917.

(111) Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; in 't Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D.; Shan, R.; Stevens, M. J.; Tranchida, J.; Trott, C.; Plimpton, S. J. *Computer Physics Communications* **2022**, *271*, 108171, DOI: 10.1016/j.cpc.2021.108171.

(112) Daw, M. S.; Baskes, M. I. *Physical Review Letters* **1983**, *50*, 1285–1288, DOI: 10.1103/physrevlett.50.1285.

(113) Jacobsen, K. W.; Stoltze, P.; Nørskov, J. K. *Surface Science* **1996**, *366*, 394–402, DOI: 10.1016/0039-6028(96)00816-3.

(114) Zhou, X. W.; Johnson, R. A.; Wadley, H. N. G. *Physical Review B* **2004**, *69*, 144113, DOI: 10.1103/PhysRevB.69.144113.

(115) Heinz, H.; Vaia, R. A.; Farmer, B. L.; Naik, R. R. *The Journal of Physical Chemistry C* **2008**, *112*, 17281–17290, DOI: 10.1021/jp801931d.

(116) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *The Journal of Physical Chemistry* **1987**, *91*, 6269–6271, DOI: 10.1021/j100308a038.

(117) Florencio, J.; Lee, M. H. *Physical Review A* **1985**, *31*, 3231–3236, DOI: 10.1103/PhysRevA.31.3231.

(118) Lee, M. H. *Symmetry* **2016**, *8*, 22, DOI: 10.3390/sym8040022.

(119) Arbouet, A.; Voisin, C.; Christofilos, D.; Langot, P.; Fatti, N. D.; Vallée, F.; Lermé, J.; Celep, G.; Cottancin, E.; Gaudry, M.; Pellarin, M.; Broyer, M.; Maillard, M.; Pileni, M. P.; Treguer, M. *Physical Review Letters* **2003**, *90*, 177401, DOI: 10.1103/PhysRevLett.90.177401.

(120) Franco, F.; Rettenmaier, C.; Jeon, H. S.; Cuenya, B. R. *Chemical Society Reviews* **2020**, *49*, 6884–6946, DOI: 10.1039/D0CS00835D.

(121) Akbashev, A. R. *ACS Catalysis* **2022**, *12*, 4296–4301, DOI: 10.1021/acscatal.2c00123.

(122) Cortright, R. D.; Davda, R. R.; Dumesic, J. A. *Nature* **2002**, *418*, 964–967, DOI: 10.1038/nature01009.

(123) Huber, G. W.; Cortright, R. D.; Dumesic, J. A. *Angewandte Chemie* **2004**, *116*, 1575–1577, DOI: 10.1002/ange.200353050.

(124) Gohda, Y.; Schnur, S.; Groß, A. *Faraday Discussions* **2009**, *140*, 233–244, DOI: 10.1039/B802270D.

(125) Carrasco, J.; Hodgson, A.; Michaelides, A. *Nature Materials* **2012**, *11*, 667–674, DOI: 10.1038/nmat3354.

(126) Liu, X.; Xiao, J.; Peng, H.; Hong, X.; Chan, K.; Nørskov, J. K. *Nature Communications* **2017**, *8*, 15438, DOI: 10.1038/ncomms15438.

(127) Xie, T.; Bodenschatz, C. J.; Getman, R. B. *Reaction Chemistry & Engineering* **2019**, *4*, 383–392, DOI: 10.1039/C8RE00267C.

(128) Sakong, S.; Groß, A. *Physical Chemistry Chemical Physics* **2020**, *22*, 10431–10437, DOI: 10.1039/C9CP06584A.

(129) Ringe, S.; Morales-Guio, C. G.; Chen, L. D.; Fields, M.; Jaramillo, T. F.; Hahn, C.; Chan, K. *Nature Communications* **2020**, *11*, 33, DOI: 10.1038/s41467-019-13777-z.

(130) Kastlunger, G.; Lindgren, P.; Peterson, A. A. *The Journal of Physical Chemistry C* **2018**, *122*, 12771–12781, DOI: 10.1021/acs.jpcc.8b02465.

(131) Hori, Y. In *Modern Aspects of Electrochemistry*, Vayenas, C. G., White, R. E., Gamboa-Aldeco, M. E., Eds.; Modern Aspects of Electrochemistry; Springer: New York, NY, 2008, pp 89–189, DOI: 10.1007/978-0-387-49489-0_3.

(132) Van der Laan, G. P.; Beenackers, A. A. C. M. *Catalysis Reviews* **1999**, *41*, 255–318, DOI: 10.1081/CR-100101170.

(133) Rood, S.; Eslava, S.; Manigrasso, A.; Bannister, C. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering* **2019**, *234*, 936–949, DOI: 10.1177/0954407019859822.

(134) Antal Jr., M.; Allen, S.; Schulman, D.; Xu, X.; Divilio, R. *Industrial and Engineering Chemistry Research* **2000**, *39*, 4040–4053.

(135) Dellago, C.; Bolhuis, P. G.; Chandler, D. *The Journal of Chemical Physics* **1998**, *108*, 9236–9245, DOI: 10.1063/1.476378.

(136) Dellago, C.; Bolhuis, P. G.; Csajka, F. S.; Chandler, D. *The Journal of Chemical Physics* **1998**, *108*, 1964–1977, DOI: 10.1063/1.475562.

(137) Bolhuis, P. G.; Swenson, D. W. H. *Advanced Theory and Simulations* **2021**, *4*, 2000237, DOI: 10.1002/adts.202000237.

(138) Peters, B. *Annual Review of Physical Chemistry* **2016**, *67*, 669–690, DOI: 10.1146/annurev-physchem-040215-112215.

(139) Peters, B. *The Journal of Chemical Physics* **2006**, *125*, 241101, DOI: 10.1063/1.2409924.

(140) Peters, B. *Chemical Physics Letters* **2010**, *494*, 100–103, DOI: 10.1016/j.cplett.2010.05.069.

(141) Chen, B.; Xing, J.; Siepmann, J. I. *The Journal of Physical Chemistry B* **2000**, *104*, 2391–2401, DOI: 10.1021/jp993687m.

(142) Lorentz, H. A. *Annalen der Physik* **1881**, *248*, 127–136, DOI: 10.1002/andp.18812480110.

(143) Enkovaara, J. et al. *Journal of Physics: Condensed Matter* **2010**, *22*, 253202, DOI: 10.1088/0953-8984/22/25/253202.

(144) Mortensen, J. J.; Hansen, L. B.; Jacobsen, K. W. *Physical Review B* **2005**, *71*, 035109, DOI: 10.1103/PhysRevB.71.035109.

(145) Swenson, D. W. H.; Prinz, J.-H.; Noe, F.; Chodera, J. D.; Bolhuis, P. G. *Journal of Chemical Theory and Computation* **2019**, *15*, 813–836, DOI: 10.1021/acs.jctc.8b00626.

(146) Brotzakis, Z. F.; Bolhuis, P. G. *The Journal of Chemical Physics* **2016**, *145*, 164112, DOI: 10.1063/1.4965882.

(147) Laio, A.; Parrinello, M. *Proceedings of the National Academy of Sciences* **2002**, *99*, 12562–12566, DOI: 10.1073/pnas.202427399.

(148) Ghoussoub, M.; Yadav, S.; Ghuman, K. K.; Ozin, G. A.; Singh, C. V. *ACS Catalysis* **2016**, *6*, 7109–7117, DOI: 10.1021/acscatal.6b01545.

(149) Creazzo, F.; Luber, S. *Chemistry – A European Journal* **2021**, *27*, 17024–17037, DOI: 10.1002/chem.202102356.

(150) Piana, S.; Laio, A. *The Journal of Physical Chemistry B* **2007**, *111*, 4553–4559, DOI: 10.1021/jp067873l.

(151) Barducci, A.; Chelli, R.; Procacci, P.; Schettino, V.; Gervasio, F. L.; Parrinello, M. *Journal of the American Chemical Society* **2006**, *128*, 2705–2710, DOI: 10.1021/ja057076l.

(152) Granata, D.; Camilloni, C.; Vendruscolo, M.; Laio, A. *Proceedings of the National Academy of Sciences* **2013**, *110*, 6817–6822, DOI: 10.1073/pnas.1218350110.

(153) Réocreux, R.; Girel, É.; Clabaut, P.; Tuel, A.; Besson, M.; Chaumonnot, A.; Cabiac, A.; Sautet, P.; Michel, C. *Nature Communications* **2019**, *10*, DOI: 10.1038/s41467-019-10981-9.

(154) Dama, J. F.; Hocky, G. M.; Sun, R.; Voth, G. A. *Journal of Chemical Theory and Computation* **2015**, *11*, 5638–5650, DOI: 10.1021/acs.jctc.5b00907.

(155) Barducci, A.; Bussi, G.; Parrinello, M. *Phys. Rev. Lett.* **2008**, *100*, 020603, DOI: 10.1103/PhysRevLett.100.020603.

(156) Branduardi, D.; Bussi, G.; Parrinello, M. *Journal of Chemical Theory and Computation* **2012**, *8*, 2247–2254, DOI: 10.1021/ct3002464.

(157) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. *Computer Physics Communications* **2014**, *185*, 604–613, DOI: 10.1016/j.cpc.2013.09.018.

(158) Dell'Angela, M. et al. *Science* **2013**, *339*, 1302–1305, DOI: 10.1126/science.1231711.

(159) Doren, D. J.; Tully, J. C. *Langmuir* **1988**, *4*, 256–268, DOI: 10.1021/la00080a004.

(160) Doren, D. J.; Tully, J. C. *The Journal of Chemical Physics* **1991**, *94*, 8428–8440, DOI: 10.1063/1.460076.

(161) Peters, B. *The Journal of Physical Chemistry B* **2015**, *119*, 6349–6356, DOI: 10.1021/acs.jpcb.5b02547.

(162) Farahvash, A.; Lee, C.-K.; Sun, Q.; Shi, L.; Willard, A. P. *The Journal of Chemical Physics* **2020**, *153*, 074111, DOI: 10.1063/5.0016009.

(163) Kippelen, B.; Brédas, J.-L. *Energy & Environmental Science* **2009**, *2*, 251–261.

(164) Brabec, C. J. *Solar energy materials and solar cells* **2004**, *83*, 273–292.

(165) Thompson, J.; Blyth, R.; Mazzeo, M.; Anni, M.; Gigli, G.; Cingolani, R. *Applied Physics Letters* **2001**, *79*, 560–562.

(166) Gómez-Bombarelli, R.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Duvenaud, D.; Maclaurin, D.; Blood-Forsythe, M. A.; Chae, H. S.; Einzinger, M.; Ha, D.-G.; Wu, T., et al. *Nature materials* **2016**, *15*, 1120–1127.

(167) Castellanos, M. A.; Dodin, A.; Willard, A. P. *Physical Chemistry Chemical Physics* **2020**, *22*, 3048–3057, DOI: 10.1039/C9CP05625D.

(168) Frenkel, J. *Phys. Rev.* **1931**, *37*, 17–44, DOI: 10.1103/PhysRev.37.17.

(169) A S Davydov *Soviet Physics Uspekhi* **1964**, *7*, 145.

(170) Spano, F. C. *Annual Review of Physical Chemistry* **2006**, *57*, 217–243, DOI: 10.1146/annurev.physchem.57.032905.104557.

(171) Shi, L.; Willard, A. P. *The Journal of Chemical Physics* **2018**, *149*, 094110, DOI: 10.1063/1.5044553.

(172) Beljonne, D.; Cornil, J.; Silbey, R.; Millié, P.; Brédas, J. L. *The Journal of Chemical Physics* **2000**, *112*, 4749–4758, DOI: 10.1063/1.481031.

(173) Morrison, A. F.; You, Z.-Q.; Herbert, J. M. *Journal of Chemical Theory and Computation* **2014**, *10*, 5366–5376, DOI: 10.1021/ct500765m.

(174) Li, X.; Parrish, R. M.; Liu, F.; Kokkila Schumacher, S. I. L.; Martínez, T. J. *Journal of Chemical Theory and Computation* **2017**, *13*, 3493–3504, DOI: 10.1021/acs.jctc.7b00171.

(175) Lee, C. K.; Shi, L.; Willard, A. P. *The Journal of Physical Chemistry C* **2019**, *123*, 306–314, DOI: 10.1021/acs.jpcc.8b11504.

(176) Barford, W.; Bittner, E. R.; Ward, A. *The Journal of Physical Chemistry A* **2012**, *116*, 10319–10327, DOI: 10.1021/jp307041n.

(177) Tozer, O. R.; Barford, W. *The Journal of Chemical Physics* **2015**, *143*, 084102, DOI: 10.1063/1.4929378.

(178) Lee, E. M. Y.; Willard, A. P. *Journal of Chemical Theory and Computation* **2019**, DOI: 10.1021/acs.jctc.9b00302.

(179) Bittner, E. R. *Journal of Photochemistry and Photobiology A: Chemistry* **2007**, *190*, 328–334, DOI: 10.1016/j.jphotochem.2006.12.007.

(180) Graugnard, E.; Kellis, D. L.; Bui, H.; Barnes, S.; Kuang, W.; Lee, J.; Hughes, W. L.; Knowlton, W. B.; Yurke, B. *Nano Letters* **2012**, *12*, 2117–2122, DOI: 10.1021/nl3004336.

(181) Tretiak, S.; Middleton, C.; Chernyak, V.; Mukamel, S. *The Journal of Physical Chemistry B* **2000**, *104*, 4519–4528, DOI: 10.1021/jp9939930.

(182) Sisto, A.; Glowacki, D. R.; Martinez, T. J. *Accounts of Chemical Research* **2014**, *47*, 2857–2866, DOI: 10.1021/ar500229p.

(183) Ye, J.; Sun, K.; Zhao, Y.; Yu, Y.; Kong Lee, C.; Cao, J. *The Journal of Chemical Physics* **2012**, *136*, 245104, DOI: 10.1063/1.4729786.

(184) Da Como, E.; Loi, M. A.; Murgia, M.; Zamboni, R.; Muccini, M. *Journal of the American Chemical Society* **2006**, *128*, 4277–4281, DOI: 10.1021/ja056060s.

(185) Fitzner, R.; Mena-Osteritz, E.; Mishra, A.; Schulz, G.; Reinold, E.; Weil, M.; Körner, C.; Ziehlke, H.; Elschner, C.; Leo, K.; Riede, M.; Pfeiffer, M.; Uhrich, C.; Bäuerle, P. *Journal of the American Chemical Society* **2012**, *134*, 11064–11067, DOI: 10.1021/ja302320c.

(186) Schwabegger, G.; Djuric, T.; Sitter, H.; Resel, R.; Simbrunner, C. *Crystal Growth & Design* **2013**, *13*, 536–542, DOI: 10.1021/cg3010823.

(187) Mani, A.; Schoonman, J.; Goossens, A. *The Journal of Physical Chemistry B* **2005**, *109*, 4829–4836, DOI: 10.1021/jp045916x.

(188) Wagner, T.; Fritz, D. R.; Zeppenfeld, P. *Synthetic Metals* **2011**, *161*, 2006–2010, DOI: 10.1016/j.synthmet.2011.07.014.

(189) Djuric, T.; Hernandez-Sosa, G.; Schwabegger, G.; Koini, M.; Hesser, G.; Arndt, M.; Brinkmann, M.; Sitter, H.; Simbrunner, C.; Resel, R. *Journal of Materials Chemistry* **2012**, *22*, 15316–15325, DOI: 10.1039/C2JM32023A.

(190) Hinrichs, K.; Levichkova, M.; Wynands, D.; Walzer, K.; Eichhorn, K. J.; Bäuerle, P.; Leo, K.; Riede, M. *Thin Solid Films* **2012**, *525*, 97–105, DOI: 10.1016/j.tsf.2012.10.069.

(191) Varene, E.; Bogner, L.; Bronner, C.; Tegeder, P. *Phys. Rev. Lett.* **2012**, *109*, 207601, DOI: 10.1103/PhysRevLett.109.207601.

(192) Moser, A.; Salzmann, I.; Oehzelt, M.; Neuhold, A.; Flesch, H.-G.; Ivanco, J.; Pop, S.; Toader, T.; Zahn, D. R. T.; Smilgies, D.-M.; Resel, R. *Chemical Physics Letters* **2013**, *574*, 51–55, DOI: 10.1016/j.cplett.2013.04.053.

(193) D'Avino, G.; Muccioli, L.; Zannoni, C. *Advanced Functional Materials* **2015**, *25*, 1985–1995, DOI: 10.1002/adfm.201402609.

(194) Bogner, L.; Yang, Z.; Baum, S.; Corso, M.; Fitzner, R.; Bäuerle, P.; Franke, K. J.; Pascual, J. I.; Tegeder, P. *The Journal of Physical Chemistry C* **2016**, *120*, 27268–27275, DOI: 10.1021/acs.jpcc.6b07325.

(195) Klett, B.; Cocchi, C.; Pithan, L.; Kowarik, S.; Draxl, C. *Physical Chemistry Chemical Physics* **2016**, *18*, 14603–14609, DOI: 10.1039/C6CP01405D.

(196) Taima, T.; Shahiduzzaman, M.; Ishizeki, T.; Yamamoto, K.; Karakawa, M.; Kuwabara, T.; Takahashi, K. *The Journal of Physical Chemistry C* **2017**, *121*, 19699–19704, DOI: 10.1021/acs.jpcc.7b07953.

(197) Bronsch, W.; Wagner, T.; Baum, S.; Wansleben, M.; Zielke, K.; Ghanbari, E.; Györök, M.; Navarro-Quezada, A.; Zeppenfeld, P.; Weinelt, M.; Gahl, C. *The Journal of Physical Chemistry C* **2019**, *123*, 7931–7939, DOI: 10.1021/acs.jpcc.8b07280.

(198) Pople, J. A. *Trans. Faraday Soc.* **1953**, *49*, 1375–1385, DOI: 10.1039/TF9534901375.

(199) Chang, J. C. *The Journal of Chemical Physics* **1975**, *67*, 3901, DOI: 10.1063/1.435427.

(200) Wong, C. Y.; Curutchet, C.; Tretiak, S.; Scholes, G. D. *The Journal of Chemical Physics* **2009**, *130*, 081104, DOI: 10.1063/1.3088846.

(201) Kasha M.; Rawls H. R.; Ashraf El-Bayoumi M. *Pure and Applied Chemistry* **1965**, *11*, 371, DOI: 10.1351/pac196511030371.

(202) Madjet, M. E.; Abdurahman, A.; Renger, T. *The Journal of Physical Chemistry B* **2006**, *110*, 17268–17281, DOI: 10.1021/jp0615398.

(203) Howard, I. A.; Zutterman, F.; Deroover, G.; Lamoen, D.; Van Alsenoy, C. *The Journal of Physical Chemistry B* **2004**, *108*, 19155–19162, DOI: 10.1021/jp040417h.

(204) Krueger, B. P.; Scholes, G. D.; Fleming, G. R. *The Journal of Physical Chemistry B* **1998**, *102*, 9603–9604, DOI: 10.1021/jp983589l.

(205) Kistler, K. A.; Spano, F. C.; Matsika, S. *The Journal of Physical Chemistry B* **2013**, *117*, 2032–2044, DOI: 10.1021/jp310603z.

(206) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *Journal of the American Chemical Society* **1996**, *118*, 11225–11236, DOI: 10.1021/ja9621760.

(207) DuBay, K. H.; Hall, M. L.; Hughes, T. F.; Wu, C.; Reichman, D. R.; Friesner, R. A. *Journal of Chemical Theory and Computation* **2012**, *8*, 4556–4569, DOI: 10.1021/ct300175w.

(208) Nosé, S. *The Journal of Chemical Physics* **1984**, *81*, 511–519, DOI: 10.1063/1.447334.

(209) Hoover, W. G. *Physical Review A* **1985**, *31*, 1695–1697, DOI: 10.1103/PhysRevA.31.1695.

(210) Bowers, K. J.; Chow, D. E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. E. In *SC '06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*, 2006, pp 43–43, DOI: 10.1109/SC.2006.54.

(211) Yanai, T.; Tew, D. P.; Handy, N. C. *Chemical Physics Letters* **2004**, *393*, 51–57, DOI: 10.1016/j.cplett.2004.06.011.

(212) Lu, C.; Liu, Q.; Sun, Q.; Hsieh, C.-Y.; Zhang, S.; Shi, L.; Lee, C.-K. *The Journal of Physical Chemistry C* **2020**, *124*, 7048–7060, DOI: 10.1021/acs.jpcc.0c00329.

(213) Shao, Y.; Mei, Y.; Sundholm, D.; Kaila, V. R. I. *Journal of Chemical Theory and Computation* **2020**, *16*, 587–600, DOI: 10.1021/acs.jctc.9b00823.

(214) Laurent, A. D.; Jacquemin, D. *International Journal of Quantum Chemistry* **2013**, *113*, 2019–2039, DOI: 10.1002/qua.24438.

(215) Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; Wouters, S.; Chan, G. K.-L. *WIREs Computational Molecular Science* **2018**, *8*, e1340, DOI: 10.1002/wcms.1340.

(216) Pedregosa, F. et al. *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.

(217) Martín Abadi et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, 2015.

(218) Rupp, M.; Tkatchenko, A.; Müller, K.-R.; von Lilienfeld, O. A. *Phys. Rev. Lett.* **2012**, *108*, 058301, DOI: 10.1103/PhysRevLett.108.058301.

(219) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; von Lilienfeld, O. A. *Journal of Chemical Theory and Computation* **2015**, *11*, 2087–2096, DOI: 10.1021/acs.jctc.5b00099.

(220) Hansen, K.; Montavon, G.; Biegler, F.; Fazli, S.; Rupp, M.; Scheffler, M.; von Lilienfeld, O. A.; Tkatchenko, A.; Müller, K.-R. *Journal of Chemical Theory and Computation* **2013**, *9*, 3404–3419, DOI: 10.1021/ct400195d.

(221) Rupp, M. *International Journal of Quantum Chemistry* **2015**, *115*, 1058–1073, DOI: 10.1002/qua.24954.

(222) Behler, J.; Parrinello, M. *Physical Review Letters* **2007**, *98*, 146401, DOI: 10.1103/PhysRevLett.98.146401.

(223) Welborn, M.; Cheng, L.; Miller, T. F. *Journal of Chemical Theory and Computation* **2018**, *14*, 4772–4779, DOI: 10.1021/acs.jctc.8b00636.

(224) Schütt, K. T.; Sauceda, H. E.; Kindermans, P.-J.; Tkatchenko, A.; Müller, K.-R. *The Journal of Chemical Physics* **2018**, *148*, 241722, DOI: 10.1063/1.5019779.

(225) Schütt, K. T.; Gastegger, M.; Tkatchenko, A.; Müller, K.-R.; Maurer, R. J. *Nature Communications* **2019**, *10*, 1–10, DOI: 10.1038/s41467-019-12875-2.

(226) Grisafi, A.; Fabrizio, A.; Meyer, B.; Wilkins, D. M.; Corminboeuf, C.; Ceriotti, M. *ACS Central Science* **2019**, *5*, 57–64, DOI: 10.1021/acscentsci.8b00551.

(227) Dral, P. O. *The Journal of Physical Chemistry Letters* **2020**, *11*, 2336–2347, DOI: 10.1021/acs.jpclett.9b03664.

(228) Dral, P. O.; Barbatti, M.; Thiel, W. *The Journal of Physical Chemistry Letters* **2018**, *9*, 5660–5663, DOI: 10.1021/acs.jpclett.8b02469.

(229) Chen, W.-K.; Liu, X.-Y.; Fang, W.-H.; Dral, P. O.; Cui, G. *The Journal of Physical Chemistry Letters* **2018**, *9*, 6702–6708, DOI: 10.1021/acs.jpclett.8b03026.

(230) Westermayr, J.; Gastegger, M.; Menger, M. F. S. J.; Mai, S.; González, L.; Marquetand, P. *Chemical Science* **2019**, *10*, 8100–8107, DOI: 10.1039/C9SC01742A.

(231) Gasteiger, J.; Marsili, M. *Tetrahedron Letters* **1978**, *19*, 3181–3184, DOI: https://doi.org/10.1016/S0040-4039(01)94977-9.

(232) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. *Journal of the American Chemical Society* **1992**, *114*, 10024–10035, DOI: 10.1021/ja00051a040.

(233) Brockherde, F.; Vogt, L.; Li, L.; Tuckerman, M. E.; Burke, K.; Müller, K.-R. *Nature Communications* **2017**, *8*, 872, DOI: 10.1038/s41467-017-00839-3.

(234) Wang, C.-I.; Braza, M. K. E.; Claudio, G. C.; Nellas, R. B.; Hsu, C.-P. *The Journal of Physical Chemistry A* **2019**, *123*, 7792–7802, DOI: 10.1021/acs.jpca.9b04256.

(235) Szabo, A.; Ostlund, N., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Books on Chemistry; Dover Publications: 1996.

(236) Smith, J. S.; Nebgen, B.; Lubbers, N.; Isayev, O.; Roitberg, A. E. *The Journal of Chemical Physics* **2018**, *148*, 241733, DOI: 10.1063/1.5023802.

(237) Zhong, M. et al. *Nature* **2020**, *581*, 178–183, DOI: 10.1038/s41586-020-2242-8.

(238) Berne, B. J.; Harp, G. D. In *Advances in Chemical Physics*; John Wiley & Sons, Ltd: 1970, pp 63–227.

(239) Wang, S.; Li, Z.; Pan, W. *Soft Matter* **2019**, *15*, 7567–7582, DOI: 10.1039/C9SM01211G.

(240) Diebold, A. C.; Adelman, S. A.; Mou, C. Y. *The Journal of Chemical Physics* **1979**, *71*, 3236–3251, DOI: 10.1063/1.438753.

(241) Persson, B. N. J.; Ryberg, R. *Physical Review B* **1985**, *32*, 3586–3596, DOI: 10.1103/PhysRevB.32.3586.

(242) Förster, T. *Annalen der Physik* **1948**, *437*, 55–75, DOI: 10.1002/andp.19484370105.

(243) Dexter, D. L. *The Journal of Chemical Physics* **1953**, *21*, 836–850, DOI: 10.1063/1.1699044.