Karlsruhe Institute of Technology

Development of Anharmonic Molecular Models and Simulation of Reaction Kinetics in Zeolite Catalysis

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN
(Dr. rer. nat.)

von der KIT-Fakultät für Chemie und Biowissenschaften
des Karlsruher Instituts für Technologie (KIT)

genehmigte
DISSERTATION

von
Jonas Amsler

1. Referent: Prof. Dr. Felix Studt
2. Referentin: apl. Prof. Dr. Karin Fink
Tag der mündlichen Prüfung: 28. April 2022

Karlsruhe Institute of Technology
Department of Chemistry and Biosciences
Kaiserstr. 12
76131 Karlsruhe

Abstract

Zeolite materials catalyze a remarkable hydrocarbon chemistry through their unique topology with pores on the molecular scale and distributed catalytically active Brønsted acid sites. In particular, industrially viable processes like Ethanol-to-Olefins (ETO) and Methanol-to-Olefins (MTO) could conceivably become an integral part of the value-added chain from gasified biomass to petrochemicals.

Simulations offer insights into otherwise hardly observable processes, such as catalysts under reaction conditions, and thus improve the knowledge-based design of new materials and processes. Multiscale modeling at the cutting edge of zeolite catalysis, as presented in this work, builds on a collection of theories each tailored to different length and time scales. Fundamental catalytic properties on the molecular scale are determined by density functional theory (DFT) and further quantum mechanical theories. The partition functions of statistical mechanics – including the harmonic oscillator approximation – link these elementary results with thermodynamics. Rate constants derived thus within the mean field approximation enable kinetic simulations with ideal batch and plug flow reactor models.

Inspection of the rather unexplored but presumed poor quality of the (molecular) harmonic oscillator approximation for adsorption processes and reaction barriers in zeolites reveals the significance of anharmonicity in vibrations, hindered rotations and hindered translations (dissociation). To quantify anharmonicity, a new generally applicable method for calculating anharmonic corrections is reported. The novel concept involves thermodynamic λ-path integration (λ-TI) from the harmonic to the fully DFT-based system combined with curvilinear internal coordinates in systems with periodic boundary conditions. Its independence of any reaction coordinates makes this λ-TI method particularly useful for computational (zeolite) catalysis and advantageous for adsorption studies relative to the gas phase, since it can be applied to study arbitrary states in the free energy landscape. In two steps, the λ-TI method using DFT-based molecular dynamics simulation is applied to study adsorption inside zeolite H-SSZ-13 and then extended to study anharmonicity of reaction barriers. The prediction of experimental adsorption free energies is improved. Regarding barriers, the λ-TI method is demonstrated as a comparable alternative to an established ξ-TI method that depends on the reaction coordinate ξ. Machine learning and related descriptors are outlined as promising complements to the λ-TI method.

Reaction pathways for ethanol dehydration and a significant part of the autocatalytic olefin cycle in ETO are explored using the multiscale approach described above. The formation of diethyl ether is observed at lower temperatures with decreasing selectivity for increasing conversion. At higher temperatures, ethanol dehydration occurs much faster than ethylation of the formed ethene. Hexene isomers form on the same time scale as butene, with branched isomers being favored and 2-methylpentene isomers contributing most to propene formation through cracking. Surface alkoxy species enable the most relevant (stepwise) alkylation mechanism under the reaction conditions studied in zeolite H-SSZ-13 for both ETO and MTO. Furthermore, the relevance of a selection of common organic impurities for the initiation of the MTO process is quantified in a kinetic model comprising 107 elementary steps, including a representative part of the olefin cycle and direct initiation from methanol through CO-mediated C–C bond formation. The effect of different impurities on the olefin evolution is observed to vary with the impurity type and amount. Already extremely low quantities of impurities result in faster initiation than direct C–C bond formation, which only matters in complete absence of impurities.

Es ist beachtlich, in welchem Maß der Mensch die Natur rechenbar gemacht hat. Nur eines können wir prinzipiell nicht wissen, nämlich was wir künftig wissen werden. Denn sonst wüssten wir es bereits jetzt.

Logiker und Philosoph Sir Karl Popper (1902 - 1994)[1]

Kurzfassung

Entwicklung von anharmonischen Molekülmodellen und Kinetiksimulationen für Zeolithkatalyse. Zeolithe katalysieren eine bemerkenswerte Kohlenwasserstoffchemie durch ihre einzigartige Topologie mit Poren von molekularer Größenordnung und verteilten aktiven Brønsted Säurezentren. Insbesondere industriell nutzbare Prozesse wie "Ethanol-to-Olefins" (ETO) und "Methanol-to-Olefins" (MTO) könnten ein fester Bestandteil der Wertschöpfungskette von vergaster Biomasse zu Petrochemikalien werden.

Simulationen bieten direkte Einblicke in ansonsten kaum beobachtbare Vorgänge, wie zum Beispiel unter Reaktionsbedingungen arbeitende Katalysatoren, und erleichtern somit die zielgerichtete Gestaltung von neuen Materialien und Prozessen. Mehrskalenmodellierung von Zeolithkatalyse, wie in dieser Arbeit gezeigt, verbindet Theorien zur Beschreibung von Phänomenen auf verschiedensten Längen- und Zeitskalen. Grundlegende katalytische Eigenschaften auf molekularer Ebene werden durch Dichtefunktionaltheorie (DFT) und weitere Quantenmethoden bestimmt. Die Zustandssummen der statistischen Mechanik – einschließlich der Näherung des harmonischen Oszillators – verknüpfen diese fundamentalen Ergebnisse mit Thermodynamik. Dadurch abgeleitete Ratenkonstanten ermöglichen Kinetiksimulationen mit Reaktormodellen wie dem idealen Rührkessel und idealen Strömungsrohr.

Die Untersuchung der eher unerforschten, aber vermutlich mangelhaften Qualität der Näherung des (molekularen) harmonischen Oszillators für Adsorptionsprozesse und Reaktionsbarrieren in Zeolithen zeigt die Bedeutung von Anharmonizität für Schwingungen sowie für unterdrückte Rotationen und Translationen (Dissoziationen). Um Anharmonizität zu quantifizieren wird eine neue allgemein anwendbare Methode zur Berechnung anharmonischer Korrekturen vorgestellt. Das neue Konzept bedient sich thermodynamischer λ-Pfadintegration (λ-TI) vom harmonisch zum DFT-basiert interagierenden System kombiniert mit krummlinigen internen Koordinaten in Systemen mit periodischen Randbedingungen. Seine Unabhängigkeit von jeglichen Reaktionskoordinaten macht dieses λ-TI Verfahren besonders nützlich für computergestützte (Zeolith-)Katalyse und vorteilhaft für die Betrachtung von Adsorption relativ zur Gasphase, da es zur Untersuchung beliebiger Zustände in der freien Energiehyperfläche angewendet werden kann. In zwei Schritten wird das λ-TI Verfahren unter Verwendung von DFT-basierter Molekulardynamik Simulation zur Untersuchung von Adsorption im Zeolith H-SSZ-13 und von Reaktionsbarrieren eingesetzt. Die Vorhersage experimenteller freier Adsorptionsenthalpien wird verbessert. Für Barrieren erweist sich das λ-TI Verfahren als vergleichbare Alternative zu einer etablierten, von der Reaktionskoordinate abhängigen, Methode. Maschinelles Lernen und damit verbundene Deskriptoren werden als vielversprechende Ergänzungen zum λ-TI Verfahren skizziert.

Reaktionspfade der Ethanoldehydratisierung und eines wesentlichen Teils des autokatalytischen Olefinzyklus in ETO werden unter Verwendung der oben beschriebenen Mehrskalenmodellierung untersucht. Die Bildung von Diethylether wird bei niedrigeren Temperaturen mit abnehmender Selektivität für steigenden Umsatz beobachtet. Bei höheren Temperaturen verläuft die Ethanoldehydratisierung viel schneller als die Ethylierung des entstehenden Ethens. Hexenisomere bilden sich auf der gleichen Zeitskala wie Buten, wobei verzweigte Isomere bevorzugt werden, und 2-Methylpentenisomere am meisten zur Propenbildung durch Spaltung beitragen. Oberflächengebundene Alkoxygruppen ermöglichen den relevantesten (schrittweisen) Alkylierungsmechanismus unter den untersuchten Reaktionsbedingungen in H-SSZ-13 sowohl bei ETO als auch bei MTO. Außerdem wird die Relevanz einer Auswahl gängiger

organischer Verunreinigungen für die Initiierung des MTO Prozesses in einem kinetischen Modell mit 107 Elementarschritten, darunter ein repräsentativer Teil des Olefinzyklus und die direkte Initiierung aus Methanol durch CO-vermittelte C–C Bindungsknüpfung, quantifiziert. Die Wirkung verschiedener Verunreinigungen auf die Olefinentwicklung variiert mit ihrer Art und Menge. Bereits extrem kleine Verunreinigungsmengen führen zu schnellerer Initiierung als die direkte C–C Bindungsknüpfung, die nur in vollständiger Abwesenheit von Verunreinigungen von Bedeutung ist.

Contents

Part I.

Introduction

1. Motivation

1.1. The Fossil Fuel Dilemma

Fossil fuels such as natural gas, crude oil and coal are needed to satisfy a major part of the ever increasing global energy demand (85 % of 576 EJ[a] in 2018).[2,3] Moreover, crude oil is an essential source of hydrocarbons for the petrochemical industry.[4] In the foreseeable future, however, a decline in crude oil production from natural reserves can be expected as depicted in Figure 1.1.[1] This dilemma of increasing demand but decreasing supply makes affordable alternatives to fossil fuels highly desirable.[b] At the same time, over 100 years of linear energy economy have emitted enough greenhouse gases to cause a global warming, bringing along increasingly problematic climate changes.[6] A circular energy economy, where the feedstock of petrochemicals is shifted from fossil fuels to renewable biomass, could potentially alleviate the accumulation of emissions and supply petrochemicals on a sustainable basis.

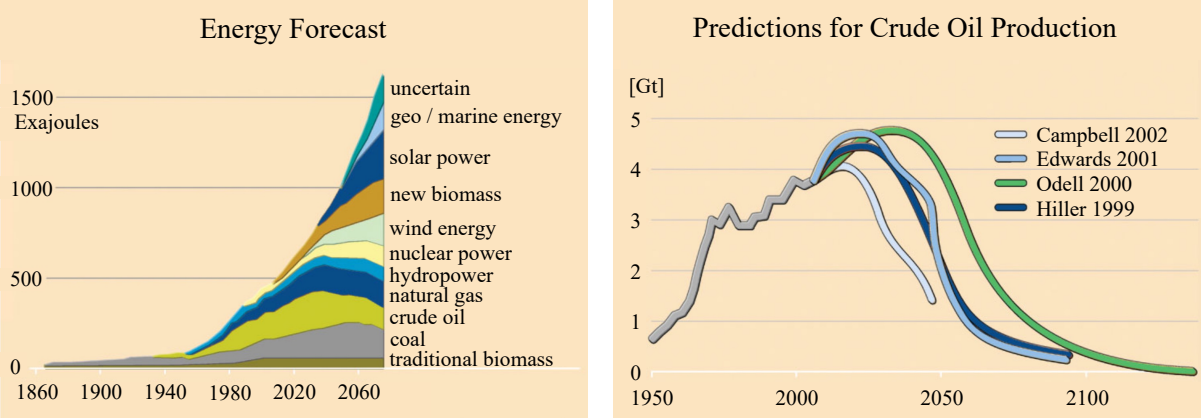


Figure 1.1.: Left: Forecast of global energy consumption partitioned by power source.[2] Right: Predictions for global crude oil production.[7–10] Adapted with permission from [1]. Copyright 2005 Wiley.

1.2. The Call for Catalysts

Catalysts are key materials[c] in any future energy scenario as over 85 % of products in the chemical industry are made in catalytic processes.[11] Only the right catalysts can ameliorate the chemistry

[a] Exajoule: 1 EJ = 10^{18} J.

[b] Hydrocarbons as a chemical energy storage are generally very practical due to their noteworthy high energy density compared to electric batteries and their uncomplicated storage and transport properties compared to hydrogen.[5]

[c] In short, the principle of Sabatier states that catalysts facilitate the molecular hurdle race by offering smoother reaction pathways. By definition, a catalyst offers an alternative, energetically favorable reaction mechanism to a non-catalytic reaction without changing the thermodynamic equilibrium and without being consumed or altered after a catalytic cycle. Catalysts can accelerate reaction rates or enable processes under more benign reaction conditions (pressure, temperature).

involved in our enormous venture, a petrochemical revolution. However, the knowledge-based design and improvement of new catalytic materials pose a major technical challenge. Catalysis, as a nanotechnology, is a highly complex field where mechanistic insight can be obscured by effects from different length and time scales. To systematically disentangle the situation, research is increasingly supported by quantum chemical calculations of reaction mechanisms and kinetic modeling of corresponding reaction rates.[12] The fossil fuel dilemma is likely not the last call for catalysts, but can it be answered?

1.2.1. Zeolites and the Chabazite Framework

Zeolite science is a promising candidate in the race to answer the call for heterogeneous catalysts in the fossil fuel dilemma. Zeolites are a unique class of oxides, consisting of microporous, crystalline aluminosilicates with the general formula $M^{z+}_{m/z} \cdot [(AlO_2)^-_m \cdot (SiO_2)_n] \cdot x\ H_2O$. Typically, M is an (earth-) alkali metal, ammonium or a proton for charge balance. Zeolites form a variety of characteristic crystal structures, some of which appear naturally, comprised of tetrahedral units AlO_4 and SiO_4 sharing the O corners.[4,11] Artificial zeolite synthesis is typically conducted in aqueous medium, for instance via hydrothermal zeolitization from supersaturated solutions in autoclaves at 300 K to 500 K.[13] Due to their large inner surface area of several tens to hundreds of $m^2\ g^{-1}$,[14] the flexible water content of some zeolite minerals can take up to 50 % of the total volume.[15] This available volume – in the order of $0.15\ mL\ g^{-1}$ to $0.35\ mL\ g^{-1}$ – also makes zeolites readily invite other guest molecules that can reside, adsorb and react inside.[11] The interior 3D topology extends in diverse open networks of channels, or in cavities linked by shared windows.[15] Protonated zeolites can be understood as gigantic inorganic polyacids with distributed single Brønsted acid sites that are particularly suitable for catalyzing hydrocarbon chemistry. Since that was recognized, industrial applications of zeolites as catalysts range from hydrocracking, fluid catalytic cracking (FCC) and exhaust gas aftertreatment to the hydrocarbon production processes introduced below.[4,11] Of the many existing topologies only the chabazite (CHA) framework type shown in Figure 1.2 is relevant for this work. Herein, zeolite H-SSZ-13 and its iso-structural zeotype H-SAPO-34 (strictly a silicoaluminophosphate) are investigated using a structural model with one acid site per unit cell. Compared to H-SSZ-13, Al atoms are replaced by Si while Si is alternately replaced by Al or P.[11] This difference in composition is clearly revealed in the lower acidity of H-SAPO-34 that is measurable indirectly by the heat of adsorption of ammonia.[18–24]

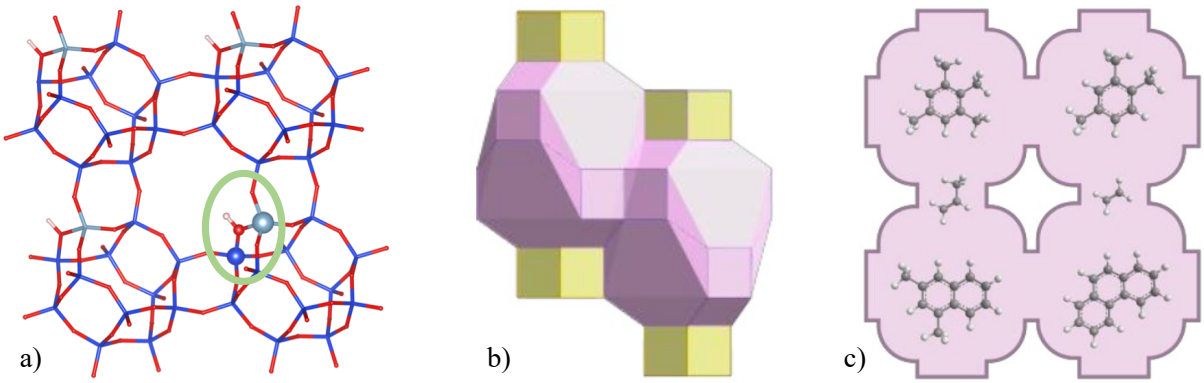


Figure 1.2.: The 8-ring structure of the CHA framework is composed of large cavities with small window openings. (a) Model structure of H-SSZ-13 (CHA) with one acid site per cavity and a Si to Al ratio of 35:1. The highlighted acid site is also the adsorption site for guest molecules. Color code: Si, dark blue; Al, light blue; O, red; H, white. (b) Shape of cavities (pink) and windows (yellow) in the CHA framework. (c) Large hydrocarbons such as aromatics can be formed inside the cavities but cannot escape them. Small molecules easily diffuse through the pores. Adapted compilation of images with permissions from [16, 17]. Copyright 2018 American Chemical Society (a) and 2021 Springer Nature (b, c), respectively.

In general, the CHA framework features large cavities with small 8-ring window openings entailing consequences for the selectivity of any hydrocarbon chemistry inside. Its unique pore size restricts access and exit to molecules below a certain size (reactant and product selectivity). The shape and confinement of the cavities further restricts the possible reactions (reaction selectivity). This oftentimes advantageous property can also turn into a problem when trapped polycondensed species and carbon deposition lead to catalyst deactivation via internal and external coking.[16]

1.2.2. Hydrocarbons from Alkanols

Short alkanols, such as methanol and ethanol, are part of the value-added chain from gasified biomass to petrochemicals. Methanol can even be obtained from direct recycling of carbon dioxide as demonstrated in a pilot plant on Iceland.[25,26] In particular, the production of hydrocarbons from these alkanols could conceivably become a mosaic building block in resolving the fossil fuel dilemma. Industrially viable processes like methanol-to-hydrocarbons (MTH) can yield mixtures of different composition tuned towards olefins (MTO) or gasoline (MTG) when operated under the right combination of catalyst and reaction conditions (H-SAPO-34 catalyst for MTO at 670 K to 820 K and 1 bar to 5 bar, H-ZSM-5 catalyst for MTG at slightly lower temperatures but typically higher pressures).[27,28] Related processes with ethanol, like ethanol-to-olefins (ETO), are in their laboratory stages or on the pilot scale.[29]

Since the proposal by Dahl and Kolboe,[30–32] the mechanistic details of such alkanol conversions are explained by the hydrocarbon pool (HCP) comprised of a variety of (unsaturated) hydrocarbons that continuously undergo oligomerization and cracking reactions.[27,33–36] Extensive experimental and theoretical studies on methanol based processes have led to the dual-cycle concept[37] depicted in Figure 1.3. Olefins are formed by repeated alkene methylation[38–40] and cracking[16,27] in the olefin cycle, but also through dealkylation of aromatics in the aromatic cycle, where aromatics are methylated.[41] Olefins and aromatics thus act as co-catalysts to form more hydrocarbons causing an autocatalytic behavior of the process. The chemical reactions in the HCP – alkylation/oligomerization, isomerization, cyclization, aromatization, hydrogen transfers, dehydration, and cracking – are catalyzed by zeolites. The Brønsted acid sites of these catalysts stabilize the carbocationic transition states of the hydrocarbons

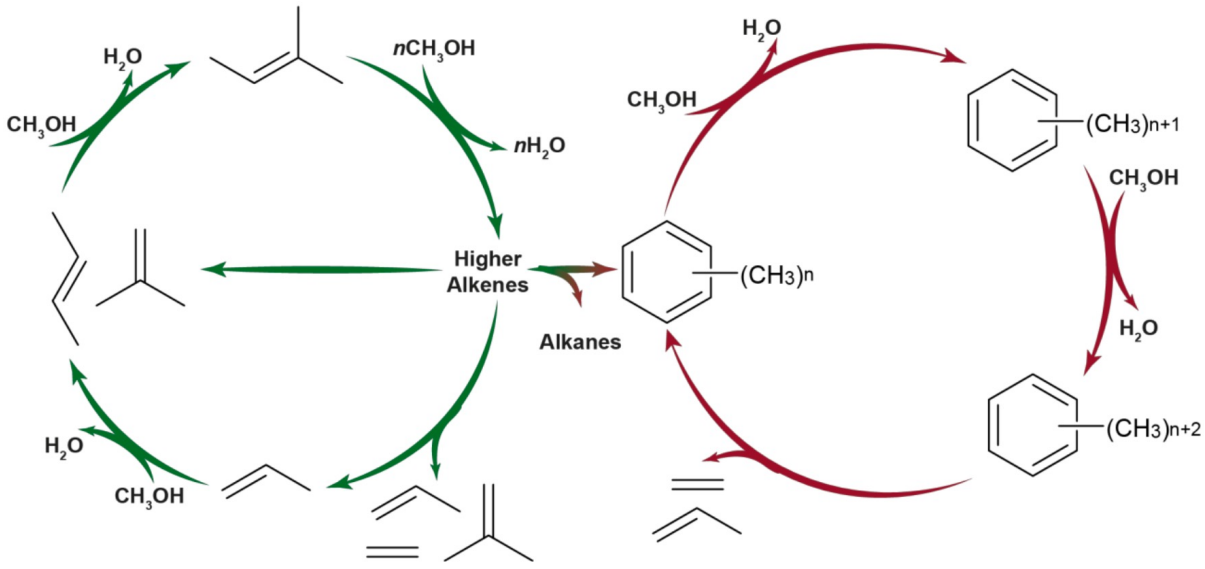


Figure 1.3.: Dual-cycle concept for the conversion of methanol during the MTH reaction. Olefin cycle (left) and aromatic cycle (right) form an autocatalytic reaction network. Reprinted with permission from [16]. Copyright 2018 Springer Nature.

thus favoring the production of branched hydrocarbons.[27] Acid strength and acid site density aside, the catalyst topology also significantly impacts the product distribution of the HCP.[42–45] Due to the small window size of around 3.7 Å,[24] catalysts of the CHA framework type have shown excellent shape selectivity and control over the product distribution of light olefins with up to four carbon atoms but also a tendency to more severe deactivation. Large hydrocarbons and particularly aromatics are trapped upon formation inside the cavities of the CHA framework. In the prevailing aromatic cycle, aromatics are thus doomed to autocatalytically incubate light olefins that can diffuse into the gas phase with high selectivities.[16] Although many aspects of MTH reactions are now largely understood,[27] certain details like the initiation of the first C–C bond formation from pure methanol are discussed controversially in the literature.[46–48] Further insights into the structure-performance correlation of zeolites for the HCP and understanding in depth how shape selectivity emerges from the confinement by the zeolite framework may lead to improved design principles that can provide answers to the call for catalysts. The search for novel materials continues where grand challenges have to be tackled.

1.3. Grand Challenges in Computational Zeolite Catalysis

The pores on the scale of molecules, the large inner surface area and the Brønsted acid sites make zeolites ideal molecular sieves, adsorbents and catalysts with versatile applications. Notably, the combination of zeolite composition, acid strength, acid site density and topology creates ample scope for catalyst fine tuning within the limits of the synthetic arts and crafts. Several characteristic length scales become evident from a fundamental modeling perspective, for instance, there are the active sites with elementary reaction steps (sub-nanometer regime), the zeolite framework (nanometers), zeolite grains and grain boundaries (mesoscale), shaped catalysts and extrudates (centimeters) and the catalyst bed in the reactor (meters). Likewise, characteristic time scales underlying zeolite catalysis range from femtoseconds for electronic motion governing intrinsic chemistry, picoseconds for the formation and breaking of bonds, microseconds for elementary steps inside zeolite pores, seconds for entire reaction cycles and (pore) diffusion, to minutes and hours for whole processes, like catalyst deactivation, in reactors.[11,49] This work generally follows a parameter-passing multiscale modeling approach (one-way, bottom-up), as depicted in Figure 1.4, covering quantum mechanics, molecular dynamics (MD) and kinetic simulation – fluid dynamics and effects of any larger scales are omitted. Although the dimensions change gradually and, on most scales, continuously, scales are commonly divided into

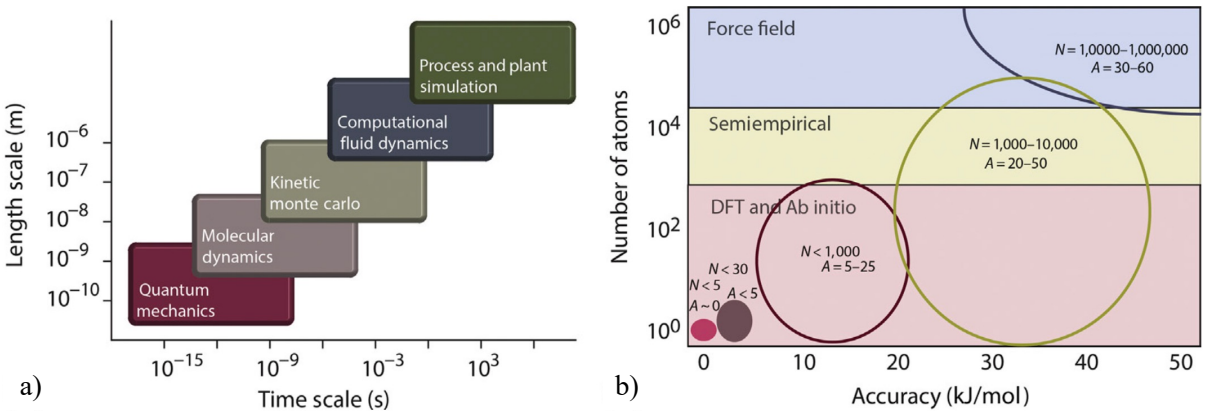


Figure 1.4.: (a) Scientific theories and models are stacked to bridge length and time scales. Results from more fundamental models are passed on as parameters to models of larger scales. (b) Hierarchy of different computational methods (N stands for the number of atoms, A is achievable accuracy in kJ mol^{-1}). Reprinted with permission from [49]. Copyright 2016 Elsevier.

Table 1.1.: Exemplary relations between uncertainty in free energy barriers and corresponding uncertainty in reaction rates.

Uncertainty of barrier ΔE (kJ mol^{-1})	30			20			5		
Temperature (K)	300	500	700	300	500	700	300	500	700
Factor by which rate would change[a]	167000	1400	170	3000	120	30	7	3	2

[a] Corresponds to the Boltzmann factor, $\exp\left(-\frac{\Delta E}{RT}\right)$, at the temperature T, where R is the ideal gas constant.

intervals, each with plausible approximations and corresponding achievable accuracy for the theories and models of the physical and chemical processes. The challenge of linking theories from different scales is tackled by scale-bridging workflows, where the aim is to stack models for different scales in a complementary way that one model produces the input parameters for the other. For instance, quantum mechanics can compute electronic energy and sometimes drive molecular dynamics, whereas statistical mechanics and molecular dynamics can estimate enthalpy and entropy thus upgrading electronic (free) energy to (configurational) free energy including zero-point vibrational energy. Macroscopic rate constants obtained from these free energies can be plugged into microkinetic models and kinetic Monte Carlo schemes. Further considerations of gradients in temperature, pressure and concentrations, as well as transport effects, including diffusion and fluid dynamics, ultimately lead to the simulation of whole reactors or even production plants. When appropriate, some effects of certain scales can be omitted by taking adequate approximations – in processes with no transport limitation, for instance, diffusion could be ignored by assuming instantaneous mixing.

On the smallest scale, computational zeolite catalysis deals with structural models representing the zeolite material. Fortunately, zeolites are equipped with well-defined catalytically active sites in comparatively rigid, well-known crystal structures that do not undergo rigorous rearrangements (as compared to supported metal catalysts with a plethora of active site motifs each changing over time with new elementary steps and different rates). Structural models typically comprise a unit cell in periodic boundary conditions (PBC) or an isolated atom arrangement centered around the active site. Zeolite grain boundaries are thus excluded *a priori*. Furthermore, the number of acid sites and their locations inside the unit cell are often chosen based on accessibility, stability or experimental data.[50] Clearly, not all acid site motifs in synthetic zeolites are well-represented by such a choice. It is known that the position of the active site matters[51] and that location and proximity influences reactivity by cooperation of adjacent sites. This aspect is important to consider, as sometimes a small number of active sites in a special configuration dominates the overall reaction rate. Nevertheless, the commonly applied model of isolated acid sites is generally believed to give a good first estimate, also for zeolites with higher Al to Si ratios.[52,53] Further aspects, like the role of metal cations inside zeolite micropores, are not relevant for this work.

One of the grandest challenges in computational catalysis in general[12] is the crucial theoretical junction between free energy barriers of reactions and corresponding rate constants. From statistical mechanics it is known that the rate constant of a reaction step changes drastically with minor changes of its reaction barrier – for instance, the rate of a reaction at 500 K changes by a factor of about 120 for an uncertainty range of 20 kJ mol^{-1}, or by a factor of approximately 3 for an uncertainty range of 5 kJ mol^{-1}, see Table 1.1. The simulation of reactions inside zeolites involves large atom arrangements with oftentimes complex transition structures. In such systems, however, only approximate methods are computationally feasible for the calculation of enthalpy and entropy of reaction barriers.

In a compromise between accuracy and computational cost, the electronic energy of zeolite models can be computed with density functional theory (DFT). Usual procedures employ generalized gradient approximation (GGA) functionals and a plane-wave basis for zeolite models with PBC. Mainly

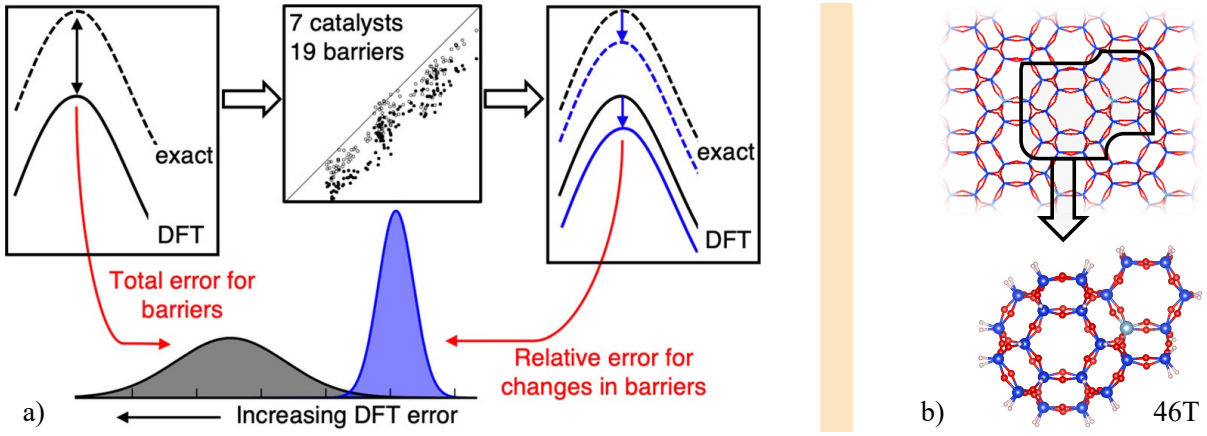


Figure 1.5.: (a) Schematic probability distribution for DFT errors in computational zeolite catalysis. DFT underestimates total energies for barriers while preserving relative energy differences between barriers well with a narrow error distribution ($\pm 5\,\mathrm{kJ\,mol^{-1}}$). Reprinted with permission from [62]. Copyright 2020 American Chemical Society. (b) Illustration of the periodic model and the 46T cluster model used in this work. The cluster model comprises one whole cavity with one active site. Terminal bonds are saturated with H atoms. Color code: Si, dark blue; Al, light blue; O, red; H, white.

van-der-Waals (vdW) interaction makes zeolite pores exert strong local confinement effects on the hydrocarbons inside.[51,54–56] An adequate description of these dispersive long-range effects originating from the attraction between induced dipoles (instantaneous electron correlation) is therefore crucial for the investigation of adsorption and of reactions inside zeolite pores. At longer length and time scales, transport effects like diffusion limitation are particularly important for larger molecules, where diffusion barriers correlate with the vdW-radius of the diffusing molecule.[57] The known shortcoming of DFT regarding vdW interactions can be countered by dispersion corrected DFT-D approaches.[58–60] Nevertheless, GGA functionals (PBE-D3) are prone to have intrinsic errors of up to $40\,\mathrm{kJ\,mol^{-1}}$ for the prediction of single reaction steps corresponding to a large uncertainty in reaction rates.[61] Fortunately, recent work has confirmed that the error is systematic and the accuracy for the prediction of differences of the same reaction barrier in different zeolites is within $\pm 5\,\mathrm{kJ\,mol^{-1}}$,[62] see Figure 1.5a. This high precision when comparing zeolites among themselves justifies the widespread use of DFT for computational studies and screenings of zeolite materials. Fast computational screenings usually require a reductionist approach to reduce the complexity by identification and analysis of principal components. Scaling relations are such a tool that enables predictive modeling of reaction energies and barriers using simple descriptors.[63–67]

A higher precision of enthalpies can be achieved by the hierarchical cluster model approach[68–73] with computationally more demanding theories, such as MP2 or DLPNO-CCSD(T) (see Section 2.2.1), that are typically only affordable for a finite subset of atoms – the 46T cluster model is shown in Figure 1.5b. Procedures including a full basis set extrapolation[61] are also used in this work.

Methods for the calculation of accurate entropies via the partition function are far less developed – analytical expressions for the vibrational partition function exist only within idealized models, such as the harmonic oscillator approximation. More accurate methods are emerging[74–77] but are often computationally demanding and imply a trade-off with respect to the accuracy of the corresponding enthalpy.[12] A good example is given by the thermodynamic λ-path integration (λ-TI) approach[17] presented in this work, where molecular dynamics (MD) simulations yield more reliable entropies, but are only feasible at the GGA level of DFT. Faster algorithms, machine learning (ML) and quantum computing[78] are expected to move the frontiers of what is possible today.[12] For instance, given

a sufficiently large and well-defined data base, ML can identify key parameters and correlations of interest – moreover, ML force fields can significantly speed up the exploration of phase space.

The complex reaction network of hydrocarbons in zeolites makes assumptions about the reaction rate based on the energetic span model[79] difficult. Microkinetic models can be invoked to translate the free energy landscape into simulated conversion and selectivity. For instance, detailed kinetic models are essential to capture the autocatalytic nature of MTH processes.[80] Linking various length and time scales with accurate free energy calculations in a multiscale modeling approach is still presenting a major challenge.[81,82] Keeping in mind the problem of fairly large errors of rate constants even for extremely well described reaction steps, the improvement of kinetic models in a multiscale approach goes along with more accurate free energy data.[12]

Providing direct information on the otherwise hardly observable state of active sites under reaction conditions is a primary epistemic goal of computational zeolite catalysis. Nevertheless, theory always thrives along complimentary experimental expertise. Eventually, the unambiguous identification of reaction intermediates and mechanisms relies on simulation combined with validation data from carefully conducted characterization and kinetic studies on well-defined systems.[12]

1.4. Scope of this Work

The grand challenges in computational zeolite catalysis discussed above set the scene for this work. Adsorption phenomena and hydrocarbon chemistry in zeolite H-SSZ-13, a potential candidate for industrial MTO and ETO processes, are simulated using a hierarchy of theories tailored to bridge length and time scales. This bridging of scales generally follows a bottom-up approach beginning with molecular models. Taking advantage of the hierarchical cluster model approach introduced above, electronic structure theories (DFT, MP2, CC) are applied to atom positions on the smallest scale. Relevant samples on the free energy landscape are computed using the standard molecular partition functions (see page 18), including the harmonic oscillator approximation. The rather unexplored but presumed poor quality of this harmonic approximation for adsorption processes in zeolites and in general is inspected in Chapter 3, where a new generally applicable computational method for the calculation of anharmonic corrections to the harmonic approximation is reported. In two steps, this new application of thermodynamic λ-path integration (λ-TI) is developed for adsorbates and then extended to reactions with transition states. Using transition state theory with the mean-field approximation, rate constants are derived from the calculated free energy profiles and fed into ideal reactor models for batch and plug flow reactors. The obtained microkinetic models are employed to explore and evaluate reaction pathways for the MTO and ETO processes that are difficult to decipher experimentally. Chapter 4 presents three studies based on the parameter passing multiscale approach from the fundamentals to the reactor model. The first study compares the dehydration of ethanol to diethyl ether and ethene in zeolite H-SSZ-13 to zeotype H-SAPO-34. In the second study, the formation of the HCP in the ETO process is explored for a significant part of the olefin cycle including ethanol dehydration, olefin ethylations up to hexene isomers and the corresponding cracking reactions. The second study concludes with a comparison of methylation and ethylation barriers in MTO and ETO, respectively. In the third study, the role of omnipresent impurities in the initiation of the MTO process is investigated and quantified using a kinetic model comprising 107 elementary steps with reaction barriers from quantum mechanical calculations. An unconventional application of machine learning related descriptors as “universal” coordinates for λ-TI is addressed in a brief intermediate communication suggesting an interesting direction of future work.

2. Computational Methods for Zeolite Catalysis

2.1. On the Hierarchy of Theories

Bridging the scales. Multiscale modeling of heterogeneous catalysts builds on a collection of theories each describing the natural phenomena at the corresponding length and time scale. Elementary catalytic properties on the submicroscopic scale are determined by electronic structure and quantum mechanical wave functions. The partition functions of statistical mechanics relate these elementary results to measurable averages of large ensembles. Rate constants computed from such molecular simulations build the bridge to the largest scale considered in this work, where kinetic models are employed to study the macroscopic evolution of partial pressures in batch and plug flow reactor models. Below is an introduction to the most relevant theories in the standard terminology of chemistry textbooks.[83–88]

High-performance computing. Scientific computing has benefited a lot from the development of computer architectures with great computational power and data storage. Modern systems are capable to routinely harness several hundreds of central processing units simultaneously for a single simulation. Thanks to the bwHPC initiative,[a] such computational resources and capacities were employed for the by far most demanding, although highly parallelized, electronic structure calculations in this work.

2.2. Electronic Structure Theory

Quantum chemistry. Chemical systems at the atomic scale can be described by quantum mechanics.[b] In *A Short History of Nearly Everything* by Bryson,[89] the comprehension inherent in the periodic table is paraphrased by stating that "protons give an atom its identity, electrons its personality." In fact, the chemical properties of molecules are widely ascribed to the involved electrons and their electromagnetic interactions. Electrons classify as indistinguishable fermions of light weight with one negative elementary charge. Moreover, electrons obey Fermi-Dirac statistics due to their intrinsic spin of $\frac{1}{2}$. Fundamental laws of physics combined with sophisticated mathematical methods reconcile their dual nature as particles and waves to a seemingly paradox picture.[c] The different time scales of the movements of atomic nuclei and electrons justify the separation of nuclear and electronic degrees of freedom in the Born-Oppenheimer approximation. In consequence, many chemical processes can be thought of the atomic nuclei moving slowly in a potential governed by the electrons which adjust quasi instantly to the positions of the atomic nuclei. The potential energy and its gradient with respect to atom positions can be computed for given nuclear coordinates from the electrons and their interactions by techniques such as wave function methods and density functional theory (DFT).

[a] In bwHPC, the infrastructure and projects are funded by the ministry of science, research and arts (MWK) of the state of Baden-Württemberg, Germany, by the German research society (DFG) and by ten universities of Baden-Württemberg.

[b] Where physical properties, such as energy and momentum, of bound systems are quantized (restricted to discrete values).

[c] As occasionally credited to Heisenberg, "Not only is the Universe stranger than we think, it is stranger than we can think."

2.2.1. Wave Function Methods

Electronic wave function. Wave function methods describe a quantum system by the concept of the quantum mechanical wave function from which physical observables can be computed as eigenvalues of the corresponding operators. The extraction of the electronic wave function $\mathbf{\Psi}$ from the wave function for the whole system is reasonable in the Born-Oppenheimer approximation. Electronic energy E can then be computed by solving the time-independent non-relativistic electronic Schrödinger eq. (2.1).

$$\mathbf{H}\mathbf{\Psi} = E\mathbf{\Psi} \tag{2.1}$$

Electronic Hamilton operator. The electronic Hamilton operator $\mathbf{H}$ comprises the kinetic energy of the electrons $\mathbf{T}_\mathrm{e}$, the attraction between nuclei and electrons $\mathbf{V}_\mathrm{ne}$ as well as the repulsion between electrons $\mathbf{V}_\mathrm{ee}$. The repulsion of the nuclei $\mathbf{V}_\mathrm{nn}$ is a constant for constant nuclear coordinates due to the Born-Oppenheimer approximation. In atomic units ($m_\mathrm{e} = e = \hbar = \frac{4\pi}{\varepsilon_0} = 1$) $\mathbf{H}$ can be expressed with the number of charges Z, the electron coordinates r and the nuclear coordinates R. The zero-point thus corresponds to a state where nuclei and electrons rest at infinite distance from each other ($\mathbf{H} = 0$).

$$\mathbf{H} = \mathbf{T}_\mathrm{e} + \mathbf{V}_\mathrm{ne} + \mathbf{V}_\mathrm{ee} + \mathbf{V}_\mathrm{nn} = -\frac{1}{2}\sum_i^{N_\mathrm{e}} \nabla_i^2 - \sum_a^{N_\mathrm{nuc}} \sum_i^{N_\mathrm{e}} \frac{Z_a}{|\mathbf{R}_a - \mathbf{r}_i|} + \sum_i^{N_\mathrm{e}} \sum_{j>i}^{N_\mathrm{e}} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_a^{N_\mathrm{nuc}} \sum_{b>a}^{N_\mathrm{nuc}} \frac{Z_a Z_b}{|\mathbf{R}_a - \mathbf{R}_b|} \tag{2.2}$$

The Hamilton operator can be rewritten to $\mathbf{H} = \sum_i^{N_\mathrm{e}} \mathbf{h}_i + \sum_{j>i}^{N_\mathrm{e}} \mathbf{g}_{ij} + \mathbf{V}_\mathrm{nn}$ with $\mathbf{h}_i = -\frac{1}{2}\nabla_i^2 - \sum_a^{N_\mathrm{nuc}} \frac{Z_a}{|\mathbf{R}_a - \mathbf{r}_i|}$ describing the motion of electron i in the field of all nuclei and the electron repulsion $\mathbf{g}_{ij} = \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$.

Slater determinant. According to the Pauli exclusion principle, electrons in a quantum system cannot occupy the same quantum state simultaneously. The wave function must be antisymmetric (change its sign) with respect to interchange of any two electron coordinates. The simplest wave function that satisfies the Pauli exclusion principle for a system with N_e electrons is the ansatz of a Slater determinant $\mathbf{\Phi}$ with orthonormalized ($\langle \phi_i | \phi_j \rangle = \delta_{ij}$) single-electron wave functions ϕ_i, called orbitals, in the columns and electron coordinates along the rows. For a molecule, these orbitals are molecular orbitals (MO), which are given as the product of a spin function α or β and a spatial orbital.

$$\mathbf{\Phi}(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{N_\mathrm{e}}) = \frac{1}{\sqrt{N_\mathrm{e}!}} \begin{vmatrix} \phi_1(\boldsymbol{x}_1) & \phi_2(\boldsymbol{x}_1) & \cdots & \phi_n(\boldsymbol{x}_1) \\ \phi_1(\boldsymbol{x}_2) & \phi_2(\boldsymbol{x}_2) & \cdots & \phi_n(\boldsymbol{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\boldsymbol{x}_{N_\mathrm{e}}) & \phi_2(\boldsymbol{x}_{N_\mathrm{e}}) & \cdots & \phi_n(\boldsymbol{x}_{N_\mathrm{e}}) \end{vmatrix} \tag{2.3}$$

Hartree-Fock theory. Hartree-Fock (HF) theory uses a single Slater determinant as the trial wave function. Additional determinants lead to more exact solutions. According to the variational principle the ground state of the system is given by the normalized wave function ($\langle \mathbf{\Phi} | \mathbf{\Phi} \rangle = 1$) with the minimum electronic energy E. The expectation value for E is expressed in terms of $\mathbf{h}_i$ and V_nn, as well as Coulomb repulsion and exchange interaction represented by the operators $\mathbf{J}_i$ and $\mathbf{K}_i$, respectively.

$$E = \langle \mathbf{\Phi} | \mathbf{H} | \mathbf{\Phi} \rangle = \sum_i^{N_\mathrm{e}} \langle \phi_i | \mathbf{h}_i | \phi_i \rangle + \sum_i^{N_\mathrm{e}} \sum_{j>i}^{N_\mathrm{e}} \langle \phi_j | \mathbf{J}_i | \phi_j \rangle - \langle \phi_j | \mathbf{K}_i | \phi_j \rangle + V_\mathrm{nn} \tag{2.4}$$

It turns out that the variation in electronic energy $\delta E = \sum_i^{N_\mathrm{e}} \langle \delta\phi_i | \mathbf{F}_i | \phi_i \rangle + \langle \phi_i | \mathbf{F}_i | \delta\phi_i \rangle$ is related to the Fock operator $\mathbf{F}_i = \mathbf{h}_i + \sum_j^{N_\mathrm{e}} \mathbf{J}_j - \mathbf{K}_j$, which is an effective one-electron energy operator describing

the kinetic energy of one electron i and its attraction to all nuclei, as well as the repulsion to all other electrons j. The set of ϕ_i that minimizes E can be obtained by performing a constrained optimization of $\mathbf{\Psi}$ using Lagrangian multipliers. In this context, ϕ_i are transformed into canonical MOs ϕ_i' to arrive at (a diagonal matrix of Lagrangian multipliers and) the HF eqs. (2.5), a set of eigenvalue eqs. where the Fock operator depends on all occupied MOs. From an initial guess for the wave function, the minimum E in this self-consistent field (SCF) has to be determined iteratively, because a specific Fock orbital can only be determined if all other occupied orbitals are known.

$$\mathbf{F}_i\phi_i' = \varepsilon_i\phi_i' \tag{2.5}$$

Basis sets. In the basis set approximation, each $\phi_i = \sum_{\alpha}^{M_{\text{basis}}} c_{\alpha i}\chi_\alpha$ is expanded with coefficients $c_{\alpha i}$ in terms of basis functions χ_α, for instance Gaussian functions. This linear combination of atomic orbitals (LCAO) leads to the Roothaan-Hall eqs. (2.6), corresponding to the HF eqs. in the atomic orbital (AO) basis set. The M_{basis} eqs. are collected in a generalized eigenvalue problem in matrix notation with the Fock matrix $\mathbf{F}$, the square matrix of expansion coefficients $\mathbf{C}$, the overlap matrix $\mathbf{S} = \left\langle \chi_\alpha \middle| \chi_\beta \right\rangle$ and the orbital energies in a diagonal matrix $\boldsymbol{\varepsilon}$. Diagonalization of $\mathbf{F}$ yields a normalized set of eigenvalues (one-electron energies) and orthogonal eigenvectors (MOs). The set of MO coefficients is varied until the energy is converged. In the unrestricted HF (UHF) approach for open shell systems, the Roothaan-Hall eqs. are solved separately for α and β spin functions, which leads to deviations (spin contamination) from the expectation value of the squared spin operator $\left\langle \mathbf{S}^2 \right\rangle$ with the exact eigenvalues $S(S+1)$. The singlet states in this work are well described by closed shell wave functions and restricted HF (RHF).

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon} \tag{2.6}$$

Electron correlation. A systematic choice of larger basis sets can improve the results systematically up to the HF complete basis set limit. However, the exact solution cannot be determined since the HF approach neglects electron correlation – the electron repulsion is only included as an average effect. The energy difference between HF and the exact solution in the given basis set is called electron correlation energy $E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}}$. Methods like Møller-Plesset second order perturbation theory (MP2) and Coupled Cluster (CC) treat electron correlation with an ansatz of multiple Slater determinants (electron configurations). Electron correlation energy is recovered from formally excited states of the HF ground state where electronic energy is higher. Comparable to the expansion of an unknown MO in the function space (similar to a coordinate system) of the basis functions, the wave function is described in the function space of Slater determinants $\mathbf{\Phi}_i$ with expansion coefficients a_i, see eq. (2.7). The basis set determines the size of the one-electron basis (thus limits the MOs), whereas the number of determinants determines the size of the many-electron basis (thus limits the description of electron correlation).

$$\mathbf{\Psi} = a_0\mathbf{\Phi}_{\text{HF}} + \sum_i a_i\mathbf{\Phi}_i \tag{2.7}$$

Møller-Plesset perturbation theory. Perturbation theory finds solutions to problems that differ by small perturbations from problems with known solutions. Møller-Plesset perturbation theory defines the unperturbed Hamiltonian operator $\mathbf{H}_0 = \sum_i^{N_e} \mathbf{F}_i$ as a sum over the Fock operators. In the zeroth order (MP0), the wave function is the HF determinant while the energy $E(\text{MP0})$ is the sum of MO energies, $E(\text{MP0}) = E_{\text{MP0}} = \langle \mathbf{\Phi}_0 | \mathbf{H}_0 | \mathbf{\Phi}_0 \rangle = \sum_i^{N_e} \varepsilon_i$. The difference to the full Hamiltonian is the perturbation operator $\mathbf{H}' = \mathbf{H} - \mathbf{H}_0 = \mathbf{V}_{\text{ee}} - 2\langle \mathbf{V}_{\text{ee}} \rangle$. Overcounting of the electron repulsion in MP0 is corrected in the first order term $E(\text{MP1}) = E_{\text{MP0}} + E_{\text{MP1}} = E_{\text{HF}}$ with $E_{\text{MP1}} = \langle \mathbf{\Phi}_0 | \mathbf{H}' | \mathbf{\Phi}_0 \rangle$. The addition of the average of $\mathbf{H}'$ over the MP0 wave function yields exactly the HF energy E_{HF}. Electron correlation energy is

obtained at second order, $E(\mathrm{MP2}) = E_{\mathrm{MP0}} + E_{\mathrm{MP1}} + E_{\mathrm{MP2}}$, and above. The MP2 energy correction E_{MP2} is obtained from the sum over doubly excited determinants $\mathbf{\Phi}_{ij}^{ab}$ where two electrons from occupied (occ) orbitals i and j are excited to virtual (vir) orbitals a and b. Møller-Plesset theory provides size extensive methods where the expansion of the many-electron wave function is truncated for finite basis sets.

$$E_{\mathrm{MP2}} = \sum_{i<j}^{\mathrm{occ}} \sum_{a<b}^{\mathrm{vir}} \frac{\left\langle \mathbf{\Phi}_0 \left| \mathbf{H}' \right| \mathbf{\Phi}_{ij}^{ab} \right\rangle \left\langle \mathbf{\Phi}_{ij}^{ab} \left| \mathbf{H}' \right| \mathbf{\Phi}_0 \right\rangle}{E_0 - E_{ij}^{ab}} \tag{2.8}$$

Coupled Cluster. Coupled Cluster (CC) methods are based on the excitation operator $\mathbf{T} = \sum_i^{N_\mathrm{e}} \mathbf{T}_i$. By acting on an HF reference wave function $\mathbf{\Phi}_0$, $\mathbf{T}_i$ generates all ith excited Slater determinants. For instance, $\mathbf{T}_1\mathbf{\Phi}_0 = \sum_i^{\mathrm{occ}} \sum_a^{\mathrm{vir}} t_i^a \mathbf{\Phi}_i^a$ and $\mathbf{T}_2\Phi_0 = \sum_{i<j}^{\mathrm{occ}} \sum_{a<b}^{\mathrm{vir}} t_{ij}^{ab} \Phi_{ij}^{ab}$ with excited determinants $\mathbf{\Phi}_i$ and amplitudes t equivalent to a_i in eq. (2.7). CC methods include all corrections of a formally excited state (singlet (S), doublet (D), triplet (T), …) to infinite order with the size consistent and size extensive exponential ansatz $\mathbf{\Psi}_{\mathrm{CC}} = \mathrm{e}^{\mathbf{T}}\mathbf{\Phi}_0$ and the Schrödinger eq. $\mathbf{H}\mathrm{e}^{\mathbf{T}}\mathbf{\Phi}_0 = E\mathrm{e}^{\mathbf{T}}\mathbf{\Phi}_0$. Size extensivity implies that the method scales properly with the number of particles which can be interacting. The exponential cluster operator $\mathrm{e}^{\mathbf{T}}$ can be written as a sum of generators for the ground state (HF) and excited states (S, D, T, …). Physically, a connected term such as $\mathbf{T}_4$ corresponds to four electrons interacting simultaneously whereas a disconnected term such as $\mathbf{T}_2^2$ corresponds to two non-interacting pairs of interacting electrons. The products with the disconnected terms make CC size extensive.

$$\mathrm{e}^{\mathbf{T}} = \sum_{k=0}^{\infty} \frac{1}{k!} \mathbf{T}^k = 1 + \mathbf{T}_1 + \left(\mathbf{T}_2 + \frac{1}{2}\mathbf{T}_1^2 \right) + \left(\mathbf{T}_3 + \mathbf{T}_2\mathbf{T}_1 + \frac{1}{6}\mathbf{T}_1^3 \right) + \cdots \tag{2.9}$$

Instead of an unmanageable variational approach with non-vanishing terms up to order N_e, the standard procedure is to project the CC Schrödinger eq. onto the reference wave function. Expanding eq. (2.9) and using the fact that $\mathbf{H}$ contains only one- and two-electron operators leads to eq. (2.10).

$$E_{\mathrm{CC}} = \left\langle \mathbf{\Phi}_0 \left| \mathbf{H}e^{\mathbf{T}} \right| \mathbf{\Phi}_0 \right\rangle = E_0 + \sum_{i}^{\mathrm{occ}} \sum_{a}^{\mathrm{vir}} t_i^a \left\langle \mathbf{\Phi}_0 \left| \mathbf{H} \right| \mathbf{\Phi}_i^a \right\rangle + \sum_{i<j}^{\mathrm{occ}} \sum_{a<b}^{\mathrm{vir}} \left(t_{ij}^{ab} + t_i^a t_j^b - t_i^b t_j^a \right) \left\langle \mathbf{\Phi}_0 \left| \mathbf{H} \right| \mathbf{\Phi}_{ij}^{ab} \right\rangle \tag{2.10}$$

The first matrix elements in eq. (2.10) are zero (Brillouin theorem), the second matrix elements are two-electron integrals over MOs for Slater determinants constructed from HF orbitals.

$$E_{\mathrm{CC}} = E_0 + \sum_{i<j}^{\mathrm{occ}} \sum_{a<b}^{\mathrm{vir}} \left(t_{ij}^{ab} + t_i^a t_j^b - t_i^b t_j^a \right) \left(\left\langle \phi_i \phi_j \left| \phi_a \phi_b \right\rangle - \left\langle \phi_i \phi_j \right| \phi_b \phi_a \right\rangle \right) \tag{2.11}$$

Truncation of the excitation operator to $T = T_1 + T_2$ leads to computationally attainable CC with single and double excitations (CCSD). The CCSD(T) method includes perturbative triple excitations approximated using Møller-Plesset fourth-order perturbation theory with amplitudes from CCSD instead of the perturbation coefficients. Coupling between singles and triples is accounted by a term from fifth order perturbation theory. A recently developed linearly scaling approximation to CCSD(T) is achieved using domain based local pair natural orbitals (DLPNO).[90–92] The number of correlated electron pairs in DLPNO is reduced by localization of internal orbitals (pair correlation decreases sharply with distance). Moreover, a significant part of virtual space for each electron pair is spanned by a truncated pair specific natural orbital basis. In consequence, electron correlation domains (virtual orbital subspaces) can be assigned to atomic centers since the pair natural orbital space of an electron pair is local and localized in the same region of space. Due to its usual accuracy of $\pm 5\,\mathrm{kJ\,mol^{-1}}$ for most systems, DLPNO-CCSD(T) is recommended for the calculation of formation enthalpies.[93] According to literature, this method recovers 99.8 % to 99.9 % of the CCSD(T) electron correlation energy.[94]

2.2.2. Density Functional Theory

Hohenberg-Kohn theorems. Density functional theory (DFT) describes a system of electrons through its electron density from which physical properties are approximated in form of functionals.[d] DFT is based on the proof that the ground state electronic energy $E[\rho(r)]$ for an external potential $V_{\text{ext}}(r)$ is determined completely by its electron density $\rho(r)$ and can be variationally obtained analogously to wave function methods. Moreover, there exists a universal functional of the density $F[\rho(r)]$ which is independent of $V_{\text{ext}}(r)$ and minimizes the electronic energy to the correct ground state energy associated with $V_{\text{ext}}(r)$. The exact form of that density functional is not known and approximated in practice.

$$E[\rho(r)] = \int V_{\text{ext}}(r)\rho(r)dr + F[\rho(r)] \tag{2.12}$$

Kohn-Sham ansatz. Early attempts to use DFT suffered from a poor representation of the electron kinetic energy T_{S}. The accuracy was ameliorated significantly by the introduction of an auxiliary set of orbitals ϕ_i representing the electron density $\rho_{\text{approx}} = \sum_i^{N_e} |\phi_i|^2$. Under the assumption of non-interacting electrons, the exact electron kinetic energy $T_{\mathbf{\Phi}} = \sum_i^{N_e} \left\langle \phi_i \left| -\frac{1}{2}\nabla^2 \right| \phi_i \right\rangle$ can be computed from a Slater determinant $\mathbf{\Phi}$. For interacting electrons, the general DFT expression for electronic energy $E_{\text{DFT}} = T_{\mathbf{\Phi}}[\rho] + E_{\text{ne}}[\rho] + J[\rho] + E_{\text{xc}}[\rho]$ consists of the additional potential energy terms E_{ne} for interaction between electrons and nuclei as well as the Coulomb interaction J. Electron interaction is then expressed in the exchange-correlation energy $E_{\text{xc}}[\rho] = (T[\rho] - T_{\mathbf{\Phi}}[\rho]) + (E_{\text{ee}}[\rho] - J[\rho])$, which is defined as the part that remains of the electronic energy after subtraction of the known energy terms from the exact kinetic energy T and Coulomb energy E_{ee}. In this approach, the exact exchange-correlation functional is the essence of the unknown terms – and roughly a factor 10 smaller than the kinetic energy that was not adequately described by ρ before. Similar to the HF and Roothaan-Hall eqs., the Kohn-Sham pseudo-eigenvalue eqs. $\mathbf{h}_{\text{KS}}\phi_i = \varepsilon_i\phi_i$ with canonical Kohn-Sham orbitals ϕ_i and the effective one-electron operator $\mathbf{h}_{\text{KS}}$ can be expanded in a basis set, $\mathbf{h}_{\text{KS}}\mathbf{C} = \mathbf{SC}\boldsymbol{\varepsilon}$, where the energy is minimized self-consistently.

$$\mathbf{h}_{\text{KS}} = \frac{1}{2}\nabla^2 + \mathbf{V}_{\text{eff}} \quad \text{with} \quad \mathbf{V}_{\text{eff}}(\mathbf{r}) = \mathbf{V}_{\text{ne}}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' + \mathbf{V}_{\text{xc}}(\mathbf{r}) \quad \text{with} \quad \mathbf{V}_{\text{xc}}(\mathbf{r}) = \frac{\delta E_{\text{xc}}[\rho]}{\delta\rho(\mathbf{r})} \tag{2.13}$$

Resolution of the identity. The Coulomb energy $J = \frac{1}{2}\int \rho(\boldsymbol{r}_i)\frac{1}{\boldsymbol{r}_{ij}}\rho(\boldsymbol{r}_j)d\tau$ can be evaluated accurately and more efficiently with an approximate representation for the electron density $\rho(\boldsymbol{r}) = \sum_\alpha c_\alpha\alpha(\boldsymbol{r})$ expanded in a nuclear-centered auxiliary basis set α.[95] This resolution of the identity (RI-J) approximation is applied in the electronic structure calculations with local basis sets reported in this work.

Rungs towards more and more accurate electron exchange-correlation energy functionals. The exact unique density functional for exchange correlation, valid for all systems, is unknown but can be approximated at different levels of theory commonly referred to as Jacob's ladder, which is a metaphor for the tendency that accuracy improves with the effort of theory although there is no systematic way for improvement. DFT methods differ in the form of this density functional which either aims to fulfill a number of theoretical premises or is adapted to fit experimental data. In practical calculations, the functional is chosen based on the investigated system and the properties of interest. In the Local Density

[d] A functional $F[f]$, denoted by brackets, is a prescription for producing a number from a function – an example is the electronic energy depending on a wave function or on the electron density. A function $f(\boldsymbol{x})$, denoted by parenthesis, does the same from a set of variables – examples are wave functions and electron density as functions of coordinates.

Approximation (LDA) the functional depends only on the local electron density $E_{\mathrm{xc}}^{\mathrm{LDA}} = E[\rho(\mathbf{r})]$. LDA tends to overestimate molecular binding energies, to underestimate lattice constants of solids and to fail for reaction energies of many reactions. The General Gradient Approximation (GGA) improves the description by including the local gradient of the electron density $E_{\mathrm{xc}}^{\mathrm{GGA}} = E[\rho, \nabla\rho]$. An example is the non-empirical Perdew-Burke-Ernzerhof (PBE) functional used in this work.[96] Generally, both LDA and GGA tend to underestimate activation barriers, because the transition structures at the barrier peaks are predicted to be too stable. Despite these flaws, their results can oftentimes be useful when estimations deviate systematically so that trends are reproduced well. In these currently known and used approximate density functionals, DFT definitions of exchange and correlation energies differ slightly from their counterparts in wave function methods. The exchange energy present in HF theory can be considered as a quantum correction to classical Coulomb repulsion, while the correlation energy describes a dynamical effect where electrons tend to avoid each other more than prescribed by the HF wave function. This reduced probability is called an exchange-correlation hole. The exchange-correlation energy E_{xc} in DFT accounts for the electron kinetic energy difference between the exact interacting system and the non-interacting approximation as well as the corrections for exchange interaction and Coulomb interactions of an electron with itself. These local definitions of exchange and correlation energy miss the (static and time-independent) long-range correlation energy (London dispersion) which can be considered in corrections such as DFT-D2 and DFT-D3 by Grimme et al.[59,60] Furthermore, the electron self-interaction energy does not cancel out like in HF theory leading to non-physical self-repulsion due to its own electron density. Some density functionals treat exchange and correlation separately whereas others follow a combined approach – however, only the combined exchange-correlation hole has a physical meaning. Further rungs on Jacob's ladder are meta-GGA functionals, such as TPSS, which include second (and higher) derivatives of the electron density and eventually hybrid functionals, such as B3-LYP, that mix in the HF exchange functional.

Periodic boundary conditions and plane waves. Extended systems with translational symmetry, like crystals, surfaces and polymers, can be modeled in unit cells with periodic boundary conditions (PBC) in 3D, 2D and 1D, respectively.[e] Translational symmetry can also be expected for the nucleus-electron interaction potentials of these systems. This fact is used in the Bloch theorem which states that the solutions of the single particle Schrödinger eq. for the wave function $\psi_{\mathbf{k}}$ in a periodic potential have the form of a plane wave with the wave vector $\mathbf{k}$ modulated by a periodic function $u_k(\mathbf{r})$ with the same periodicity as the unit cell. Moreover, the translational symmetry of the system must recur in the square of $\psi_{\mathbf{k}}$. Instead of treating an infinite number of orbitals (electrons), only the electrons within the unit cell need to be considered. $\psi_{\mathbf{k}}(\mathbf{r})$ of an electron occupying a state of the quantized wave vector $\mathbf{k}$ in a periodic potential is written as the wave vector confined to the first Brillouin zone. The first Brillouin zone is the equivalent of a Wigner-Seitz unit cell in reciprocal space. Instead of integrating over an infinite number of $\mathbf{k}$ values, the electron density of the slowly varying wave functions is approximated with a finite number of sampling points called k-points – for the systems in this work the Γ-point is sufficient. $u_k(\mathbf{r})$ can be expanded in plane waves which have the periodicity of the lattice vector of the unit cell in reciprocal space $\mathbf{G}$ and the expansion coefficient $c_{i\mathbf{k}}(\mathbf{G})$. It is sufficient to truncate the infinite series and only include the plane waves up to the cut-off energy $E_{\mathrm{cut}} = \frac{\hbar^2}{2m}\mathbf{G}_{\mathrm{cut}}^2$ because the expansion coefficient becomes smaller with increasing $|\mathbf{G}|^2$ thus reaching convergence within the required tolerance upon a certain size of the basis set.

$$\psi_{\mathbf{k}}(\mathbf{r}) = \exp(i\mathbf{k}\mathbf{r})\, u_{\mathbf{k}}(\mathbf{r}) \quad \xrightarrow{\text{expanded to}} \quad \psi_{i\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{i\mathbf{k}}(\mathbf{G}) \exp(i(\mathbf{k}+\mathbf{G})\mathbf{r}) \tag{2.14}$$

[e] A unit cell is characterized by the length and angles of the vectors spanning its physical space containing the atoms.

In analogy to the HF scheme, a matrix eigenvalue problem can be derived for systems with PBCs by considering only the orbitals within the first Brillouin zone, $\mathbf{F}^{\mathbf{k}}\mathbf{C}^{\mathbf{k}} = \mathbf{S}^{\mathbf{k}}\mathbf{C}^{\mathbf{k}}\boldsymbol{\varepsilon}^{\mathbf{k}}$. The solutions are continuous as a function of $\mathbf{k}$ and provide a range of energies called a band. In contrast to nuclear-centered Gaussian-type basis sets, which scale linearly with system size, plane-wave basis sets depend only on the size of the unit cell, not the size of the actual system described within the cell. Plane-wave basis sets become more favorable for large systems. While plane-wave basis sets have primarily been used for systems with translational symmetry, they can also be used for localized systems (molecules) if the system is placed in a sufficiently large unit cell such that it does not interact with its own image in the neighboring cells.

Pseudopotentials. The singularity of the nucleus-electron interaction potential in the proximity of the nuclei makes all-electron calculations with a plane-wave basis rather expensive (many plane waves are needed). Therefore, this type of basis set is used in connection with pseudopotentials for modeling the effect of core electrons and describing nuclear charge. A pseudopotential is also required for describing the potential near the nucleus in hydrogen, even though hydrogen does not have core electrons. The projector augmented wave (PAW) method is a way to solve this electronic structure challenge by introducing projectors on smooth valence functions and auxiliary localized functions. For the PAW method, space is divided into two regions, namely non-overlapping augmentation spheres around the atoms and the interstitial region, where the Kohn-Sham wave functions are expected to be smooth. Inside the augmentation spheres, the rapidly oscillating atomic wave functions are modeled by spherical augmentation functions called partial waves. With Gaussian-type basis sets, effective core potentials (ECPs) are often used for heavier elements to replace a number of core electrons by an effective potential. Core electrons are less important for chemical reactions, but a large number of basis functions is necessary to expand the corresponding orbitals required for a proper description of the valence orbitals. ECPs are particularly advantageous for heavier atoms with relativistic effects.

2.3. Statistical Mechanics

Macroscopic thermodynamic properties represent the average behavior of a large number of particles. Typical molecular calculations are performed on microscopic systems with relatively few particles. Electronic structure calculations usually yield electronic (free) energies for a given set of nuclear coordinates. The theory of statistical mechanics employs the concepts of ensembles and partition functions to connect the microscopic results with measurable macroscopic properties such as thermodynamic state functions.

Ensemble statistics. An ensemble is a collection of imaginary replications of a thermodynamic system distributed over its unique microscopic states. Macroscopic observables can be calculated from the average of the equilibrium distribution of microstates. The classification of thermodynamic systems with respect to their surroundings applies likewise for ensembles. The ensemble of an isolated system with specified composition, volume and energy is called microcanonical ensemble – sometimes the term NVE ensemble is used because all replications have a fixed composition N, volume V and energy E in common. In the canonical ensemble, the isolated systems are replaced by closed systems in thermal contact – due to the exchange of energy the replications in this NVT ensemble have fixed composition, volume and temperature T in common. Further ensemble types exist beyond the scope of this work.

Partition functions. The canonical partition function q for a closed system of particles at a temperature T is the summation of all accessible quantum energy states where the degeneracy factor g_i accounts for equal but distinct energy levels ε_i. The Boltzmann constant k_B relates temperature to energy.

$$q = \sum_i g_i \exp\left(-\frac{\varepsilon_i}{k_\mathrm{B}T}\right) \tag{2.15}$$

Quantum effects can be neglected at elevated temperatures where the energy levels are narrowly spaced. If the state distribution is treated as continuous, the discrete sum is replaced by an integral over the phase space comprised of coordinates $\boldsymbol{r}$ and momenta $\boldsymbol{p}$ where $\mathcal{H}$ is the classical Hamiltonian.

$$q = \int \exp\left(-\frac{\mathcal{H}(\boldsymbol{r},\boldsymbol{p})}{k_\mathrm{B}T}\right) d\boldsymbol{r}d\boldsymbol{p} \tag{2.16}$$

Boltzmann distribution. At any temperature above 0 K the microstates of a closed system are distributed in all accessible energy levels. In thermal equilibrium, the probability p for finding a microstate in a certain energy level is given by the Boltzmann distribution. In opposing tendencies with increasing temperature, this probability decreases exponentially whereas the density of states usually increases. Depending on g_i, the most probable microstate is not necessarily the one with the lowest energy.

$$p(\varepsilon_i) = g_i \exp\left(-\frac{\varepsilon_i}{k_\mathrm{B}T}\right)\frac{1}{q} \tag{2.17}$$

Ideal gas approximation. The partition function $Q = \frac{q^N}{N!}$ for a collection of N non-interacting indistinguishable particles (ideal gas) can be computed from the partition function q for a single particle. Gaseous systems are often approximated as ideal gases because their macroscopic properties can be computed from the molecular partition functions for a single molecule. Without further approximations, systems of condensed matter typically require explicit sampling of the phase space.

Essence of statistical mechanics. The thermodynamic state functions can be defined in terms of the partition function Q. Therefore, macroscopic observables such as pressure, volume V and heat capacity can be computed. Most important for this work are the Helmholtz free energy $A = -k_\mathrm{B}T\ln Q$ and the Gibbs free energy $G = H - TS$ comprising the enthalpy H and the entropy S defined in eqs. (2.18).

$$H = k_\mathrm{B}T^2\left(\frac{\partial \ln Q}{\partial T}\right)_V + k_B T V\left(\frac{\partial \ln Q}{\partial V}\right)_T \quad \text{and} \quad S = k_\mathrm{B}\ln Q + k_\mathrm{B}T\left(\frac{\partial \ln Q}{\partial T}\right)_V \tag{2.18}$$

2.3.1. Molecular Partition Functions

The ground state and the rigid-rotor harmonic-oscillator approximation. The minimum energy of a molecule at 0 K, the molecular ground state, is often taken as a reference for relative energies. Besides electronic energy states, molecular energy levels appear in molecular motion with varying degrees of freedom such as translation of the center of mass, rotation of the molecule or vibration of its atoms. Assuming these degrees of freedom are independent and separable, the molecular partition function q for an isolated molecule can be computed from the product of partition functions for the independent degrees of freedom, see eq. (2.19) – and thermodynamic quantities can be computed from the sum of independent contributions. Within the rigid-rotor harmonic-oscillator (RRHO) approximation, closed

analytical forms exist for these partition functions. The separate contributions to the molecular partition function and their quantum (and classical) expressions are elucidated below.

$$q = q_{\text{elec}} \cdot q_{\text{trans}} \cdot q_{\text{rot}} \cdot q_{\text{vib}} \tag{2.19}$$

Electronic partition function. In all cases in this work, the energy difference between the electronic ground state ε_0 with multiplicity g_0 and excited electronic states is large compared to $k_\text{B}T$. The electronic partition function q_{elec} therefore corresponds to the sum in eq. (2.15) terminated after the first term.

$$q_{\text{elec}} = g_0 \exp\left(-\frac{\varepsilon_0}{k_\text{B}T}\right) \quad (g_0 = 1 \text{ for a singlet, 2 for a doublet, 3 for a triplet, } \ldots) \tag{2.20}$$

Translational partition function. The translational partition function q_{trans} for the translation of a particle with mass m in D degrees of freedom ($D = 3$ in a gas, $D = 2$ for a free translator on a surface) depends on the available characteristic length L (volume of a gas, area of a surface) and the thermal de Broglie wavelength Λ. Classical treatment comes to an equivalent result. The approximation holds as long as Λ – typically in the order of pm – is much smaller than the smallest dimension in space.

$$q_{\text{trans}} = \frac{L^D}{\Lambda^D} \quad \text{with} \quad \Lambda = \frac{h}{\sqrt{2\pi m k_\text{B} T}} \tag{2.21}$$

Rotational partition function. The model of the rigid rotor assumes a fixed molecular geometry independent of the rotational energy state. Rotational symmetry, expressed in the symmetry number σ of equivalent molecular orientations, has to be considered in the rotational partition function q_{rot} which depends on the degenerate moment of inertia I for linear molecules and on the principal moments of inertia (I_1, I_2, I_3) for non-linear molecules. Classical treatment comes to an equivalent result at practical temperatures, where the large values of the moments of inertia characteristic of polyatomic molecules make the quantum of rotational energy $\frac{h^2}{8\pi^2 I_i}$ much smaller than the thermal energy $k_\text{B}T$.

$$q_{\text{rot}} = \begin{cases} \frac{8\pi^2 k_\text{B} T}{h^2} \frac{I}{\sigma} & \text{(linear molecule)} \\ \left(\frac{8\pi^2 k_\text{B} T}{h^2}\right)^{3/2} \frac{(\pi I_1 I_2 I_3)^{1/2}}{\sigma} & \text{(non-linear molecule)} \end{cases} \tag{2.22}$$

Vibrational partition function. The forces in a harmonic oscillator are proportional to the displacement from the equilibrium positions. From spectroscopy experiments it is known that the model of harmonic oscillators for molecular vibrations performs well for low vibrational energy states. The vibrational partition function q_{vib} for molecules with more than two atoms can be approximated by a superposition of harmonic vibrations, the normal vibrations. Classical treatment, where zero-point vibrational energy must not be included, fails at ambient temperature because the characteristic temperature $\frac{h\nu_i}{k_\text{B}}$ is too high. Quantum and classical treatment converge slowly with increasing temperature. The harmonic approximation (HA) for adsorbates can be problematic due to weak interactions where dissociation and internal rotations are important. Anharmonic corrections to the HA are introduced in Chapter 3.

$$q_{\text{vib}} = \prod_i q_{\text{vib},i} \quad \text{with} \quad q_{\text{vib},i} = \begin{cases} \frac{\exp(-h\nu_i/2k_\text{B}T)}{1-\exp(-h\nu_i/k_\text{B}T)} & \text{(quantum treatment)} \\ \frac{k_\text{B}T}{h\nu_i} & \text{(classical treatment)} \end{cases} \tag{2.23}$$

2.3.2. Chemical Equilibrium and Transition State Theory

Chemical reactions. Under conservation of mass and stoichiometry, a chemical reaction is expressed by its reactive species A_i and their stoichiometric coefficients $\nu_i < 0$ for reactants and $\nu_i > 0$ for products. All chemical reactions are equilibrium reactions with forward and reverse reaction.

$$|\nu_1| A_1 + |\nu_2| A_2 + \ldots \rightleftharpoons \ldots + |\nu_i| A_i \tag{2.24}$$

Equilibrium constants. A system of normalized volume $v = \frac{V}{V_0}$ is in thermodynamic equilibrium if its thermodynamic state functions do not change over time – the equilibrium is macroscopically stationary although submicroscopically dynamic where convergent and divergent flows are balanced. For a chemical reaction between the constituents of the system, the chemical equilibrium composition with concentrations $[A_i] = \frac{N_i}{v}$ can be expressed in the equilibrium constant K_c as a function of the partition functions. Note that the molecular partition functions q_i are referenced to the same zero-point of energy while the molecular partition functions $q_{i,0}$ are referenced to the ground states $\varepsilon_{i,0}$ of the molecules, which differ by the energy $\Delta\varepsilon_0$. Furthermore, K_c is often expressed as $K_c = \exp\left(-\frac{\Delta G}{k_B T}\right)$.

$$K_c = \prod_i [A_i]^{\nu_i} = \prod_i (q_i/v)^{\nu_i} = \prod_i \left(q_{i,0}/v\right)^{\nu_i} \exp\left(-\frac{\Delta\varepsilon_0}{k_B T}\right) \quad \text{with} \quad \Delta\varepsilon_0 = \sum_i \nu_i \varepsilon_{i,0} \tag{2.25}$$

Potential energy surface. Atomic coordinates determine the potential energy of a molecular system. Many chemical reactions can be thought of the nuclei moving on a potential energy surface (PES) governed by the theory for the underlying force field. A local minimum on the PES indicates a (meta-)stable state for the structure with the corresponding coordinates.

Transition states. In the semi-classical harmonic transition state theory (TST), reactions from one minimum to another are analyzed via points on the minimum energy path. The highest point along the minimum energy path corresponds to the transition structure, a first order saddle point on the potential energy surface. Defined by the coordinates of the transition structure, the transition state (TS) is a hyperplane that separates the phase spaces of reactants and products – it is a macroscopic ensemble with a Boltzmann energy distribution. Note that one degree of freedom is removed in the TS due to the constraint by the hyperplane. The dynamics along the reaction coordinate, which describes the progress of a reaction as a function of the coordinates, are treated classically, but vibrational and rotational energy states are quantized. It is assumed that molecules acquire energy by collisions and that the internal (vibrational) energy redistribution is significantly faster than the time scale for breaking or forming a bond. In TST the PES is approximated to be locally harmonic (second order Taylor series) around the reactant and around the transition structure.

Link to kinetics. TST can be used to compute the rate constant k_{TST} for a chemical transition from reactants to products. The Eyring eq. (2.26) is obtained assuming that the reactants r are in equilibrium with the TS and that all reactants which pass through the hyperplane perpendicular to the reaction coordinate at the TS go on to the product with no return. The equilibrium constant $K_c^{\ddagger}$ between the reactants and the TS is defined as in eq. (2.25) but the defect degree of freedom in the partition function $q_{TS,0}^{\ddagger}$ for the TS is removed. Treatment of the defect degree of freedom at the TS as a classical vibration with a low force constant or as translation results in $q' = k_B T h^{-1}$ for the decay rate constant for the motion from the activated complex in the hyperplane towards the products. The ground states of

reactants and TS differ by the activation energy $\Delta\varepsilon_0^{\ddagger}$. In this work, the relation is used to compute rate constants from Gibbs free energy barriers $\Delta G^{\ddagger}$ via $k_{\mathrm{TST}} = k_{\mathrm{B}} T h^{-1} \exp(-\Delta G^{\ddagger} k_{\mathrm{B}}^{-1} T^{-1})$.

$$k_{\mathrm{TST}} = q' K_c^{\ddagger} = \frac{k_{\mathrm{B}} T}{h} \frac{\left(q_{\mathrm{TS},0}^{\ddagger}/v\right)}{\prod_r \left(q_{r,0}/v\right)^{|\nu_r|}} \exp\left(-\frac{\Delta\varepsilon_0^{\ddagger}}{k_{\mathrm{B}} T}\right) \tag{2.26}$$

Transition state search. The key to the simulation of reaction kinetics with rate constants from harmonic TST is to find the minima and TS on the PES of the given system because the rate constant is completely determined by the height of the reaction barrier $\Delta G^{\ddagger}$. Local minima are discovered by automated optimization algorithms, for instance by following the gradient of the PES. However, the search of a TS is more challenging. Some TS search algorithms climb along a degree of freedom, whereas others interpolate between reactant and product structures. In this work, the TS are discovered by applying Automated Relaxed Potential Energy Surface Scans (ARPESS).[97] Given a good starting geometry and some chemical intuition or experience, the ARPESS algorithm climbs along a pre-defined holonomic constraint resembling the reaction coordinate to obtain the transition structure.

2.3.3. Molecular Dynamics

The concept. Molecular dynamics (MD) simulation is a computational technique to study the time evolution of a many-body system, where the nuclear motion of the constituent particles obeys the laws of classical mechanics. For an initial sample of the model system, the eqs. of motion are integrated, where the system energy and its gradient with respect to the particle positions are derived from the underlying model for particle interactions. These driving forces are typically computed from classical (spring-like) force fields or *ab initio* methods. After an equilibration period during which the thermodynamic equilibrium is reached, the simulation enters the production period, where the observable of interest can be measured (sampling). In the simulation, the system explores its phase space, which is the full collection of accessible system states given by positions and momenta. The continuous curve in phase space as a function of time, the trajectory, is used to analyze observables.

Andersen thermostat. Usually, the particle positions $\boldsymbol{r}$ and momenta $\boldsymbol{p}$ at a given point in time are computed by numerical integration of the eqs. of motion using a finite time step Δt at which forces are evaluated and positions are updated. The choice of Δt is a compromise between accuracy of the integration (numerical stability) and computational cost for repeated evaluations of the forces. Depending on the integration algorithm, a good choice for Δt should be at least one order of magnitude shorter than the fastest motion, often caused by vibrations of the lightest particles. Reliable algorithms, like the leapfrog and velocity Verlet variants, are stable, time reversible and conserve the total energy in the NVE ensemble. The NVT ensemble can be obtained by employing numerical thermostats, such as the Andersen thermostat. Using this thermostat, constant temperature is imposed by stochastic collisions with a virtual heat bath. With a (small) probability the collisions act occasionally on velocity components of randomly selected particles. Upon a collision, the new velocity is drawn from the Maxwell-Boltzmann distribution at the corresponding temperature. The system is then integrated numerically at constant energy according to the Newtonian laws of motion. The collision probability – the coupling strength to the heat bath – is defined as the average number of collisions per atom and time step. The algorithm generates a canonical distribution.[98] However, due to the random decorrelation of velocities, the dynamics are unphysical and cannot represent dynamical properties like diffusion or viscosity. Another disadvantage is that the collisions are stochastic in nature, so repeating the simulation will ordinarily not give exactly the same trajectory unless the random seed is fixed.

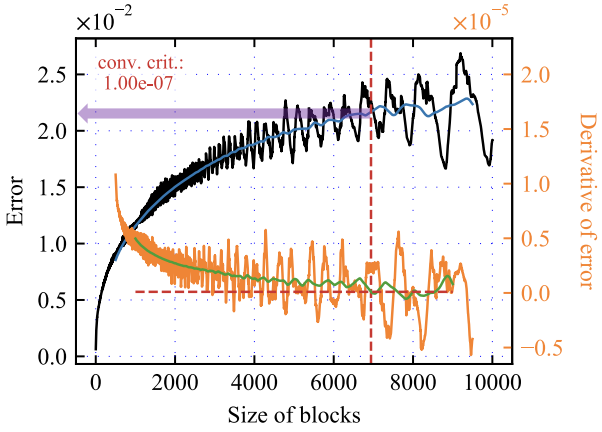


Figure 2.1.: Illustration of the blocking method for the error estimation of correlated data. The error of the correlated data (black) converges for increasing block size, as blocks become more and more uncorrelated. Automated error analysis can be done by monitoring the smoothened derivative (green) of the smoothened block error (blue). In the shown example, the error converges to around $2.2 \cdot 10^{-2}$ in units of the quantity of interest.

The ensemble average equals the time average. An ergodic system eventually passes through all points of phase space constituting the real ensemble with the correct statistical weights. For ergodic systems the ensemble average $\langle O \rangle$ of a position or momentum dependent system property $O(\boldsymbol{r}(t), \boldsymbol{p}(t))$ can be obtained from a sufficiently long MD simulation, discretized in practice with a sufficiently large number of time steps M. The ergodic hypothesis is thus fundamental for MD simulation.

$$\langle O \rangle = \lim_{t\to\infty} \frac{1}{t} \int_0^t O(\boldsymbol{r}(t), \boldsymbol{p}(t)) dt = \lim_{M\to\infty} \frac{1}{M} \sum_{n=1}^{M} O_n \tag{2.27}$$

Observables. Observables in MD simulations must be expressed as functions of positions or momenta. Two examples appearing in this work are the probability distribution function $P(\xi') = \langle \delta(\xi(\boldsymbol{r}, \boldsymbol{p}) - \xi') \rangle$ of a position dependent geometric variable ξ and the radial distribution function (RDF). The RDF is an expression for the average relative locations of the particles. It defines the probability of finding a particle j in a distance r_{ij} from a particle i by describing how the normalized particle density ρ_j of a particle j varies as a function of r_{ij}. The RDF g_{ij} is computed from the number dN_j of particles j inside the spherical shell with thickness dr and volume $4\pi r^2 dr$ as a function of r_{ij}. Deviations from unity reflect correlations between particles due to particle interactions, thus $g_{ij} = 1$ for ideal gases.

$$g_{ij}(r_{ij}) = \frac{\rho_j(r_{ij})}{\bar{\rho}_j} = \lim_{dr\to 0} \left\langle \frac{dN_j(r_{ij})}{4\pi r_{ij}^2 dr} \right\rangle_M \frac{1}{\bar{\rho}_j} \tag{2.28}$$

Error estimation by blocking analysis. MD trajectories are time correlated data, and so are the instantaneous system properties $O(t)$. The statistical error of $\langle O \rangle$ can be estimated using the blocking method of Flyvbjerg and Petersen,[99] where the time correlated data is repeatedly replaced by lesser correlated data, namely, the averages of blocks of a number of consecutive samples of $O(t)$. With increasing block size, the correlation becomes lesser, as can be seen in Figure 2.1. A number of N uncorrelated block averages with the standard deviation σ are obtained once the standard error ε as a function of block size is converged to a plateau. The procedure is also called binning analysis.[100]

$$\varepsilon = \frac{\sigma}{\sqrt{N}} \quad \text{with} \quad \sigma = \sqrt{\frac{1}{N-1} \sum_{i}^{N} (O_i - \langle O \rangle)^2} \tag{2.29}$$

Free energy methods. Free energy differences between a reference and a target system can be determined upon exploration of the configurational space of the reference system such that relevant, low-energy states of the target system are adequately sampled. Conventional MD simulation, however, by means of Boltzmann sampling is typically not sufficient. A usual strategy to escape Boltzmann sampling is multistage sampling which relies on the construction of a chain of configurational energies that bridge the reference and target states whenever their low-energy regions overlap poorly. In this stratification approach the total free energy difference is split into a sum of free energy differences between intermediate states that overlap considerably better than the initial and final states. Another common enhanced sampling strategy is called importance sampling, where certain important regions of low visitation probability are sampled frequently in a biased distribution with known weights. The results of the simulation are then reweighted to obtain the unbiased estimator. The arsenal of computational methods for free energy calculations ranges from perturbative methods and methods based on probability distributions to thermodynamic integration (TI) and non-equilibrium methods. Notable among the popular methods are Umbrella sampling together with the weighted histogram analysis method, TI using (un-)constrained dynamics and Bennett's acceptance ratio (BAR). The TI technique relies on calculating and subsequently integrating the derivatives of the free energy over a transformation path defined by an order parameter. Typically, the order parameter is either a function of atomic coordinates or a parameter in the Hamiltonian.[101] Chapter 3 presents a newly developed application of TI to compute anharmonic contributions to the free energy of a harmonic reference system. The introduction of the TI method in detail is continued in Section 3.1.2.

2.4. Kinetics and Surface Kinetics

Reaction kinetics. The dynamics of reaction processes over time are described and analyzed by reaction mechanisms and kinetic models. A reaction mechanism is a network of reaction steps representing chemical conversions from reactants to products. Elementary reactions correspond to individual reaction steps on the molecular scale, while global reactions summarize over multiple steps and have no such direct physical meaning. Separation and neglect of transport effects (diffusion, convection) leads to a microkinetic model. The microkinetic models in this work are composed of elementary reactions.

Reaction rate. For a chemical reaction with reactive species A_i and their stoichiometric coefficients ν_i as defined in eq. (2.24), the reaction rate r at constant volume is defined by the stoichiometry-normalized change in concentration of any reactive species over time, see eq (2.30).

$$r = \frac{1}{\nu_i} \frac{d[A_i]}{dt} \tag{2.30}$$

Reaction order. Empirical reaction rate laws are often expressed by a rate constant k times the concentrations $[A_i]$ to the power of the reaction order m_i. The overall reaction order is $m = \sum_i m_i$. Elementary reactions follow simple rate laws where the integral reaction order follows from the reaction mechanism. In unimolecular reactions ($A_1 \longrightarrow A_2$) the reaction orders are $m_1 = 1$ and $m_2 = 0$. In bimolecular reactions ($A_1 + A_2 \longrightarrow A_3$) the reaction orders are $m_1 = 1$, $m_2 = 1$ and $m_3 = 0$. Transition state theory (see Section 2.3.2) can be used to compute rate constants from molecular properties.

$$r = k \cdot \prod_i [A_i]^{m_i} \tag{2.31}$$

Concentrations over time. The concentration changes of all reactive species can be described by a system of coupled (non-linear) ordinary differential eqs. of first order. Integration from initial concentrations $[A_i](t = 0)$ yields the evolution of the reaction processes over time $[A_i] = [A_i](t)\ \forall\ i$. Numerical integration over small time steps can be used when analytical expressions are missing.

Sorption processes. Adsorption and desorption processes at the interface of a solid surface are typically much faster than surface reactions between adsorbates. Associative adsorption leaves the adsorbate intact, while the adsorbate breaks during dissociative adsorption. The surface reactions in this work are modeled with the mean-field approximation and the Langmuir model for adsorption.

Coverage. Adsorption of a molecular species A can formally be considered as a reaction with an available adsorption site denoted as $*$. With a maximum of one adsorbate per site, the ratio between the occupied surface area covered by adsorbates $[A^*]$ to the generally available surface area $[*]_0$ is described by the surface coverage θ_A. The remaining fraction of available surface area is expressed as $\theta_* = 1 - \theta_A$. Coverage decreases with increasing temperature due to the entropic penalty.

$$\theta_A = \frac{[A^*]}{[*]_0} \tag{2.32}$$

Mean-field approximation. The mean-field approximation for surface kinetics omits spatial correlations like diffusion limitations or interactions between adsorbates. A random distribution of adsorbates over the adsorption sites is assumed. In consequence, coverages can be interpreted as the equivalent to concentrations – and eq. (2.31) applies analogously. In other words, coverage in zeolites with no interaction between adjacent adsorption sites can be understood as a probability for their occupation.

Langmuir model for adsorption. The Langmuir isotherm relates coverage of a particular species to its partial pressure above the surface. This adsorption model by Langmuir assumes coverage-independent equivalence of a finite number of adsorption sites, no interactions between adsorbates and $\theta_A \leq 1$. In the example of associative adsorption of a species A, the rate of adsorption $r_a = k_a p_A \theta_*$ depends on the partial pressure p_A of species A, while the rate of desorption $r_d = k_d \theta_A$ depends only on its coverage.

2.4.1. Ideal Batch Reactor Model

The ideal batch reactor is a model for a discontinuously operated well-stirred tank reactor with no feed and outlet over time, the residence time of the initial input mass thus equals the reaction time. Local gradients of temperature and concentration are absent due to instantaneous mixing. Reactor shape and wall effects are neglected. For isothermal reaction control, instantaneous heat exchange with the environment is assumed. Reaction kinetics for this reactor obey the microkinetic rate laws. The gas phase reactions in this work are simulated isothermal for the ideal batch reactor at constant volume.

2.4.2. Ideal Plug Flow Reactor Model

A continuously operated tube reactor of any cross-section shape can be simulated with the ideal plug flow reactor model for a stationary state with constant feed and outlet. No back-mixing is assumed for the non-dispersive one-dimensional flow. Axial diffusion of any quantity is assumed negligible

relative to convection. Wall effects are neglected. The model assumes no gradients and no variation in transverse direction. Therefore, the approximations in an infinitesimally narrow segment moving along the reactor are comparable to those of the ideal batch reactor model, except that the catalyst bed is fixed in place. For isothermal reaction control, the system of differential algebraic eqs. describing the plug flow reactor consists of the continuity eq. (2.33), the species conservation eq. (2.34) and the state eq. (2.35). Required parameters are the area A_c of the cross-section of the channel, the density ρ, the velocity u, the reactor length z, the surface area A_s per unit length, the number K_g of gas phase species k, the molar production rate $\dot{s}_k$ of species k by surface reactions, the molar mass M_k of species k, the mass fraction Y_k of species k, the molar production rate $\dot{\omega}_k$ of species k by gas phase reactions, the pressure p, the average molar weight $\overline{M}$, the ideal gas constant R and the temperature T.[102]

$$A_c \frac{d(\rho u)}{dz} = A_s \sum_{k=1}^{K_g} \dot{s}_k M_k \tag{2.33}$$

$$A_c \frac{d(\rho u Y_k)}{dz} = M_k (A_s \dot{s}_k + A_c \dot{\omega}_k) \tag{2.34}$$

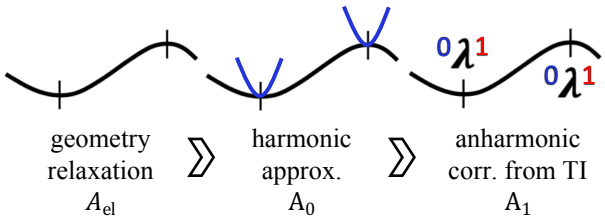
$$p\overline{M} = \rho RT \tag{2.35}$$

Part II.

Contents

3. Anharmonic Corrections from Molecular Dynamics using Thermodynamic Integration

Scope. This chapter presents thermodynamic λ-path integration (λ-TI) as a tool to compute anharmonic corrections to the commonly used harmonic approximation. Section 3.1 explains the concept and how λ-TI is combined with curvilinear internal coordinates in systems with periodic boundary conditions. The λ-TI method is then applied to study adsorption inside zeolite H-SSZ-13 in Section 3.2 and to study reaction barriers in Section 3.3. Parts of this chapter have been published and are adapted and reprinted with permission from [17, 103]. Copyright 2021, 2023 American Chemical Society. The work was compiled with the co-authors Philipp N. Plessow and Felix Studt in advisor roles, and Tomáš Bučko, who implemented the λ-TI method in the VASP software package to make it available for the broader scientific community. Our λ-TI method using molecular dynamics (MD) simulation driven by density functional theory (DFT) is independent of the reaction path and can thus be applied to study arbitrary states in a free energy landscape. The common workflow in which energy contributions are added to the total energy of stationary points on the potential energy surface is thus extended by a third step as shown in Scheme 3.1.



Scheme 3.1.: Common workflow for the computation of free energies in the harmonic approximation extended by the calculation of anharmonic contributions through λ-TI as a third step. The Helmholtz free energies are denoted as A_{el}, A_0 and A_1, respectively. Reprinted with permission from [17]. Copyright 2021 American Chemical Society.

3.1. Method Development and Theory

3.1.1. Overview

MD simulations, typically based on DFT, in combination with enhanced sampling techniques can be used to determine changes in free energy beyond the harmonic (oscillator) approximation. Inherently, established methods like Blue Moon[104,105] or Umbrella[106–108] sampling require the integration over a numerical descriptor for the reaction path called collective variable (CV) or reaction coordinate ξ. The major attention of these tools in heterogeneous catalysis[109–115] has been paid to surface reaction steps rather than adsorption processes. This unilateral focus is unsurprising as the MD free energy calculations are very revealing for chemical transformations but not quite suitable for desorption during which adsorbate and adsorbent must separate ideally infinitely along the CV. Nevertheless, Li et al. demonstrated that the integration of the potential of mean force from DFT-based MD can be used to

estimate adsorption free energy.[77] Obviously, the challenge is to sample anharmonic contributions to the free energy independent of a CV.

In contrast to aforementioned MD techniques, thermodynamic integration (TI)[83,101,116] can be used to calculate free energies relative to a reference system over transformation paths other than the reaction path.[117–122] Temperature and Hamiltonian-based λ-paths have been exploited to compute the free energy of classical crystals[123,124] and to study phase transitions.[125–130] Very recently, Jinnouchi et al. proposed an effective combination of machine learning with TI using the ideal gas reference allowing for accurate ab initio calculations of the chemical potential of LiF dissolved in liquid water.[131] TI can be used to obtain anharmonic contributions with respect to a harmonic reference system. This idea has been put into practice for crystals by performing TI from Debye models to anharmonically interacting systems.[132,133] Its realization for adsorption processes and catalysis in general is limited by several obstacles. Most importantly, conventional TI techniques use Cartesian coordinates which are inherently unsuitable for systems with rotational and translational symmetries. An illustrative presentation of this problem is given in Section 3.1.3. While this difficulty can be overcome in simulations of simple systems by fixing the overall translations and rotations, the solution for the more general cases with internal rotational degrees of freedom (such as, for instance, a weakly interacting molecule adsorbed on a substrate) is less straightforward. Furthermore, Cartesian coordinates fail to represent molecular vibrational motions of large amplitudes[134–136] potentially leading to unnecessarily large differences in energies between harmonic reference and anharmonically interacting system in practical simulations. This latter effect may then lead to a significant efficiency reduction of TI calculations even for relatively simple molecular systems.

A solution to some of the limitations of λ-TI by making use of rotationally and translationally invariant curvilinear coordinates is presented below. This new variant of a λ-TI formalism allows to compute the anharmonic contribution to the free energy for any system that can be described with the harmonic approximation – independent of any reaction path. Moreover, the approach allows for the theoretically exact calculation of a classical anharmonic correction to the harmonic – quantum or classical – approximation. In particular, all anharmonic quantum effects are neglected while the quantum effects described by the quantum harmonic oscillator model can be, if appropriate, directly taken into account. This is especially important, for instance, when adsorption at a wide range of temperatures is considered. Quantum corrections will be important at low temperatures where anharmonicity is small (and thus the error for harmonic treatment is small as well) but with increasing temperature the system will gradually shift to the classical regime for which our treatment is exact. Note that a canonical (NVT) ensemble is assumed throughout this work although a generalization to other ensembles is also possible.

3.1.2. Thermodynamic Integration with a Harmonic Reference System

In the TI technique,[83,101,116] the free energy of a system 1 is expressed in terms of a reference system 0:

$$A_1 = A_0 + \Delta A_{0\to 1} \tag{3.1}$$

where $\Delta A_{0\to 1}$ designates the Helmholtz free energy difference between 1 and 0. Depending on the properties of the investigated system, a variety of reference systems is used in the literature.[124,132,133,137–140] Motivated by the fact that the harmonic approximation is commonly used for calculations of adsorption thermodynamics in solid sorbents[63,74,134,135,141–145] and also in the field of computational zeolite catalysis,[55–57,80,112,113,146] we focus on the harmonic reference system throughout this work. Classical and quantum harmonic approximation deviate significantly at low temperatures where the free energy of matter is dominated by quantum effects. For relative free energies (adsorption, reaction barriers),

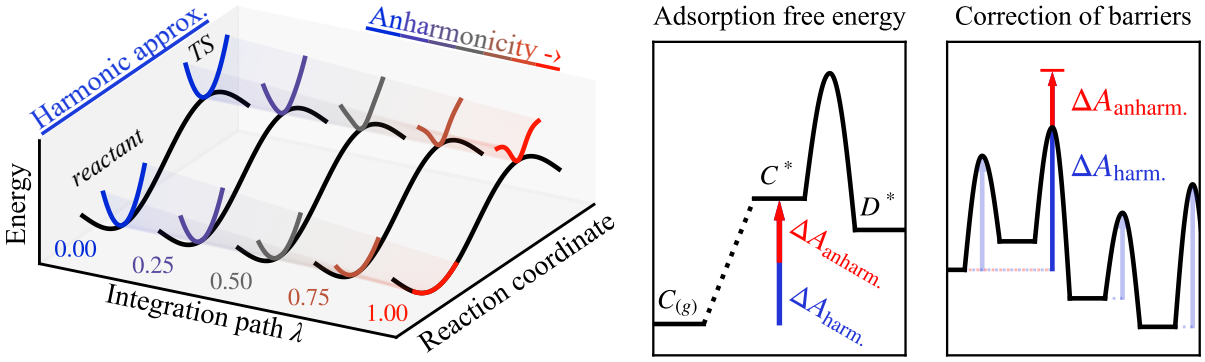


Figure 3.1.: A progressive parameter λ couples the harmonic reference system with the anharmonic DFT-based system in our λ-TI method. Anharmonic corrections to the harmonic approximation can be computed for adsorption free energies and for (selected) free energy barriers. Adapted and reprinted with permission from [17, 103]. Copyright 2021, 2023 American Chemical Society.

however, this difference between theories converges rapidly to zero with increasing temperature. For the outlined anharmonic correction, we are free to choose between the classical and quantum harmonic reference system (and their difference is negligible). Since the anharmonic correction by DFT-based MD is always classical we decided to use the classical harmonic reference system, for which the quasi-classical free energy expression[147] can be written as

$$A_{0,\boldsymbol{x}} = A_{\mathrm{el}}(\boldsymbol{x}_0) - k_{\mathrm{B}}T \sum_{i=1}^{N_{\mathrm{vib}}} \ln \frac{k_{\mathrm{B}}T}{\hbar\omega_i} \tag{3.2}$$

with the electronic free energy $A_{\mathrm{el}}(\boldsymbol{x}_0)$ for the configuration corresponding to the potential energy minimum with the atomic position vector $\boldsymbol{x}_0$, the number of vibrational degrees of freedom N_{vib}, and the angular frequency ω_i of vibrational mode i. Note that a harmonic dependence on Cartesian coordinates $\boldsymbol{x}$ is assumed for the harmonic potential energy $V_{0,\boldsymbol{x}}$ in the derivation of this eq., that is

$$V_{0,\boldsymbol{x}}(\boldsymbol{x}) = V_{0,\boldsymbol{x}}(\boldsymbol{x}_0) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_0)^T \underline{\boldsymbol{H}}^{\boldsymbol{x}} (\boldsymbol{x} - \boldsymbol{x}_0) \tag{3.3}$$

where $\underline{\boldsymbol{H}}^{\boldsymbol{x}}_{i,j} = \frac{\partial^2 V_{0,\boldsymbol{x}}(\boldsymbol{x})}{\partial x_i \partial x_j}\big|_{\boldsymbol{x}=\boldsymbol{x}_0}$ is the Hessian matrix evaluated for the structure $\boldsymbol{x}_0$ belonging to the potential energy minimum. By definition,[87] $A_{\mathrm{el}}(\boldsymbol{x}_0) = -k_{\mathrm{B}}T \ln\left[g_0 \exp\left(-V_{0,\boldsymbol{x}}\left(\boldsymbol{x}_0\right)/k_{\mathrm{B}}T\right)\right]$, where g_0 is the electronic multiplicity of the ground state. Hence, the particularly simple relation $A_{\mathrm{el}}(\boldsymbol{x}_0) = V_{0,\boldsymbol{x}}(\boldsymbol{x}_0)$ holds when the electronic state of the system of interest is a singlet, which is the case for all systems discussed in this work. The correction term $\Delta A_{0\to1} \equiv \Delta A_{0,\boldsymbol{x}\to1}$ can be computed using λ-TI[83,101,116] according to eq. (3.4), where λ is the coupling strength between the true physical system 1 whose potential energy V_1 is computed by DFT and the harmonic reference system 0 with its potential energy defined in eq. (3.3), and $\langle \ldots \rangle_\lambda$ represents the NVT ensemble average of the system driven by the classical Hamiltonian $\mathcal{H}_\lambda = \lambda\mathcal{H}_1 + (1-\lambda)\mathcal{H}_{0,\boldsymbol{x}}$.

$$\Delta A_{0,\boldsymbol{x}\to1} = \int_0^1 d\lambda \langle \mathcal{H}_1 - \mathcal{H}_{0,\boldsymbol{x}} \rangle_\lambda \tag{3.4}$$

Since the Hamiltonians $\mathcal{H}_{0,\boldsymbol{x}}$ and $\mathcal{H}_1$ differ only in their respective potential energy contributions $V_{0,\boldsymbol{x}}$ and V_1, eq. (3.4) can be cast into the form of eq. (3.5). The procedure is illustrated in Figure 3.1.

$$\Delta A_{0,\boldsymbol{x}\to1} = \int_0^1 d\lambda \langle V_1 - V_{0,\boldsymbol{x}} \rangle_\lambda \tag{3.5}$$

In practice, the ensemble averages are computed using MD simulations, like in this work, or Monte Carlo simulations. Typically, it is observed that the phase space volume effectively spanned by the system at the given conditions increases as λ is increased. This increase is a consequence of the transformation from harmonic to anharmonic vibrations or even to non-vibrational degrees of freedom such as internal rotations or hindered translations (for instance, when adsorbates start to diffuse through the substrate). For this reason, the integrand of eq. (3.4) often tends to increase in absolute value as λ reaches its upper limit, which must be taken into account when choosing the mesh of integration points for the evaluation of $\Delta A_{0,\boldsymbol{x}\to 1}$.

3.1.2.1. TI in Internal Coordinates

One practical problem for the use of λ-TI in chemistry is related to the fact that the commonly used Cartesian coordinates $\boldsymbol{x}$ are not insensitive to the overall rotations and translations (or any other symmetry operation) that must leave the total energy invariant. Therefore, Cartesian coordinates are intrinsically unsuitable for the use in λ-TI simulations of gas phase molecules and weakly bound adsorption complexes.

Here we propose to overcome this problem in replacing the Cartesian coordinates by rotationally and translationally invariant internal coordinates $\boldsymbol{q} = \boldsymbol{q}(\boldsymbol{x})$, such as bond lengths, angles, torsions, and their more complex combinations, for instance, weighted sums of coordination number functions as in eq. (3.29). Since the force field which is harmonic in $\boldsymbol{x}$ is not necessarily equivalently harmonic in $\boldsymbol{q}$ (and *vice versa*), we express eq. (3.1) in the form of eq. (3.6) where $\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}}$ corresponds to the transformation from the system harmonic in $\boldsymbol{x}$ to the system harmonic in $\boldsymbol{q}$, while $\Delta A_{0,\boldsymbol{q}\to 1}$ is the contribution by the transformation from the system harmonic in $\boldsymbol{q}$ to the anharmonically interacting system 1.

$$A_1 = A_{0,\boldsymbol{x}} + \Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}} + \Delta A_{0,\boldsymbol{q}\to 1} \tag{3.6}$$

In consequence of the fact that eq. (3.6) is a sum, the anharmonic correction by λ-TI extends the common workflow in which energy contributions are added to the total energy of stationary points on the potential energy surface, as already outlined in Scheme 3.1. The relations between the harmonic approximation, the harmonic reference systems and the anharmonic system are illustrated in Scheme 3.2.

For the special case of phase volume conserving coordinates (such as interatomic distances) the term $\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}}$ equals zero. In general, however, this contribution does not vanish (albeit it often is very small) and its value should be taken into account. The term $\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}}$ corresponds to work due to a force field to force field transformation and is computationally inexpensive. For this reason, we determine this term numerically using eq. (3.4) combined with the Hamiltonian $\mathcal{H}_\lambda$ in eq. (3.7).

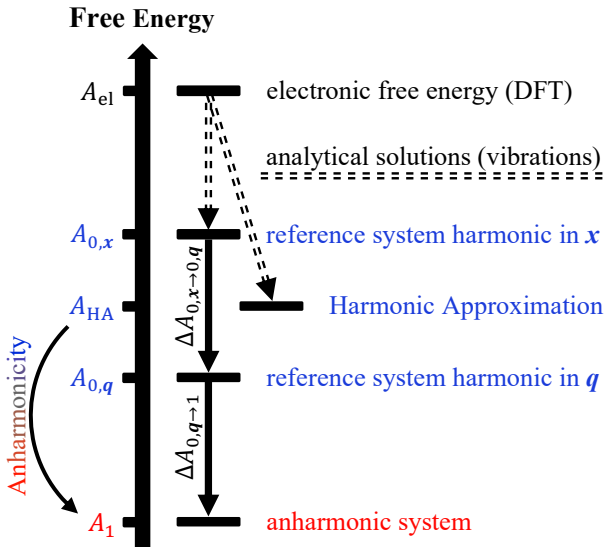
$$\mathcal{H}_\lambda = \lambda\mathcal{H}_{0,\boldsymbol{q}} + (1-\lambda)\mathcal{H}_{0,\boldsymbol{x}} \tag{3.7}$$

Since the kinetic energy is independent of the choice of coordinates, the Hamiltonians $\mathcal{H}_{0,\boldsymbol{x}}$ and $\mathcal{H}_{0,\boldsymbol{q}}$ for the systems harmonic in $\boldsymbol{x}$ and $\boldsymbol{q}$, respectively, differ only in their respective potential energy contributions $V_{0,\boldsymbol{x}}(\boldsymbol{x})$ and $V_{0,\boldsymbol{q}}(\boldsymbol{q})$ given by eqs. (3.3) and (3.8).

$$V_{0,\boldsymbol{q}}(\boldsymbol{q}) = V_{0,\boldsymbol{q}}(\boldsymbol{q}_0) + \frac{1}{2}(\boldsymbol{q}-\boldsymbol{q}_0)^T\underline{\boldsymbol{H}}^{\boldsymbol{q}}(\boldsymbol{q}-\boldsymbol{q}_0) \tag{3.8}$$

In eq. (3.8) $\boldsymbol{q}_0$ are the internal coordinates evaluated for the potential energy minimum, and the Hessian matrix $\underline{\boldsymbol{H}}^{\boldsymbol{q}}$ is related[148–152] to $\underline{\boldsymbol{H}}^{\boldsymbol{x}}$ via eqs. (3.9) with $\underline{\boldsymbol{A}}$ being the Moore-Penrose pseudoinverse of the Wilson B-matrix[136,153] given in eq. (3.10).

$$\underline{\boldsymbol{H}}^{\boldsymbol{q}} = \underline{\boldsymbol{A}}^T_{\mathbf{x}_0}\underline{\boldsymbol{H}}^{\boldsymbol{x}}\underline{\boldsymbol{A}}_{\mathbf{x}_0} \quad \text{and} \quad \underline{\boldsymbol{H}}^{\boldsymbol{x}} = \underline{\boldsymbol{B}}^T_{\mathbf{x}_0}\underline{\boldsymbol{H}}^{\boldsymbol{q}}\underline{\boldsymbol{B}}_{\mathbf{x}_0} \tag{3.9}$$



Scheme 3.2.: Relations between the electronic free energy, the harmonic approximation, the harmonic reference systems and the anharmonic system. With the assumption of harmonic oscillators in Cartesian coordinates there are analytical solutions for the free energy contributions to $A_{0,\boldsymbol{x}}$ and A_{HA}. Reprinted with permission from [103]. Copyright 2023 American Chemical Society.

$$\underline{\boldsymbol{B}}_{i,j} = \frac{\partial q_i}{\partial x_j} \tag{3.10}$$

It should be emphasized that all terms in eqs. (3.9) are evaluated for the stationary point $\mathbf{x} = \mathbf{x}_0$ whose structure can be described by the internal coordinates $\mathbf{q} = \mathbf{q}_0$. The forces needed in the evaluation of the eqs. of motion for the atoms in a λ-TI calculation of the term $\Delta A_{0,\boldsymbol{x}\to0,\boldsymbol{q}}$ are defined in eq. (3.11).

$$-\frac{\partial \mathcal{H}_\lambda}{\partial \boldsymbol{x}} = -\lambda \frac{\partial \mathcal{H}_{0,\boldsymbol{q}}}{\partial \boldsymbol{x}} - (1-\lambda)\frac{\partial \mathcal{H}_{0,\boldsymbol{x}}}{\partial \boldsymbol{x}} \tag{3.11}$$

With the transformation relations[148–152] in eqs. (3.12) for the forces derived from any arbitrary potential V eq. (3.11) can be rearranged into the form of eq. (3.13) with matrix $\underline{\boldsymbol{B}}_{\boldsymbol{x}}$ and vector $\boldsymbol{q}$ being evaluated for the current position vector $\boldsymbol{x}$.

$$-\frac{\partial V}{\partial \boldsymbol{x}} = -\underline{\boldsymbol{B}}_{\mathrm{x}}^T \left(\frac{\partial V}{\partial \boldsymbol{q}}\right) \quad \text{and} \quad -\frac{\partial V}{\partial \boldsymbol{q}} = -\underline{\boldsymbol{A}}_{\mathrm{x}} \left(\frac{\partial V}{\partial \boldsymbol{x}}\right) \tag{3.12}$$

$$-\frac{\partial \mathcal{H}_\lambda}{\partial \boldsymbol{x}} = -\lambda \underline{\boldsymbol{B}}_{\mathrm{x}}^T \underline{\boldsymbol{H}}^{\boldsymbol{q}} (\boldsymbol{q} - \boldsymbol{q}_0) - (1-\lambda)\underline{\boldsymbol{B}}_{\tilde{\boldsymbol{x}}_0}^T \underline{\boldsymbol{H}}^{\boldsymbol{q}} \underline{\boldsymbol{B}}_{\tilde{\boldsymbol{x}}_0} (\boldsymbol{x} - \tilde{\boldsymbol{x}}_0) \tag{3.13}$$

Since the molecular systems might rotate during the MD simulation, it is important to match the position of the minimum energy structure with the orientation of the current position vector $\boldsymbol{x}$, otherwise incorrect forces corresponding to the harmonic model would be generated – see Section 3.1.3 for an illustration of this problem. In fact, the term $\underline{\boldsymbol{B}}_{\tilde{\boldsymbol{x}}_0}^T \underline{\boldsymbol{H}}^{\boldsymbol{q}} \underline{\boldsymbol{B}}_{\tilde{\boldsymbol{x}}_0}$ in eq. (3.9) corresponds to the Hessian matrix of the unperturbed system expressed in rotated Cartesian coordinates. Such a rotated minimum energy structure, which we label $\tilde{\boldsymbol{x}}_0$, is determined iteratively by discretization of the linear relation $d\boldsymbol{x} = \underline{\boldsymbol{A}}\, d\boldsymbol{q}$. In this procedure, commonly used in the context of geometry optimization in internal coordinates,[148–152] we compute the rotated coordinates $\tilde{\boldsymbol{x}}_0$ self-consistently using eq. (3.14) with repeated evaluations of $\tilde{\boldsymbol{q}}^{i+1}$ corresponding to the trial Cartesian vector $\tilde{\boldsymbol{x}}_0^{i+1}$ as well as $\underline{\boldsymbol{A}}_{\tilde{\boldsymbol{x}}_0^i}$ corresponding to $\tilde{\boldsymbol{x}}_0^i$ and $\tilde{\boldsymbol{q}}^i$.

$$\tilde{\boldsymbol{x}}_0^{i+1} = \tilde{\boldsymbol{x}}_0^i + \underline{\boldsymbol{A}}_{\tilde{\boldsymbol{x}}_0^i} (\boldsymbol{q}_0 - \tilde{\boldsymbol{q}}^i) \tag{3.14}$$

Likewise, the term $\Delta A_{0,\boldsymbol{q}\to 1}$ is computed using λ-TI according to eq. (3.4) with the Hamiltonian in eq. (3.15). The forces generated by the Hamiltonian in eq. (3.15) are used in the evaluation of the eqs. of motion for the atoms in the MD simulations.

$$\mathcal{H}_\lambda = \lambda\mathcal{H}_1 + (1-\lambda)\mathcal{H}_{0,\boldsymbol{q}} \tag{3.15}$$

Employing the transformation relation from eq. (3.12), these forces can be expressed according to eq. (3.16) where $V_1(\boldsymbol{x})$ is the electronic energy computed using DFT.

$$-\frac{\partial\mathcal{H}_\lambda}{\partial\boldsymbol{x}} = -\lambda\frac{\partial V_1(\boldsymbol{x})}{\partial\boldsymbol{x}} - (1-\lambda)\underline{\boldsymbol{B}}_{\boldsymbol{x}}^T\underline{\boldsymbol{H}}^{\boldsymbol{q}}(\boldsymbol{q}-\boldsymbol{q}_0) \tag{3.16}$$

3.1.2.2. Harmonic Reference Systems with Holonomic Constraints

Naturally, transition states (TS) are instable along one degree of freedom which can often be described by an internal coordinate or by a linear combination of internal coordinates. A TS can thus be preserved using holonomic constraints on the corresponding internal coordinate. When choosing a harmonic reference system for a TS, this constrained degree of freedom must be projected out from the Hessian matrix to remove its free energy contribution entirely. In other words, the projected Hessian matrix should not generate forces for displacements of the constrained coordinate (coupling elements of the Hessian are dealt with as well). The necessary projection of the Hessian matrix blends in perfectly with our λ-TI approach where we transform the Hessian matrix using the Wilson B-matrix as projector. A proper modification of this projector can thus include any holonomic constraints.

Procedures for the preparation of a projector with constraints have been reported in the literature, mostly in the context of delocalized internal coordinates. Farkas and Schlegel describe the treatment of frozen internal coordinates as follows:[154] "Constrained internal coordinates need extra attention because of possible redundancy in the coordinates and, therefore, potential overlap between the variable and frozen subspace. We found it practical to form an orthonormal frozen coordinate space by applying a Gram–Schmidt orthogonalization on the corresponding rows of the B-matrix and then subtracting the overlap from the rows related to variable internal coordinates. If the number of frozen variables is much less than the total number of coordinates, this extra treatment does not affect the scaling of the transformations."

Here, we adapt a procedure of Peng et al.,[155] derived from Pulay and Fogarasi[156] using our set of redundant internal coordinates and holonomic constraints. We construct a square projector $\underline{\boldsymbol{P}} = \underline{\boldsymbol{B}}\,\underline{\boldsymbol{B}}^T$ from the Wilson B-matrix and modify $\underline{\boldsymbol{P}}$ with the projector for the constraints $\underline{\boldsymbol{C}}$ according to eq. (3.17).

$$\underline{\boldsymbol{P}}' = \underline{\boldsymbol{P}} - \underline{\boldsymbol{P}}\,\underline{\boldsymbol{C}}(\underline{\boldsymbol{C}}\,\underline{\boldsymbol{P}}\,\underline{\boldsymbol{C}})^{-1}\underline{\boldsymbol{C}}\,\underline{\boldsymbol{P}} \tag{3.17}$$

$\underline{\boldsymbol{C}}$ is a diagonal matrix with 1 on the diagonal for indices of the constrained coordinates in $\underline{\boldsymbol{B}}$ and 0 elsewhere.

$$\underline{\boldsymbol{C}}_{ij} = \begin{cases} \delta_{ij} & \text{if vector } \underline{\boldsymbol{B}}_i \text{ corresponds to a constrained coordinate} \\ 0 & \text{elsewhere} \end{cases} \tag{3.18}$$

Eventually, we arrive at the modified Wilson B-matrix $\underline{\boldsymbol{B}}_c$ including the constraint(s) for our set of redundant internal coordinates using eq. (3.19).

$$\underline{\boldsymbol{B}}_c = \underline{\boldsymbol{A}}\,\underline{\boldsymbol{P}}' \tag{3.19}$$

Technically, this procedure corresponds to a temporary treatment of the system in a set of non-redundant delocalized internal coordinates for the reasons explained by Farkas and Schlegel (see

above).[154] Henceforth, $\underline{\boldsymbol{B}}_c$ and its Moore-Penrose pseudoinverse $\underline{\boldsymbol{A}}_c$ are used in the initial setup of the Hessian matrices in Cartesian and internal coordinate space as usual, see eqs. (3.9). An alternative and elegant implementation has been described by Bučko.[152]

MD trajectories of a constrained system sample phase space according to a biased configurational distribution, the so-called Blue Moon ensemble.[157] Observables have thus to be reweighted with the inverse mass metric tensor Z given in eq. (3.20) where m_i is the mass of atom i, μ represents the component of the Cartesian position vector r and ξ is the constrained coordinate.

$$Z = \sum_{i=1}^{N} \frac{1}{m_i} \sum_{\mu=x,y,z} \left(\frac{\partial \xi}{\partial r_{i,\mu}} \right)^2 \tag{3.20}$$

The λ-TI method can be adapted for Blue Moon ensembles when reweighted according to eq. (3.21).

$$\Delta A_{0\rightarrow 1} = \int_0^1 d\lambda \frac{1}{\left\langle Z^{-1/2} \right\rangle_{\lambda,\xi}} \left\langle Z^{-1/2} \left(V_1 - V_0 \right) \right\rangle_{\lambda,\xi} \tag{3.21}$$

3.1.2.3. Strategies to Determine Anharmonic Free Energy Barriers

In consequence of the introduced constraint, the prepared Hessians do not generate forces for distortions along the constrained coordinates and thus the imaginary vibrational mode of the TS is eliminated to zero. The formalism works equivalently for relaxed structures with real vibrational modes. This can be useful in designing strategies to compute anharmonic free energy barriers, since the free energy difference $\Delta A_{\mathrm{free_R}\rightarrow\xi_{\mathrm{ref,R}}}$ connecting the states of a free and a constrained (reactant) structure is relatively simple, see eq. (3.22).[158]

$$\Delta A_{\mathrm{free_R}\rightarrow\xi_{\mathrm{ref,R}}} = -k_\mathrm{B}T \ln \left(\frac{h}{k_\mathrm{B}T} \frac{\left\langle \left| \dot{\xi}^* \right| \right\rangle}{2} P(\xi_{\mathrm{ref,R}}) \right) \tag{3.22}$$

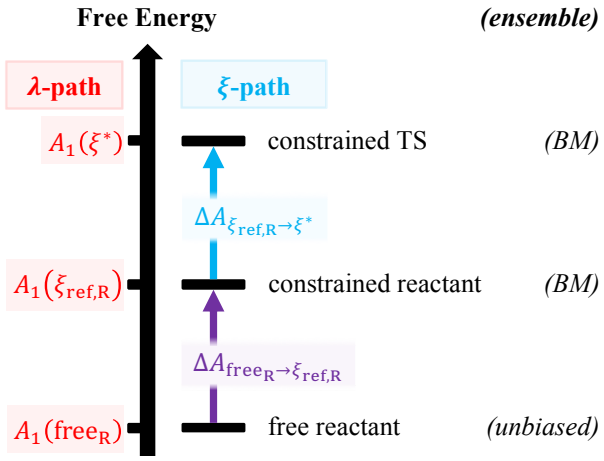
The term $\Delta A_{\mathrm{free_R}\rightarrow\xi_{\mathrm{ref,R}}}$ is computed from the temperature T, Planck constant h, Boltzmann constant k_B, the probability density of the reactant reference state $P(\xi_{\mathrm{ref,R}})$ as a function of the constrained coordinate in the probability distribution obtained from an unconstrained MD simulation of the reactant and the generalized TS velocity $\left\langle \left| \dot{\xi}^* \right| \right\rangle$ computed using a constrained MD simulation for the TS and eq. (3.23) with the inverse mass metric tensor Z defined as in eq. (3.20).

$$\left\langle \left| \dot{\xi}^* \right| \right\rangle = \frac{1}{\left\langle Z^{-1/2} \right\rangle_{\xi^*}} \sqrt{\frac{2k_\mathrm{B}T}{\pi}} \tag{3.23}$$

The contribution in eq. (3.22) is usually used together with the established thermodynamic ξ-path integration (ξ-TI) between Blue Moon ensembles by Ciccotti et al.,[105,157] which we use in this work to benchmark our λ-TI approach. While λ-TI can assign absolute values to any thermodynamic state, ξ-TI can only compute the difference between constrained systems. In detail, ξ-TI is based on the integration of free energy gradients along the reaction path connecting two states $\xi_{\mathrm{ref,R}}$ and ξ^* as in eq. (3.24).

$$\Delta A_{\xi_{\mathrm{ref,R}}\rightarrow\xi^*} = \int_{\xi_{\mathrm{ref,R}}}^{\xi^*} \left(\frac{\partial A}{\partial \xi} \right)_{\xi'} d\xi' \tag{3.24}$$

The free energy gradients are determined using MD simulations with the constraint of a constant ξ' according to eq. (3.25), where λ is the Lagrange multiplier associated with ξ' used in the SHAKE



Scheme 3.3.: Absolute anharmonic free energies A_1 for the constrained TS, the constrained reactant state and the free reactant state can be calculated with the λ-TI procedure (red). The free energy difference $\Delta A_{\xi_{\mathrm{ref,R}}\rightarrow\xi^*}$ (blue) of the constrained TS and constrained reactant state can be computed via the ξ-TI technique. The free energy contribution $\Delta A_{\mathrm{free_R}\rightarrow\xi_{\mathrm{ref,R}}}$ (violet) connects the thermodynamic states of the constrained and unconstrained reactant via eq. (3.22).

algorithm,[159] and Z is defined as in eq. (3.20). For further details on ξ-TI we recommend the original work of Carter et al. and Ciccotti and Ferrario[104,157] as well as the textbook by Chipot and Pohorille.[101]

$$\left(\frac{\partial A}{\partial \xi}\right)_{\xi'} = \frac{1}{\left\langle Z^{-1/2}\right\rangle_{\xi'}} \left\langle Z^{-1/2}\left[-\lambda + \frac{k_\mathrm{B}T}{Z}\sum_{i=1}^{N}\sum_{\mu=x,y,z}\frac{1}{m_i}\frac{\partial \xi}{\partial r_{i,\mu}}\frac{\partial Z}{\partial r_{\mu,i}}\right]\right\rangle \tag{3.25}$$

In fact, the calculation of a free energy barrier requires only two known thermodynamic states, namely the free reactant state and the TS. The constrained reactant with the same constraint as the TS is introduced as a third thermodynamic state, which is mandatory for ξ-TI and can also be computed directly with λ-TI, as illustrated in Scheme 3.3. Moreover, the illustration in Scheme 3.3 inspires three distinct and reasonable ways to compute the same anharmonic free energy barrier. The three different strategies are denoted as ξ-TI in eq. (3.26), λ-TI$_\mathrm{a}$ in eq. (3.27) and λ-TI$_\mathrm{b}$ in eq. (3.28).

$$\xi\text{-TI:} \quad \Delta A_{\mathrm{free_R}\rightarrow\xi^*} = \Delta A_{\xi_{\mathrm{ref,R}}\rightarrow\xi^*} + \Delta A_{\mathrm{free_R}\rightarrow\xi_{\mathrm{ref,R}}} \tag{3.26}$$

$$\lambda\text{-TI}_\mathrm{a}\text{:} \quad \Delta A_{\mathrm{free_R}\rightarrow\xi^*} = A_1(\xi^*) - A_1(\xi_{\mathrm{ref,R}}) + \Delta A_{\mathrm{free_R}\rightarrow\xi_{\mathrm{ref,R}}} \tag{3.27}$$

$$\lambda\text{-TI}_\mathrm{b}\text{:} \quad \Delta A_{\mathrm{free_R}\rightarrow\xi^*} = A_1(\xi^*) - A_1(\mathrm{free_R}) \tag{3.28}$$

Its ability to compute absolute anharmonic free energies for thermodynamic states is the major distinction of the λ-TI method from the reaction path-dependent ξ-TI method. Regarding the choice of λ-TI$_\mathrm{a}$ and λ-TI$_\mathrm{b}$ for calculations in practice, anharmonic corrections for the constrained reactant are presumably faster to converge due to the missing degree of freedom. However, an unconstrained MD simulation must be run to calculate the term $\Delta A_{\mathrm{free_R}\rightarrow\xi_{\mathrm{ref,R}}}$. On the other hand, the unconstrained MD simulation is not necessary, if the anharmonic correction is computed directly for the unconstrained reactant state. In this work, we use both λ-TI$_\mathrm{a}$ and λ-TI$_\mathrm{b}$ benchmarked against ξ-TI and report on their (dis-)advantages.

3.1.2.4. Choice of the Harmonic Reference System

The free energy of the harmonic reference system $A_{0,\boldsymbol{x}}$ can always be computed analytically for any harmonic potential. Furthermore, λ-TI determines $\Delta A_{0,\boldsymbol{x}\rightarrow 0,\boldsymbol{q}}$ and $\Delta A_{0,\boldsymbol{q}\rightarrow 1}$ with respect to the given

reference state so that the sum $A_{0,\boldsymbol{x}} + \Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}} + \Delta A_{0,\boldsymbol{q}\to 1}$ equals the free energy of the DFT-based system A_1. Therefore, the harmonic reference system can be chosen arbitrarily with no consequence for A_1. This choice can be used to improve the sampling efficiency[83] or to avoid numerical problems that occur at low λ when one or more harmonic force constants of the system are too weak. Specifically, displacements along soft modes may result in small changes in $V_{0,\boldsymbol{q}}$ but possibly cause huge changes in V_1, thus worsening the convergence of $\langle V_1 - V_{0,\boldsymbol{q}}\rangle_\lambda$, or even leading to problems in the convergence of the electronic structure calculations. In this work, the harmonic reference system is derived from the Hessian matrix computed numerically via a finite difference scheme. The eigenvalue spectrum (the force constants) is then inspected and all values lower than a certain limit (typically $1\,\mathrm{eV\,Å^{-2}}$ to $2\,\mathrm{eV\,Å^{-2}}$) are increased to that limit while preserving the original eigenvectors. Upon such a modification, the individual contributions $A_{0,\boldsymbol{x}}$, $\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}}$ and $\Delta A_{0,\boldsymbol{q}\to 1}$ are repartitioned in such a way that the resulting A_1 remains unchanged, provided all integrals over $\langle V_{0,\boldsymbol{q}} - V_{0,\boldsymbol{x}}\rangle_\lambda$ and $\langle V_1 - V_{0,\boldsymbol{q}}\rangle_\lambda$ are converged. A demonstration of this property for the HF molecule is presented in Section 3.2.2.1. Note that anharmonicity is an intrinsic property of the physical system. Just like A_1, anharmonicity is independent of the choice of the reference system used in the λ-TI calculation. Therefore, anharmonicity must always be referenced to the harmonic free energy determined on the basis of the actual Hessian matrix of the physical system, regardless of the choice of the harmonic reference system used in the λ-TI calculations.

3.1.2.5. Choice of Internal Coordinates

Although well converged results of the λ-TI method should be independent of the coordinates used in the simulations, the choice of coordinates still affects the sampling efficiency. Therefore, some care should be taken when choosing the coordinates $\boldsymbol{q}$ for the problem at hand. In this work, the following simple rules have been used to generate $\boldsymbol{q}$: (i) bond lengths, bond angles, and torsion angles are used for molecules and for special flexible sites of a substrate (such as Brønsted sites in acidic zeolites), (ii) interatomic distances with the atoms from the first three coordination spheres are used for all other atoms of a substrate and (iii) the position and orientation of a molecule adsorbed on the substrate is described using a function representing the coordination number (CN)[160] of an atom i of the adsorbate with respect to an atom j from the substrate defined according to eq. (3.29) where the summations are over all or selected substrate atoms N_{sub} contained in a single unit cell and all translations of the unit cell $\boldsymbol{L} = \{\pm l_1, \pm l_2, \pm l_3\}$, $r_{ij,\mathbf{L}}$ is a distance between an atom i and atom j shifted from the original unit cell ($\boldsymbol{L} = 0$) by a translation $l_1\boldsymbol{a}_1 + l_2\boldsymbol{a}_2 + l_3\boldsymbol{a}_3$ along the lattice vectors $\boldsymbol{a}_i$, and $R_{i,j}$ is a reference distance between the atoms i and j.

$$CN_i = \sum_{j}^{N_{\mathrm{sub}}} \sum_{\mathbf{L}} \frac{1 - (r_{ij,\mathbf{L}}/R_{ij})^9}{1 - (r_{ij,\mathbf{L}}/R_{ij})^{14}} \tag{3.29}$$

Since the function $\frac{1-(r_{ij,\boldsymbol{L}}/R_{ij})^9}{1-(r_{ij,\boldsymbol{L}}/R_{ij})^{14}}$ decays rapidly with increasing $r_{ij,\boldsymbol{L}}$, the (infinite) lattice sum in eq. (3.29) can be truncated when $r_{ij,\mathbf{L}}$ exceeds a certain cutoff radius (here 30 Å). Note that the local symmetry can be imposed by replacing individual primitive coordinates by their suitably chosen linear combinations. As is a common practice in atomic relaxation using internal coordinates,[155,156,161–165] the coordinate selection in point (i) is done automatically in three steps: First, distances between all pairs of atoms are computed, out of which those being shorter than a certain cutoff radius (usually based on the sum of scaled covalent radii) are used to define bond lengths. Second, all bond lengths sharing one common atom are used to define bond angles. To avoid numerical instability due to a singularity in the derivative of angles close to 180°, all straight angles (>165°) are excluded. Third, bond lengths and angles sharing one common atom are used to define dihedral angles. A similar scheme is also

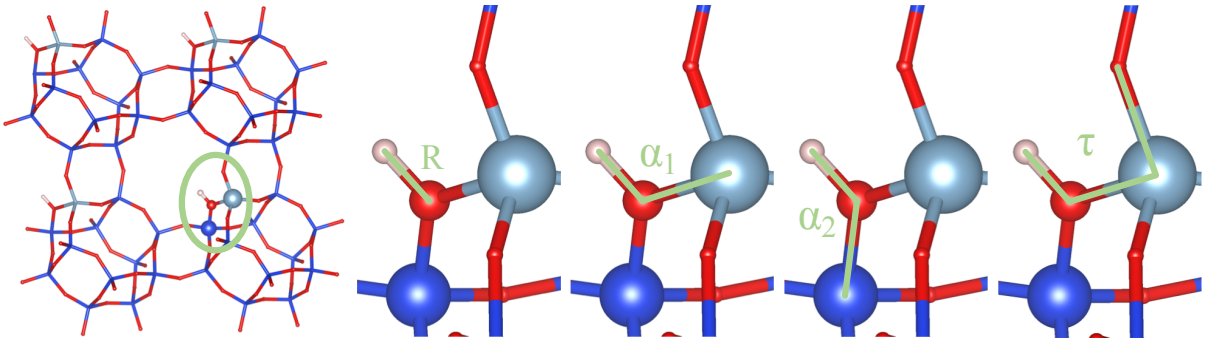


Figure 3.2.: Definition of a set of redundant internal coordinates fully determining the position of the proton of the acidic site in H-SSZ-13 without adsorbate: bond length (R), two bond angles (α_1 and α_2), and torsion angle (τ). The other atoms of the substrate are defined by interatomic distances to the atoms of their first three coordination spheres. Color code: Si: dark blue, Al: light blue, O: red, H: white. Reprinted with permission from [17]. Copyright 2021 American Chemical Society.

used for point (ii) but bond angles and dihedrals are replaced by distances linking the two terminal atoms. This strategy is used to avoid the above-mentioned numerical problems with singularities in the derivatives of angles approaching 180° during the MD. An illustration of the coordinate choice in H-SSZ-13, the acidic aluminum-substituted zeolite with the chabazite topology, is provided in Figure 3.2. Some modifications of this strategy might be needed for different adsorption problems. Future work is planned to explore the use of a potentially more universal set of coordinates based on the smooth overlap of atomic positions (SOAP)[166] or Coulomb-type matrices,[167] which are successfully employed in machine learning schemes in chemistry.[168] See Section 3.4 for a brief communication on this topic.

3.1.3. Exemplary Coordinate Transformation Solving the Rotational Problem

The following illustration is intended to demonstrate the usefulness of a coordinate transformation for solving the rotational problem. Consider a harmonic spring with force constant C, equilibration length R_0 and instantaneous length R as sketched in Figure 3.3. The spring existing in 2D is attached to a rigid wall while the other end is free. The position vector $\boldsymbol{x}$ and the equilibrium position vector $\boldsymbol{x}_0$ are defined as in eqs. (3.30).

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} R \\ 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{x}_0 = \begin{pmatrix} x_{1,0} \\ x_{2,0} \end{pmatrix} = \begin{pmatrix} R_0 \\ 0 \end{pmatrix}. \tag{3.30}$$

The force acting on the free end point of the spring in Cartesian coordinates can be obtained from the Cartesian Hessian matrix $\underline{\boldsymbol{H}}^{\boldsymbol{x}}$ and the displacement from the equilibration position, see eqs. (3.31).

$$\boldsymbol{f}^{\boldsymbol{x}} = -\underline{\boldsymbol{H}}^{\boldsymbol{x}} (\boldsymbol{x} - \boldsymbol{x}_0) = -\begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} R - R_0 \\ 0 \end{pmatrix} = -\begin{pmatrix} C (R - R_0) \\ 0 \end{pmatrix} \quad \text{with} \quad \underline{\boldsymbol{H}}^{\boldsymbol{x}} = \begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} \tag{3.31}$$

Let us now consider the same spring rotated by an angle α, without changing R. The new position $\boldsymbol{x}'$ is obtained by application of the rotation matrix, see eq. (3.32).

$$\boldsymbol{x}' = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} R \\ 0 \end{pmatrix} = \begin{pmatrix} R\cos\alpha \\ R\sin\alpha \end{pmatrix}. \tag{3.32}$$

Since the force is a covector, it must transform just like $\boldsymbol{x}$ upon rotation, see eq. (3.33).

$$\boldsymbol{f}^{\boldsymbol{x}\prime} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \boldsymbol{f}^{\boldsymbol{x}} = -\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} C (R - R_0) \\ 0 \end{pmatrix} = -C (R - R_0) \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} \tag{3.33}$$

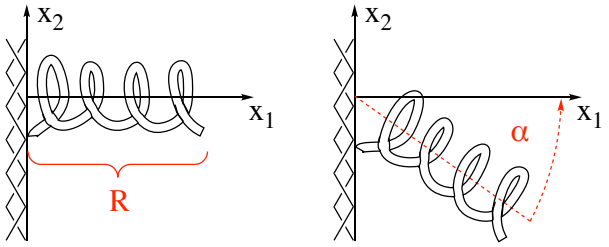


Figure 3.3.: Left: Harmonic spring attached to a rigid wall. Right: Harmonic spring rotated by an angle α (note that here $\alpha < 0$). Reprinted with permission from [17]. Copyright 2021 American Chemical Society.

Note that $\left|\boldsymbol{f}^{\boldsymbol{x}\prime}\right| = |\boldsymbol{f}^{\boldsymbol{x}}|$ as it should be since the rotation cannot change the magnitude of the force. However, a naive usage of the Hessian matrix leads to incorrect forces, see eq. (3.34).

$$\boldsymbol{f}^{\boldsymbol{x}'}_{\text{naive}} = -\underline{\boldsymbol{H}}^{\boldsymbol{x}} \left(\boldsymbol{x}' - \boldsymbol{x}_0\right) = -\begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} R\cos\alpha - R_0 \\ R\sin\alpha \end{pmatrix} = -\begin{pmatrix} C\left(R\cos\alpha - R_0\right) \\ 0 \end{pmatrix}. \tag{3.34}$$

This is obviously wrong since $\left|\boldsymbol{f}^{\boldsymbol{x}'}_{\text{naive}}\right| \neq |\boldsymbol{f}^{\boldsymbol{x}}|$ for $\alpha \neq 0$. The problem is tackled easily by the coordinate transformation $\boldsymbol{x} \rightarrow \boldsymbol{q}$, that is $\{x_1, x_2\} \rightarrow \{R\}$ through which the forces in Cartesian coordinates $\boldsymbol{x}$ can be obtained from the forces in internal coordinates $\boldsymbol{q}$. The coordinate transformation of the Hessian matrix ($\underline{\boldsymbol{H}}^{\boldsymbol{x}} \rightarrow \underline{\boldsymbol{H}}^{\boldsymbol{q}}$) is achieved via the transformation relation[148–152] in eq. (3.9) with the matrices given by eqs. (3.35) and (3.36).

$$\underline{\boldsymbol{B}}_{\boldsymbol{x}_0} = \begin{pmatrix} \left.\frac{\partial R}{\partial x_1}\right|_{\boldsymbol{x}_0} & \left.\frac{\partial R}{\partial x_2}\right|_{\boldsymbol{x}_0} \end{pmatrix} = \begin{pmatrix} 1 & 0 \end{pmatrix} \tag{3.35}$$

$$\underline{\boldsymbol{A}}_{\boldsymbol{x}_0} = \left(\underline{\boldsymbol{B}}^T_{\boldsymbol{x}_0}\underline{\boldsymbol{B}}_{\boldsymbol{x}_0}\right)^{-1}\underline{\boldsymbol{B}}^T_{\boldsymbol{x}_0} = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \end{pmatrix}\right)^{-1}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{3.36}$$

For the example with the spring, we obtain the Hessian matrix in eq. (3.37) and the force in eq. (3.38) for internal coordinates.

$$\underline{\boldsymbol{H}}^{\boldsymbol{q}} = \begin{pmatrix} 1 & 0 \end{pmatrix}\begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} C & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = C. \tag{3.37}$$

$$\boldsymbol{f}^{\boldsymbol{q}} = -\underline{\boldsymbol{H}}^{\boldsymbol{q}}\left(\boldsymbol{q} - \boldsymbol{q}_0\right) = -C\left(R - R_0\right) \tag{3.38}$$

The forces in internal coordinates are then transformed back to Cartesian coordinates using eq. (3.39) where we made use of the relations $\left.\frac{\partial R}{\partial x_1}\right|_{\boldsymbol{x}'} = \frac{x_1'}{R} = \cos\alpha$ and $\left.\frac{\partial R}{\partial x_2}\right|_{\boldsymbol{x}'} = \frac{x_2'}{R} = \sin\alpha$.[148–152] By comparison with eq. (3.33) we see that eq. (3.39) yields the correct result for the forces.

$$\boldsymbol{f}^{\boldsymbol{x}'} = \underline{\boldsymbol{B}}^T_{\boldsymbol{x}}\boldsymbol{f}^{\boldsymbol{q}} = -\begin{pmatrix} \left.\frac{\partial R}{\partial x_1}\right|_{\boldsymbol{x}'} \\ \left.\frac{\partial R}{\partial x_2}\right|_{\boldsymbol{x}'} \end{pmatrix} C\left(R - R_0\right) = -C\left(R - R_0\right)\begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} \tag{3.39}$$

3.2. Anharmonic Corrections to Adsorption Free Energies

Preface. Adsorption (and desorption) processes are often governed by weak interactions for which the estimate of entropy contributions by means of the harmonic approximation is prone to be inaccurate. Thermodynamic λ-path integration (λ-TI) from the harmonic to the anharmonically interacting system, as presented in the previous section, can be used to compute anharmonic corrections. The application

of λ-TI is now discussed on three model systems and compared with exact semi-analytical solutions. Moreover, the importance of anharmonic vibrations, hindered rotations and hindered translations (dissociation) is illustrated and quantified. Finally, the λ-TI method is applied to study the adsorption of small particles in zeolite H-SSZ-13.

3.2.1. Overview and Simulation Details

The free energy of a system has contributions from both its enthalpy and entropy. Hence both contributions need to be considered to accurately calculate changes in free energies. An accurate description of entropy changes is particularly important for adsorption and desorption processes, as these often occur with a large loss (or gain) of entropy of an adsorbate. Adsorption steps are key for many important processes, for instance, they are essential for heterogeneous catalysis where the interaction of a solid catalyst with a gas phase molecule is at the core of its function.[169] Unfortunately, the accurate treatment of entropy changes during these processes is also very challenging and typically approximated by consideration of harmonic potentials for the adsorbed system.

However, weakly interacting adsorbates often exhibit anharmonic degrees of freedom such as dissociation (translation) and rotation which are extremely poorly described[109,110,170] by the most popular and conceptually simple harmonic (oscillator) approximation,[85] that is also commonly used for zeolites. Moreover, this static approach and descriptors based on confinement[171] do neither consider adjacent local minima nor multiple adsorption sites. Little advances have been achieved to properly account for adsorption entropy beyond the harmonic approximation although its necessity has been demonstrated for adsorption on metal surfaces and at acidic sites of zeolites.[144,172] In fact, recent computational studies indicate that there can be large deviations in the prediction of adsorption free energies, as it has been shown, for instance, for ethanol adsorption in zeolites.[144,173] Several strategies have evolved over the past years to tackle these problems. By identification and separate treatment of the most important anharmonic degrees of freedom some approaches attempt to improve the accuracy of the static approach, where the thermodynamics of the system are deduced from the properties of the potential energy surface in vicinity to the stationary state of interest. For hindered rotational motion, for instance, potential energy models of different complexity (ranging from a single goniometric function up to explicitly sampled energy profiles) are used in classical,[75,174,175] semi-classical,[75,174,175] or quantum mechanical[176] treatment and similar models exist also for hindered translations.[175] Nonetheless, the reliability of these approaches depends strongly on the thermodynamic conditions considered in the simulation. An elegant demonstration of this problem has been provided in a recent study of Jørgensen and Grönbeck, where predictions of harmonic, hindered, and free translator approximations in the calculation of the entropy of CO and O adsorbed on Pt(111) have been tested against the complete potential energy sampling approach and experiments.[76] In the latter work, it was clearly shown that none of the simple approximations provided reliable predictions over a wide range of temperatures. Perhaps the most sophisticated correction to the static approach has been developed by Sauer and coworkers in which anharmonic vibrational partition functions have been determined by solving 1D Schrödinger eqs. defined for the potential energies explicitly sampled along individual vibrational eigenmodes expressed in terms of internal coordinates.[74,134,135,144,177,178] According to Piccini and Sauer, anharmonic contributions stabilize ethanol in zeolite H-MFI by 13 kJ mol^{-1} at 300 K (20 kJ mol^{-1} at 400 K).[144] Using a quasi-harmonic approximation (QHA) for ethanol adsorption in zeolite H-ZSM-5 Alexopoulos et al. claim the traditional harmonic approximation overestimates adsorption free energies by 20 kJ mol^{-1} to 50 kJ mol^{-1}.[173] The QHA relies on a vibrational density of states from molecular dynamics (MD), thereby taking into account raised temperature and the average of adjacent local minima. To date, accurate theoretical (but also experimental) reports on adsorption free energies are

rare mostly in lack of accurate accounts for entropy. Even simple internal rotations in n-alkanes are incorrectly described by the harmonic approximation and have therefore been subject to numerous investigations.[179–181]

Our newly developed λ-TI method is used below to study increasingly complex problems from diatomic molecules up to the adsorption of small adsorbates in zeolites, which has been thoroughly investigated experimentally[182] as well as computationally[134,144,172,173,177,178,183–192] since the 1970s.

3.2.1.1. Simulation Details

Periodic density functional theory (DFT) calculations have been performed using the VASP code.[193–196] The Kohn-Sham eqs. have been solved variationally in a plane-wave basis set using the PAW method of Blöchl, as adapted by Kresse and Joubert with standard PAW potentials.[197,198] The PBE[96] density functional with the D2 dispersion correction (zero damping) of Grimme[60] (PBE-D2) as implemented in VASP[199] was applied for geometry optimizations and MD simulations using an energy cutoff of 400 eV for all computations but the third case study with the default value of 266.408 eV for argon. Convergence criteria of 10^{-7} eV and 0.005 eV $Å^{-1}$ were applied to SCF-cycles and geometry optimizations, respectively. Constrained geometry relaxations have been performed using the program GADGET.[151,152] The Brillouin-zone sampling was restricted to the Γ-point.[200] Full Hessian matrices were computed for all systems using a centered finite difference scheme. Adsorption free energies were computed at a concentration of one adsorbate per unit cell (corresponding to one adsorbate per Brønsted acid site) for the adsorbed state relative to the ideal gas state at a reference pressure of 101.325 kPa. DFT-based MD simulations in the NVT ensemble were performed using the Andersen thermostat.[98] Similar to previous studies,[111,172] hydrogen atoms were treated as tritium (mass = 3 u) because this allows to use larger time steps. Consistently, the mass of tritium was also used for the computation of harmonic frequencies. Equilibration periods in MD were determined with the Mann-Kendall test for trend and variation.[201] The individual specific settings, the length of MD trajectories or the integration step size, are discussed for each test case separately. The numerical integration of eq. (3.4) has been performed using the Simpson integration scheme adapted to irregularly spaced grid points.[202–204] The statistical error has been determined as described in the next section.

3.2.1.2. Integration Scheme and Error Estimation

Numerical integration. Throughout this work the composite Simpson integration adapted for irregular spaced data[202–204] was used to evaluate the integrals over the λ-path involved in our λ-TI method. In this integration scheme, the integral I is determined from eq. (3.40) where f_k denotes the value of the integrand (in our case corresponding to $\langle V_1 - V_{0,\boldsymbol{q}}\rangle_\lambda$ or $\langle V_{0,\boldsymbol{q}} - V_{0,\boldsymbol{x}}\rangle_\lambda$) evaluated at the k-th of a total of N grid points defined for different values of λ.

$$I = \int_a^b f(\lambda)d\lambda \approx \sum_{i=0}^{N/2-1} (\alpha_i f_{2i+2} + \beta_i f_{2i+1} + \eta_i f_{2i}) \tag{3.40}$$

The coefficients α_i, β_i and η_i are defined in eqs. (3.41) to (3.43).

$$\alpha_i = \frac{2h_{2i+1}^3 - h_{2i}^3 + 3h_{2i}h_{2i+1}^2}{6h_{2i+1}\left(h_{2i+1} + h_{2i}\right)} \tag{3.41}$$

$$\beta_i = \frac{h_{2i+1}^3 + h_{2i}^3 + 3h_{2i+1}h_{2i}\left(h_{2i+1} + h_{2i}\right)}{6h_{2i+1}h_{2i}} \tag{3.42}$$

$$\eta_i = \frac{2h_{2i}^3 - h_{2i+1}^3 + 3h_{2i+1}h_{2i}^2}{6h_{2i}\left(h_{2i+1} + h_{2i}\right)} \tag{3.43}$$

Note that for even N the last interval is treated with the coefficients in eqs. (3.44).

$$\alpha = \frac{2h_{N-1}^2 + 3h_{N-1}h_{N-2}}{6\left(h_{N-2} + h_{N-1}\right)} \qquad \beta = \frac{h_{N-1}^2 + 3h_{N-1}h_{N-2}}{6h_{N-2}} \qquad \eta = \frac{h_{N-1}^3}{6h_{N-2}\left(h_{N-2} + h_{N-1}\right)} \tag{3.44}$$

Statistical error estimation. The standard error ϵ_I in I is estimated from error propagation theory applied to eq. (3.40), see eq. (3.45).

$$\epsilon_I = \sqrt{\sum_{k=1}^{N}\left(\frac{\partial I}{\partial f_k}\epsilon_{f_k}\right)^2} \tag{3.45}$$

In our calculations, the partial derivatives $\frac{\partial I}{\partial f_k}$ were computed numerically via a finite difference scheme with a numerical step size of $\Delta f \approx 10^{-3}$. Upon elimination of the time correlation from the MD data via the procedure of block averaging,[99] ϵ_{f_k} at each grid point k was obtained.

3.2.2. Results and Discussion

3.2.2.1. Simple Model Systems

To demonstrate the validity of our approach, we start our discussion by considering three simple model systems representing one-dimensional (1D) or quasi-1D problems for which exact semi-analytical solutions for $\Delta A_{0,\boldsymbol{q}\to 1}$ can be derived from their 1D potential energy profiles. Furthermore, the three cases discussed here also illustrate the qualitative differences and importance of anharmonic effects for the free energy contribution of vibrations (HF molecule), hindered translations (argon dimer), and hindered rotations (ethane).

HF molecule. Due to its strongly anharmonic ($\omega_e\chi_e = 89.9\,\mathrm{cm}^{-1}$)[205] bond, the HF molecule represents a prototype system to study the effect of anharmonicity on the free energy of vibration. As this effect is relatively small, a high temperature ($T = 2000\,\mathrm{K}$) was chosen to observe a significant deviation from the harmonic model. With the interatomic distance R being the only variable of the corresponding potential energy, this system represents a truly 1D problem. Since R is a phase volume conserving coordinate, the term $\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}}$ is zero in this case, in other words, $A_{0,\boldsymbol{x}} = A_{0,\boldsymbol{q}}$. The anharmonic potential energy profile $V_1(R)$ in Figure 3.4a was computed using DFT on a regular mesh of points defined on the interval between $R_{\min} = 0.70\,\text{Å}$ and $R_{\max} = 1.40\,\text{Å}$ with grid points separated by an increment $\Delta R = 0.01\,\text{Å}$. The harmonic potential energy takes the form $V_{0,\boldsymbol{q}}(R) = V_{0,\boldsymbol{q}}(R_0) + \frac{1}{2}C(R-R_0)^2$ whereby the equilibrium bond length $R_0 = 0.938\,\text{Å}$ was determined using structural relaxation, the force constant $C = 56.157\,\mathrm{eV\,Å}^{-2}$ was identified using a finite difference scheme and the term $V_{0,\boldsymbol{q}}(R_0) = V_{0,\boldsymbol{x}}(\boldsymbol{x}_0) = V_1(\boldsymbol{x}_0)$ is the ionic

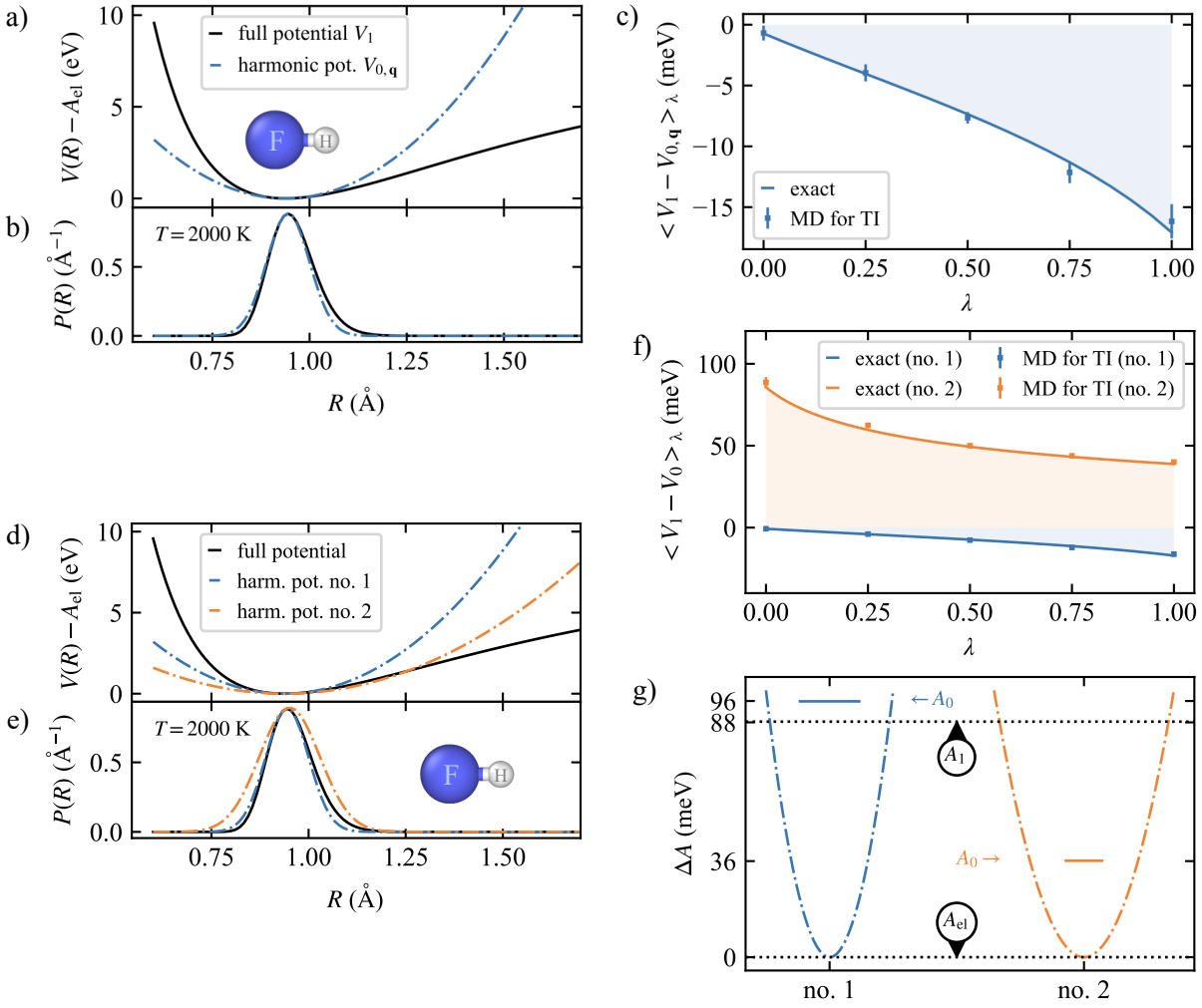


Figure 3.4.: Case study 1 on anharmonic vibration of a covalent bond in the HF molecule. (a) Potential of the interacting system (DFT) compared to its harmonic approximation. (b) Semi-analytic probability densities emerging from the potentials. (c) λ-TI: Comparison of the exact semi-analytic solution to the result obtained by MD runs at 2000 K. Integration leads to the anharmonic correction $\Delta A_{0,\boldsymbol{q}\to 1}$ listed in Table 3.1. Reprinted (a, b, c) with permission from [17]. Copyright 2021 American Chemical Society. (d) Comparison of the first harmonic potential to a second one arbitrarily set to one half of the first. (e) Comparison of semi-analytic probability densities for the two harmonic potentials. (f) The anharmonic corrections to the two harmonic reference systems differ, but the semi-analytic solutions match in each case. (g) Illustration of free energy contributions from different reference systems with different harmonic free energies A_0 resulting in the same one and only anharmonic free energy A_1.

ground state energy. With these prerequisites the exact value of $\Delta A^{\text{ref}}_{0,\boldsymbol{q}\to 1}$ can be determined from eq. (3.46).

$$\Delta A^{\text{ref}}_{0,\boldsymbol{q}\to 1} = -\frac{1}{\beta} \ln \left\{ \frac{\int_{R_{\text{min}}}^{R_{\text{max}}} dR\, R^2 e^{-\beta V_1(R)}}{\int_{R_{\text{min}}}^{R_{\text{max}}} dR\, R^2 e^{-\beta V_{0,\boldsymbol{q}}(R)}} \right\} \tag{3.46}$$

Note that the integration limits of eq. (3.46) were chosen so as to fully cover all values generated in the MD simulation. Indeed, the contribution of values outside the interval $R_{\text{min}} \leq R \leq R_{\text{max}}$ to $\Delta A^{\text{ref}}_{0,\boldsymbol{q}\to 1}$ is

Table 3.1.: Vibrational free energy contributions determined in case studies 1-3. The classical harmonic contribution ($A_{0,\boldsymbol{x}} - A_{\mathrm{el}}$) was calculated using eq. (3.2). The anharmonic corrections ($\Delta A_{0,\boldsymbol{q}\to 1}$) determined by λ-TI based on MD are compared to the semi-analytic solutions calculated using eq. (3.46) for case studies 1 and 2, and eq. (3.50) for case study 3. Reprinted with permission from [17]. Copyright 2021 American Chemical Society.

Case study	T (K)	$A_{0,\boldsymbol{x}} - A_{\mathrm{el}}(\boldsymbol{x}_0)$	$\Delta A_{0,\boldsymbol{q}\to 1}$ (meV)	
		(meV)	λ-TI of MD	reference
HF molecule[a]	2000	95.9	-8.1 ± 0.5	-7.6
argon dimer[a]	300	-39.2	-28.2 ± 2.5	-27.4
ethane[b]	400	762.9	-3.3 ± 2.2	-4.1

[a] trivially $\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}} = 0$, hence $A_{0,\boldsymbol{x}} = A_{0,\boldsymbol{q}}$. [b] $\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}} = (-0.5 \pm 0.4)$ meV.

negligible at our target temperature, see Figure 3.4b. Furthermore, it follows from eq. (3.46) that the exact reference value of $\langle V_1 - V_{0,\boldsymbol{q}}\rangle_\lambda$ can be obtained from $V_1(R)$ and $V_{0,\boldsymbol{q}}(R)$ according to eq. (3.47).

$$\langle V_1 - V_{0,\boldsymbol{q}}\rangle_\lambda^{\mathrm{ref}} = \frac{\int_{R_{\min}}^{R_{\max}} dR\, R^2 \left[V_1(R) - V_{0,\boldsymbol{q}}(R)\right] e^{-\beta\left[\lambda V_1(R) + (1-\lambda) V_{0,\boldsymbol{q}}(R)\right]}}{\int_{R_{\min}}^{R_{\max}} dR\, R^2\, e^{-\beta\left[\lambda V_1(R) + (1-\lambda) V_{0,\boldsymbol{q}}(R)\right]}} \tag{3.47}$$

We have exploited the latter expression to evaluate the accuracy of $\langle V_1 - V_{0,\boldsymbol{q}}\rangle_\lambda$ determined from MD in Figure 3.4c. The energy differences needed for λ-TI in eq. (3.4) were obtained for five evenly spaced values of λ by MD simulations with the integration step of 0.25 fs and the lengths of production runs of at least 2.5 ps with an Andersen thermostat probability of 0.05. In all simulations, a cubic unit cell with an edge length of 15 Å containing one molecule was used. As evident from Figure 3.4c, the MD results for $\langle V_1 - V_{0,\boldsymbol{q}}\rangle_\lambda$ are close to the semi-analytic solution of eq. (3.47). Using eq. (3.5), we find the anharmonic correction to the vibrational free energy $\Delta A_{0,\boldsymbol{q}\to 1} = (-8.1 \pm 0.5)$ meV which is, within the statistical error, identical to the reference value $\Delta A_{0,\boldsymbol{q}\to 1}^{\mathrm{ref}} = -7.6$ meV. Moreover, this number is in excellent agreement with fully quantum mechanical treatment, as shown in Section 3.2.2.2. Combined with the harmonic free energy $A_{0,\boldsymbol{x}} - A_{el}(\boldsymbol{x}_0) = 95.9$ meV, our calculations yield the total free energy of $A_1 - A_{el}(\boldsymbol{x}_0) = 87.8$ meV, while the reference value is 88.3 meV (see also Table 3.1).

Harmonic approximation $A_{0,\boldsymbol{x}}$ and anharmonic correction $\Delta A_{0,\boldsymbol{x}\to 1} = \Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}} + \Delta A_{0,\boldsymbol{q}\to 1}$ in the calculations presented above were based on the actual force constant of HF ($C = 56.157\,\mathrm{eV\,Å^{-2}}$) determined using a centered finite difference scheme on the DFT level of theory. We are, however, not bound to this choice. As discussed in Section 3.1.2.4, the parameters of the harmonic reference force field can be set arbitrarily without affecting the resulting value of A_1, provided all integrals of $\langle V_{0,\boldsymbol{q}} - V_{0,\boldsymbol{x}}\rangle_\lambda$ and $\langle V_1 - V_{0,\boldsymbol{q}}\rangle_\lambda$ over λ are converged. This property can easily be demonstrated. To this end, we define a new harmonic reference system with a force constant of $28.079\,\mathrm{eV\,Å^{-2}}$ which is one half of the constant determined for the original reference system, see Figure 3.4d and e. While $A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}}$ remains zero, the values of $A_{0,\boldsymbol{x}} - A_{\mathrm{el}}$ and $\Delta A_{0,\boldsymbol{q}\to 1}$ computed with this new reference are 36.1 meV and 52.8 meV, respectively. Despite this dramatic change in both terms, the resulting value of $A_1 = 88.9$ meV is virtually identical to the value obtained with the original harmonic reference system (87.8 meV), as it should be (see Figure 3.4f and g).

Argon dimer. In analogy to the HF molecule discussed above, the potential energy of the argon dimer (Ar_2) discussed in this section is a 1D function of the interatomic distance R. In contrast to HF, however, the Ar atoms of the dimer are attracted by only a weak dispersion interaction and therefore the latter system collapses already at relatively low temperatures well below 300 K considered here. Upon the

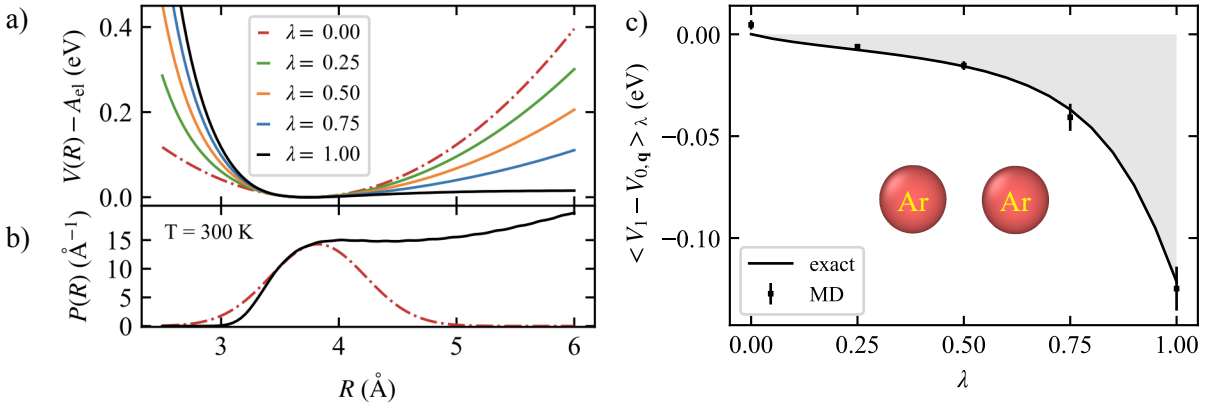


Figure 3.5.: Case study 2 on hindered translation (dissociation) of the argon dimer. (a) Visualization of potential morphing due to the linear combination of Hamiltonians. (b) Semi-analytic probability densities for harmonic ($\lambda = 0$) and DFT-based ($\lambda = 1$) system. (c) λ-TI: Comparison of exact semi-analytic solution to averaged energy differences obtained by MD runs at 300 K. Integration leads to the anharmonic correction $\Delta A_{0,\boldsymbol{q}\to 1}$. Reprinted with permission from [17]. Copyright 2021 American Chemical Society.

collapse, the vibrational degree of freedom is converted into hindered translations of the atoms which cannot be reasonably well described by a harmonic model. We therefore consider the argon dimer as a prototype system to study the free energy change due to dissociation. This transformation of a bond into two disjoint fragments occurs frequently in adsorption problems and chemical reactions.

As in the case of HF, the term $\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}}$ equals zero and our MD results are compared to analytical solutions represented by eqs. (3.46) and (3.47). If the dimer collapses, the value of $\Delta A_{0,\boldsymbol{q}\to 1}$ will strongly depend on the volume available for the atoms to move. In our calculations, we considered only the configurations where the mutual distance of two Ar atoms did not exceed 6 Å, so $R_{\max} = 6.00$ Å was used in eq. (3.46). The harmonic potential $V_{0,\boldsymbol{q}}(R)$ was built using $C = 0.154$ eV Å^{-2} and $R_0 = 3.705$ Å yielding a harmonic contribution of $A_{0,\boldsymbol{x}} - A_{el}(\boldsymbol{x}_0) = -39.2$ meV to the free energy. The anharmonic potential energy profile $V_1(R)$ was evaluated using DFT on a regular mesh of grid points defined on the interval between $R_{\min} = 2.50$ Å and $R_{\max} = 6.00$ Å with grid points separated by an increment $\Delta R = 0.10$ Å. The λ-dependent morphing of the driving potential is shown in Figure 3.5a. In excellent agreement with the quantum mechanical description in Section 3.2.2.2, the classical reference value of $\Delta A^{\text{ref}}_{0,\boldsymbol{q}\to 1}$ evaluated using eq. (3.46) is -27.4 meV resulting in the total free energy of $A_1 - A_{\text{el}}(\boldsymbol{x}_0) = -66.6$ meV. For this dissociation, the anharmonic correction is in the order of magnitude of the contribution from the harmonic approximation. In comparison to the analogous HF system with a purely vibrational degree of freedom, the anharmonic contribution is, despite the significantly lower temperature, much higher for the transformation of a vibration into translation.

Owing to the large atomic mass of Ar, the MD simulations have been conducted with an integration step size of 5 fs and an Andersen thermostat probability of 0.1. The length of the production period for each value of λ was at least 45 ps. A cubic box with an edge size of 15 Å containing one Ar_2 was used and the condition $R \leq R_{\max}$ was applied *a posteriori* by expressing the average $\langle V_1 - V_{0,\boldsymbol{q}}\rangle_\lambda$ on the subset of MD frames where the condition was met using eq. (3.48) with $\boldsymbol{x}$ and $\boldsymbol{p}$ being the atomic positions and momenta, respectively.

$$\langle V_1 - V_{0,\boldsymbol{q}}\rangle_\lambda = \frac{\int_0^{R_{\max}} dR \int d\boldsymbol{p}\, d\boldsymbol{x}\, (V_1(\boldsymbol{x}) - V_{0,\boldsymbol{q}}(\boldsymbol{x}))\delta(R - R(\boldsymbol{x}))e^{-\mathcal{H}_\lambda(\boldsymbol{p},\boldsymbol{x})/k_{\text{B}}T}}{\int_0^{R_{\max}} dR \int d\boldsymbol{p}\, d\boldsymbol{x}\, \delta(R - R(\boldsymbol{x}))e^{-\mathcal{H}_\lambda(\boldsymbol{p},\boldsymbol{x})/k_{\text{B}}T}} \tag{3.48}$$

A visual inspection of the configurations generated for $\lambda = 1$ (corresponding to MD driven by a full DFT potential) confirmed that the argon dimer was indeed unstable and collapsed. As shown in Figure 3.5c,

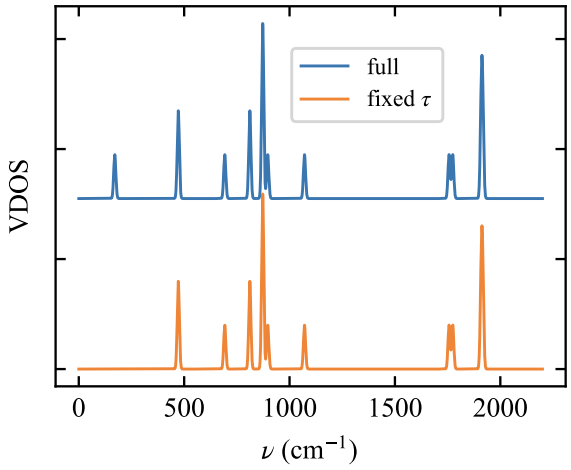


Figure 3.6.: Vibrational density of states (VDOS) computed from the (un-)constrained Hessian matrix for ethane without (full) and with (fixed τ) a constraint imposed on the parameter τ defined by eq. (3.49). Consistently, the mass of tritium was used for H atoms. Reprinted with permission from [17]. Copyright 2021 American Chemical Society.

the average values $\langle V_1 - V_{0,\boldsymbol{q}} \rangle_\lambda$ evaluated for five evenly spaced values of λ are close to the reference results obtained from eq. (3.47). The value of $\Delta A_{0,\boldsymbol{q}\to 1} = (-28.2 \pm 2.5)$ meV computed from MD is therefore in excellent agreement with $\Delta A^{\text{ref}}_{0,\boldsymbol{q}\to 1}$ (see also Table 3.1).

Ethane molecule. Unlike the two systems discussed above, the ethane molecule possesses multiple vibrational degrees of freedom, which can be seen in the vibrational density of states (VDOS) in Figure 3.6. However, it also contains one vibrational degree of freedom that can be converted into an internal rotation corresponding to a mutual rotation of two CH_3 groups upon a thermal excitation at a modest temperature. Since the anharmonic contribution of vibrations is by orders of magnitude smaller than that of rotations or translations, the former can be neglected at low temperatures and the term $\Delta A^{\text{ref}}_{0,\boldsymbol{q}\to 1}$ can thus be effectively determined from a 1D potential energy depending on a suitably chosen geometric parameter driving the internal rotation. For this purpose, we choose a linear combination τ of the torsion angles τ_1, τ_2 and τ_3 shown in Figure 3.7c and eq. (3.49).

$$\tau = \frac{1}{3}(\tau_1 + \tau_2 + \tau_3) \tag{3.49}$$

Clearly, the potential $V_1(\tau)$ is invariant with regard to internal rotation by 120°, thus reflecting the basic symmetry of ethane, see Figure 3.7. In fact, τ as defined above is nearly a perfect representation of the vibrational eigenmode corresponding to a hindered rotation of the CH_3 groups, which is evident from Figure 3.6 showing the VDOS of ethane before and after imposing a constraint on the value of τ. Obviously, fixing τ leads to a complete elimination of one vibrational mode while the remaining modes remain virtually unaffected.

The analogue of eq. (3.46) representing our semi-analytic reference for the anharmonic contribution to the free energy is expressed in eq. (3.50) where the first summand corresponds to the term $\Delta A'_{0,\boldsymbol{q}\to 1}$ that is obtained by the λ-TI approach while the latter ($-\frac{1}{\beta}\ln\frac{1}{\sigma_r}$) originates from the intrinsic permutational symmetry ($\sigma_r = 3$) of ethane and needs to be considered in both reference calculation and λ-TI approach.

$$\Delta A^{\text{ref}}_{0,\boldsymbol{q}\to 1} = -\frac{1}{\beta}\ln\left\{\frac{\frac{1}{\sigma_r}\int_{-\pi}^{\pi} d\tau e^{-\beta V_1(\tau)}}{\int_{-\pi}^{\pi} d\tau e^{-\beta V_{0,\boldsymbol{q}}(\tau)}}\right\} = -\frac{1}{\beta}\ln\left\{\frac{\int_{-\pi}^{\pi} d\tau e^{-\beta V_1(\tau)}}{\int_{-\pi}^{\pi} d\tau e^{-\beta V_{0,\boldsymbol{q}}(\tau)}}\right\} - \frac{1}{\beta}\ln\left(\frac{1}{\sigma_r}\right) \tag{3.50}$$

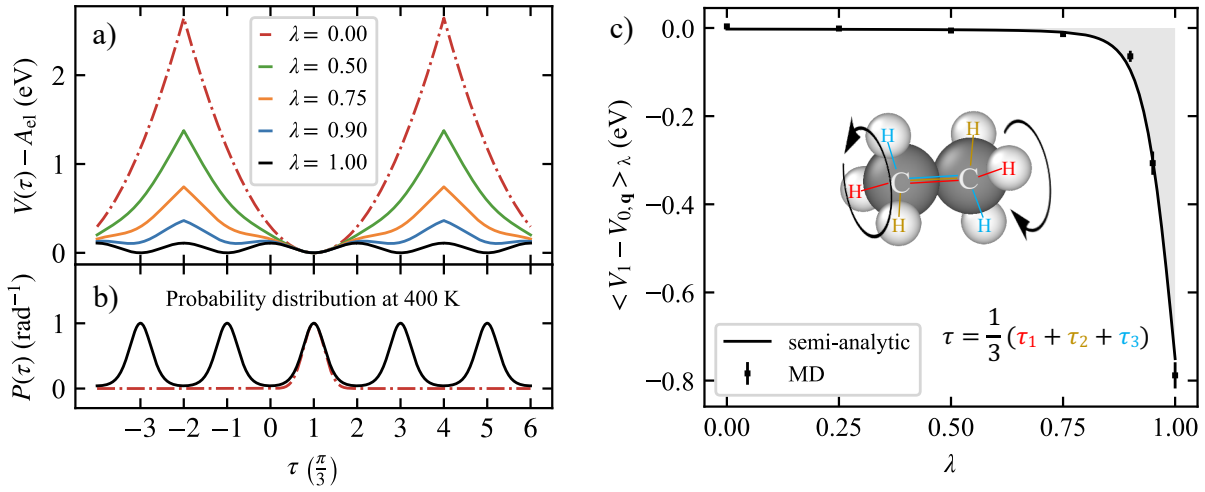


Figure 3.7.: Case study 3 on the hindered internal rotation of ethane. (a) Visualization of potential morphing due to the linear combination of Hamiltonians. (b) Semi-analytic probability densities for harmonic ($\lambda = 0$) and anharmonically interacting ($\lambda = 1$) system. (c) λ-TI: Comparison of semi-analytic solution (solid line) to averaged energy differences obtained by MD runs for seven values of λ at 400 K. Integration of the latter yields $\Delta A_{0,\boldsymbol{q}\to 1} + \frac{1}{\beta}\ln\left(\frac{1}{\sigma_r}\right)$. Reprinted with permission from [17]. Copyright 2021 American Chemical Society.

The harmonic potential is $V_{0,\boldsymbol{q}}(\tau) = V_{0,\boldsymbol{q}}(\tau_0) + \frac{1}{2}C(\tau - \tau_0)^2$ whereby the equilibrium value $\tau_0 = -1.047$ rad was determined using structural relaxation, the force constant $C = 0.508\,\text{eV rad}^{-2}$ was identified using a finite difference scheme at the DFT level of theory, and the term $V_{0,\boldsymbol{q}}(\tau_0) = V_{0,\boldsymbol{x}}(\boldsymbol{x}_0) = V_1(\boldsymbol{x}_0)$ is the ionic ground state energy. The anharmonic potential $V_1(\tau)$ has been evaluated using DFT on a grid defined in the range $-\pi \leq \tau \leq \pi$, with grid points separated by ~0.15 rad. Specifically, the value of V_1 for each grid point was determined by means of constrained relaxations with fixed τ. It can be seen in Figure 3.7a how the potential morphs with shifting values of λ so that rotations become energetically feasible. Furthermore, Figure 3.7b shows the probability density as a function of the torsion angle from which it is clear that only one out of three minima is sampled in MD driven by the harmonic potential. Moreover, it becomes clear that the discontinuity of the driving potential that emerges for $\lambda < 1$ is visited with a very low likelihood during our simulations and hence it does not pose a practical problem. The harmonic contribution to the free energy determined at 400 K is $A_{0,\boldsymbol{x}} - A_{el}(\boldsymbol{x}_0) = 762.9$ meV.

The free energy $\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}}$ representing the free energy difference between the system harmonic in internal coordinates (driven by the Hessian matrix $\underline{\boldsymbol{H}}^{\boldsymbol{q}}$) and the system harmonic in Cartesian coordinates (driven by the Hessian matrix $\underline{\boldsymbol{H}}^{\boldsymbol{x}}$) was computed as described in Section 3.1.2.1. For this purpose, MD simulations for five different regularly spaced values of λ have been performed with an integration step of 0.5 fs. The length of each simulation was 50 ps and an Andersen thermostat probability of 0.05 was used. The Hessian matrix $\underline{\boldsymbol{H}}^{\boldsymbol{x}}$ has been computed for the relaxed geometry. The internal coordinates $\boldsymbol{q}$ used to define the harmonic force field consisted of the geometric parameter τ from eq. (3.49), and the set of interatomic distances and angles generated as described in Section 3.1.2.5. The computed value of $\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}} = (-0.5 \pm 0.4)$ meV is not zero in contrast to the case studies with truly 1D potentials. Combined with $\Delta A^{\text{ref}}_{0,\boldsymbol{q}\to 1} = -4.1$ meV (see Table 3.1), the total ionic free energy $A_1 - A_{el}(\boldsymbol{x}_0)$ is 758.3 meV. Again, our classical reference value is in excellent agreement with quantum mechanical treatment as shown in Section 3.2.2.2. Hence, in our study the anharmonic contribution of a hindered rotation is noticeable already at a modest temperature but one order of magnitude smaller than that of a hindered translation.

DFT-based MD simulations were performed to determine $\Delta A_{0,\boldsymbol{q}\to 1}$ representing the free energy difference between the anharmonically interacting system (as described by DFT) and the system harmonic in internal coordinates, using an integration mesh with seven different values of λ. In addition to the five points of the regular grid used in the $\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}}$ calculations, two extra points ($\lambda = 0.90$ and 0.95) were added to improve the sampling in the region with rapidly changing $\langle V_1 - V_{0,\boldsymbol{q}}\rangle_\lambda$, see Figure 3.7c. The eqs. of motion were integrated with a step size of 0.5 fs. Due to the slow and infrequent internal rotations of CH_3 occurring when the harmonic potential is replaced by the full potential, the convergence of $\langle V_1 - V_{0,\boldsymbol{q}}\rangle_\lambda$ slowed down with increasing λ. We therefore used production periods of variable length which ranged between 15 ps ($\lambda = 0.0$) and 150 ps ($\lambda = 1.0$). A cubic box with an edge size of 15 Å containing one molecule was used in the simulations. As shown in Figure 3.7c, the $\langle V_1 - V_{0,\boldsymbol{q}}\rangle_\lambda$ computed from MD agree well with the reference data. The computed $\Delta A_{0,\boldsymbol{q}\to 1} = (-3.3 \pm 2.2)$ meV is therefore in a reasonable agreement with the reference value of −4.1 meV obtained from eq. (3.50) (see also Table 3.1).

Summary of the case studies. The simple models discussed in this section can be considered a test for the correctness of our implementation of λ-TI with internal coordinates. To this end, we compared our numerical data with the reference semi-analytical results and excellent agreement has been achieved. Furthermore, the three case studies showed that different types of degrees of freedom contribute differently to anharmonicity. From the study on the HF molecule, it can be learned that the anharmonic contribution of vibration in covalently bound systems is very small at ambient temperatures and it can be neglected in most cases. Our results for the argon dimer show that the harmonic approximation fails in the description of hindered translations (dissociation) and this failure is both qualitative and quantitative. For this reason the anharmonic correction to the free energy is large in this case. The hindered internal rotation of ethane contributes noticeably to anharmonicity already at 400 K but this contribution is one order of magnitude smaller than that of the hindered translation. For the adsorption study in Section 3.2.2.3 we shall keep in mind that the harmonic reference system can be chosen arbitrarily, as explained in Section 3.1.2.4.

3.2.2.2. Comparison of Classical Free Energies of Model Systems with Quantum Mechanical Solutions

The main purpose of the model systems presented in Section 3.2.2.1 is to demonstrate the validity of our λ-TI approach and to quantify anharmonic contributions from different degrees of freedom. Besides λ-TI, the classical anharmonic correction for these systems was also determined alternatively by semi-analytic solutions. The validation of our approach then consisted of a direct comparison of its results against the semi-analytical references. Albeit the following comparison with the exact quantum solutions is irrelevant in this respect, it is nevertheless interesting from a conceptual point of view.

HF molecule. To assess the quantum mechanical solution for the HF molecule, we adapted the commonly used approximation that individual rotational and vibrational degrees of freedom can be treated as independent motions. With the rotational and vibrational partition functions (q_{rot} and q_{vib}) the free energy difference between the harmonic reference (0) and the DFT-based (1) system can be computed via eq (3.51).

$$\Delta A_{0,\boldsymbol{x}\to 1} = \Delta A_{0,\boldsymbol{q}\to 1} = -k_\text{B}T \ln \frac{q_{\text{rot},1}\, q_{\text{vib},1}}{q_{\text{rot},0}\, q_{\text{vib},0}}. \tag{3.51}$$

Note that the systems 0 and 1 do not differ in translational degrees of freedom. Hence, the corresponding partition functions cancel in the expression in eq. (3.51). The quantum mechanical expressions for the vibrational partition functions of the harmonic oscillator ($q_{\text{vib},0}$) and the anharmonically interacting

system ($q_{\mathrm{vib},1}$) are obtained via eqs. (3.52)[87] with the harmonic frequency $\omega = 72.8$ THz, the degeneracy g_j of quantum state j, and ε_j being the energy obtained by numerically solving the 1D Schrödinger eq. (3.53) by the Fourier grid Hamiltonian method.[206]

$$q_{\mathrm{vib},0} = \frac{\exp\left(-\frac{\hbar\omega}{2k_\mathrm{B}T}\right)}{1-\exp\left(-\frac{\hbar\omega}{k_\mathrm{B}T}\right)} \quad \text{and} \quad q_{\mathrm{vib},1} = \sum_j g_j \exp\left(-\beta\varepsilon_j\right) \tag{3.52}$$

$$-\frac{\hbar^2}{2\mu}\frac{d^2}{dR^2}\Psi + V_1(R)\Psi = \varepsilon\Psi \tag{3.53}$$

A reduced mass of $\mu = 4.30 \cdot 10^{-27}$ kg is obtained for the HF molecule using the mass of tritium instead of protium, as described in Section 3.2.1.1. The full potential (Figure 3.4a) was expressed in the form of the inverse power expansion $V_1 = \sum_{n=0}^{6} c_n R^{-n}$ whereby the expansion coefficients c_n (in eVÅ^{n+1}: -17.63, 177.91, -496.64, 660.52, -481.49, 187.48, -30.06) have been obtained by fitting the potential energies evaluated on a regular 1D grid with 129 grid points. With $T = 2000$ K we obtain $\frac{q_{\mathrm{vib},1}}{q_{\mathrm{vib},0}} = 1.0157$. For the rotational degrees of freedom, we adapted the commonly used rigid rotor approximation for our system with two equivalent rotational axes to eq. (3.54)[87] with the symmetry number $\sigma = 1$ and the moment of inertia, $I_\lambda = \mu R_{\mathrm{eq},\lambda}^2$, where $R_{\mathrm{eq},\lambda}$ is the equilibrium distance between the H and F atoms evaluated for the harmonic ($\lambda = 0$) and the anharmonic ($\lambda = 1$) system.

$$q_{\mathrm{rot},\lambda} = \frac{2\pi k_\mathrm{B} T I_\lambda}{\sigma\hbar^2} \tag{3.54}$$

Since the expressions for $q_{\mathrm{rot},0}$ and $q_{\mathrm{rot},1}$ differ only in I, the fraction $\frac{q_{\mathrm{rot},1}}{q_{\mathrm{rot},0}}$ reduces to $\frac{R_{\mathrm{eq},1}^2}{R_{\mathrm{eq},0}^2}$. Due to the symmetry of the problem we express the partition function in spherical coordinates, see eq. (3.55), and substitute $R_{\mathrm{eq},0}$ and $R_{\mathrm{eq},1}$ by the average interatomic distances 0.9447 Å and 0.9581 Å, respectively, to find $\frac{q_{\mathrm{rot},1}}{q_{\mathrm{rot},0}} = 1.0286$.

$$R_{\mathrm{eq},\lambda} = \frac{\int dR R^3 \exp\left(-\beta V_\lambda(R)\right)}{\int dR R^2 \exp\left(-\beta V_\lambda(R)\right)} \tag{3.55}$$

Combining all terms we obtain $\Delta A_{0,\boldsymbol{x}\to 1} = \Delta A_{0,\boldsymbol{q}\to 1} = -0.0075$ eV which is in excellent agreement with the reference value of −7.6 meV reported in Section 3.2.2.1. Finally, we note that a very similar result (−8.2 meV) is obtained when the mass of protium instead of tritium is used in the calculation, which follows from the fact that the system is close to the classical limit and the classical canonical partition function is independent of the mass of the atoms.

Ar dimer. Applying the quantum treatment for the HF molecule to the Ar dimer we find the values 1.9183 and 1.5700 for $\frac{q_{\mathrm{vib},1}}{q_{\mathrm{vib},0}}$ and $\frac{q_{\mathrm{rot},1}}{q_{\mathrm{rot},0}}$, respectively. Here, the reduced mass of $\mu = 3.32 \cdot 10^{-26}$ kg, the harmonic vibrational frequency of the relaxed dimer of $\omega = 1.37$ THz, and the symmetry number $\sigma = 2$ were used, while the expansion coefficients c_n (in eVÅ^{n+1}) obtained by fitting the full potential V_1 in Figure 3.5a by the inverse power expansion were -7.4, 192.8, -2053.2, 11478.2, -35386.0, 56684.2 and -36523.4 for $n \in [0, 6]$, respectively. As in the case of HF, the 1D Schrödinger eq. (3.53) was solved numerically by the Fourier grid Hamiltonian method with a regular grid of 129 points. These values yield $\Delta A_{0,\boldsymbol{x}\to 1} = \Delta A_{0,\boldsymbol{q}\to 1} = -0.0285$ eV in excellent agreement with the reference value of −27.4 meV reported in Section 3.2.2.1.

Ethane molecule. Assuming that the individual degrees of freedom are independent, the free energy difference between the model harmonic in internal coordinates (0,$\boldsymbol{q}$) and the hindered rotor model (1) of ethane can be computed according to eq. (3.56) where $q_{\mathrm{hr},1}$ is the partition function of the 1D hindered rotor and $q'_{\mathrm{vib},0,\boldsymbol{q}}$ is that for its harmonic approximation.[87]

$$\Delta A_{0,\boldsymbol{q}\to 1} = -k_{\mathrm{B}}T \ln \frac{q_{\mathrm{hr},1}}{q'_{\mathrm{vib},0,\boldsymbol{q}}} \tag{3.56}$$

The overall translations and rotations as well as all vibrational degrees of freedom apart from the specific vibrational mode corresponding to the frustrated internal rotation in the harmonic model are assumed to be identical in the models 0 and 1, hence their contributions to the partition functions cancel. The quantum mechanical expressions for the involved partition functions are given in eqs. (3.57)[87] where $\sigma_r = 3$ is the symmetry number of the internal rotation, g_j is the degeneracy of the quantum state j with eigenenergy ε_j, and $\frac{\omega_{\boldsymbol{q}}}{2\pi} = \frac{\sigma_r}{2\pi}\sqrt{\frac{W}{2I}} = 5.04\,\mathrm{THz}$ (for example, see McClurg et al.[175]) is the harmonic frequency of the hindered rotor defined via the potential energy barrier for the internal rotation ($W = 0.110\,\mathrm{eV}$) and the effective moment of inertia I.

$$q'_{\mathrm{vib},0,\boldsymbol{q}} = \frac{\exp\left(-\frac{\hbar\omega_{\boldsymbol{q}}}{2k_{\mathrm{B}}T}\right)}{1-\exp\left(-\frac{\hbar\omega_{\boldsymbol{q}}}{k_{\mathrm{B}}T}\right)} \quad \text{and} \quad q_{\mathrm{hr},1} = \frac{1}{\sigma_r}\sum_j g_j \exp\left(-\beta\varepsilon_j\right) \tag{3.57}$$

Neglecting Coriolis contributions, the latter is computed using eq. (3.58)[176] where G_1 and G_2 indicate the two rotating groups of atoms (here both G_1 and G_2 represent the two methyl groups of ethane), d_i is the distance between the rotation axis (in this case passing through the C–C bond) and the atom i with mass m_i that belongs to the group G_j.

$$\frac{1}{I} = \frac{1}{I_{G_1}} + \frac{1}{I_{G_2}} \quad \text{with} \quad I_{G_j} = \sum_{i\in G_j} m_i d_i^2 \tag{3.58}$$

The effective moment of inertia ($I = 7.89\cdot 10^{-47}\,\mathrm{kg\,m^2}$) is then used in eq. (3.54) for the internal rotation involving two methyl groups. The eigenenergies ε_j are obtained by numerically solving the periodic 1D Schrödinger eq. (3.59).

$$-\frac{\hbar^2}{2\mu}\frac{d^2}{d\tau^2}\Psi + V_1(\tau)\Psi = \varepsilon\Psi \tag{3.59}$$

To this end a method described in Chapter 7 of Kittel[86] (section 'Solution of the central equation') was adapted whereby the full potential V_1 (Figure 3.7a) was expressed in the form of a Fourier series with 9 terms obtained by fitting the potential energies evaluated on a 1D grid. With $T = 400\,\mathrm{K}$ we obtained $\frac{q_{\mathrm{hr},1}}{q'_{\mathrm{vib},0,\boldsymbol{q}}} = 1.119$ yielding in turn the result $\Delta A_{0,\boldsymbol{q}\to 1} = -0.0039\,\mathrm{eV}$ which is in excellent agreement with our reference value of −4.1 meV.

The model described in Hill (1986)[207] represents an approximation to the treatment described above, in particular, only a single dominant term of the Fourier expansion $V_1 \approx \frac{W}{2}(1-\cos(\sigma_r\tau))$. Since the actual V_1 is very well represented by a single cosine function, the result of $\Delta A_{0,\boldsymbol{q}\to 1} = -0.0044\,\mathrm{eV}$ obtained by the quantum mechanical treatment of a potential of the latter form is very similar to the less approximate result discussed above.

Finally, the free energy contribution due to the transformation from the potential harmonic in Cartesian coordinates to that harmonic in internal coordinates can be computed by eq. (3.60) where the quantum mechanical expression for $q'_{\mathrm{vib},0,\boldsymbol{q}}$ is given by eq. (3.57).[87]

$$\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}} = -k_\mathrm{B}T \ln \frac{q'_{\mathrm{vib},0,\boldsymbol{q}}}{q'_{\mathrm{vib},0,\boldsymbol{x}}} \quad \text{with} \quad q'_{\mathrm{vib},0,\boldsymbol{x}} = \frac{\exp\left(-\frac{\hbar\omega_{\boldsymbol{x}}}{2k_\mathrm{B}T}\right)}{1-\exp\left(-\frac{\hbar\omega_{\boldsymbol{x}}}{k_\mathrm{B}T}\right)} \tag{3.60}$$

The harmonic frequency $\frac{\omega_{\boldsymbol{x}}}{2\pi} = 5.12\,\mathrm{THz}$ is computed by diagonalization of the dynamical matrix defined in Cartesian coordinates and selection of the vibrational mode corresponding to the frustrated rotation. The value of $\frac{q_{\mathrm{vib},0,\boldsymbol{q}}}{q'_{\mathrm{vib},0,\boldsymbol{x}}} = 1.015$ is leading to $\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}} = -0.0005\,\mathrm{eV}$, which is basically identical with the classical value obtained via λ-TI that was reported in Section 3.2.2.1.

3.2.2.3. Adsorption of Ar and N_2 in Zeolite H-SSZ-13

The relevance of anharmonicity for adsorption processes in heterogeneous catalysis is often ignored or at best estimated by the more or less coarse techniques mentioned in Section 3.2.1. Our presented λ-TI approach can be used to calculate anharmonic corrections for adsorption processes on the DFT level of theory using MD. Here, we apply the concept to study two adsorbates (N_2, Ar) in zeolite H-SSZ-13 at 200 K. We will see that these two systems exhibit qualitatively different behavior with respect to the stability of the complexes formed upon adsorption. Experimental Gibbs free energies of adsorption ($\Delta_\mathrm{ads}G$) at 200 K were deduced from data reported by Barrer and Davies.[182] The linear regression of the experimental results, shown in the Appendix, resulted in values of −24 meV and −45 meV for Ar and N_2, respectively. According to our calculations, the harmonic approximation evaluated at the PBE-D2 level tends to overestimate these values with +26 meV and −24 meV for the adsorption free energies of Ar and N_2, respectively. As shown in our analysis of the model system Ar_2 in Section 3.2.2.1, the harmonic approximation restricts the motion of atoms that are not bound by chemical bonds leading to a significant overestimation of the free energy. Similarly, the harmonic approximation restricts the adsorbate@substrate systems to configurations corresponding to the adsorption complexes in which the adsorbate is attached to the proton of the Brønsted acid site (see insets in Figure 3.8).

Just as in the model Ar_2 case, the shift from harmonic ($V_{0,\boldsymbol{x}}$) to full DFT (V_1) interaction causes significant changes in the behavior of the adsorbate. These qualitative changes can be seen best from radial distribution functions (RDF) for the H−Ar and H−N pairs in Figure 3.8 computed using MD. The peak of the RDF for Ar adsorption in Figure 3.8 computed using the potential harmonic in Cartesian coordinates $\boldsymbol{x}$ (blue line) is relatively narrow, approximately Gaussian-shaped, and centered at the distance found in the relaxed structure (2.4 Å). The same peak computed from MD driven by the full PBE-D2 potential (green line) is broadened, its maximum becomes less pronounced and is shifted towards longer interatomic separation (2.7 Å). This broadening of the RDF and the non-vanishing probability at long range clearly signify a higher entropy of the system driven by the anharmonic model compared to its harmonic counterpart. The specific interaction between N_2 and the Brønsted site of the zeolite is more significant than in the case of Ar adsorption (164.0 meV vs. 36.8 meV) and the adsorption complex in Figure 3.8 similar to the potential energy minimum indeed exists with a greater likelihood at our target temperature of 200 K. Nevertheless, the complex is not stable all the time – rather it collapses and is recreated frequently during the MD. The shift from the potential harmonic in Cartesian coordinates to the potential harmonic in internal coordinates and eventually to the full interaction potential is accompanied by a broadening of the RDF peaks and a change in their shape from relatively narrow Gaussians to broader asymmetric bands. Furthermore, in consequence of temporal

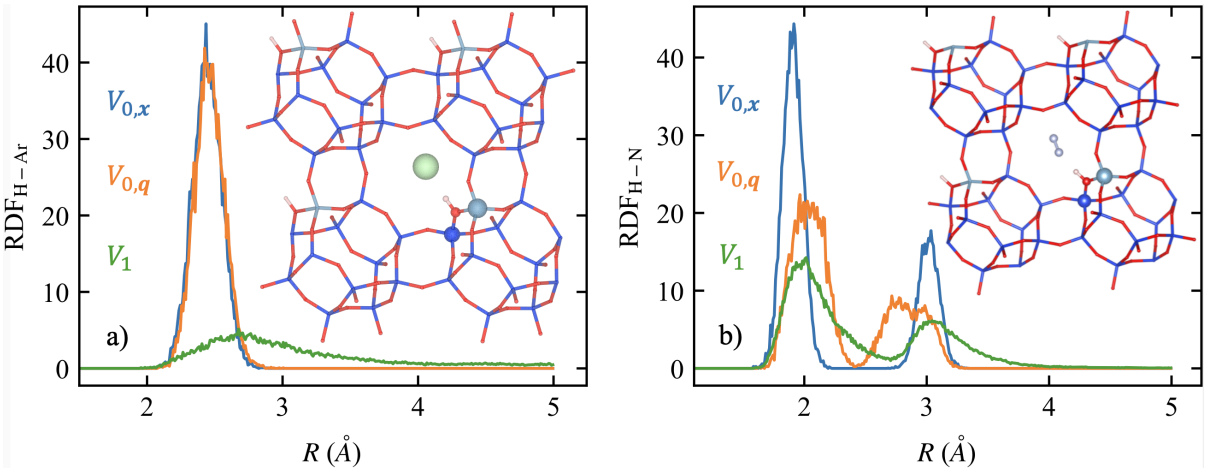


Figure 3.8.: Partial radial distribution functions (RDF) versus interatomic distance R computed for the pairs of proton and adsorbate atoms for Ar in zeolite H-SSZ-13 (left) and N_2 in H-SSZ-13 (right). The results were obtained from MD at 200 K driven by the full PBE-D2 potential (V_1), by the potential harmonic in Cartesian coordinates ($V_{0,\boldsymbol{x}}$), and by the potential harmonic in internal coordinates ($V_{0,\boldsymbol{q}}$). The depicted zeolite structure (color code similar to Figure 3.2) is represented by a stick model while the central adsorbate atoms (Ar and N_2) and the framework atoms at the acidic site are shown as spheres. Reprinted with permission from [17]. Copyright 2021 American Chemical Society.

collapses of the adsorption complex and in contrast to the harmonic model, the RDF has non-vanishing values in the region between the two maxima. Despite the distinctions, the structure of the RDF and the positions of its maxima remain similar.

In our MD calculations, the three systems, namely the clean substrate, the adsorbate in the substrate and the free adsorbate in gas phase were treated separately. A consistent simulation setting has been used in the simulations of all three systems. In particular, all simulations were performed under periodic boundary conditions with the identical simulation cell corresponding to the primitive unit cell of zeolite H-SSZ-13 (see Figure A.2 in the Appendix), and the basis set was fixed by setting the plane-wave cutoff to 400 eV. A Si to Al ratio of 11:1 has been used in the structural model of zeolite H-SSZ-13 where the proton occupied position O(1),[50] which belongs to the most populated H sittings in this zeolite.[208] All MD simulations pertaining to the λ-TI method have been performed for seven different values of λ (0.00, 0.25, 0.50, 0.75, 0.90, 0.95, 1.00) plus several additional points defined when the $\langle V_1 - V_{0,\boldsymbol{q}}\rangle_\lambda$ term changed rapidly with λ (see below). As we have shown above, the anharmonicity of a single bond is very small even at temperatures as high as 2000 K. We therefore neglect this contribution to the free energy of the N_2 molecule in gas phase, the N–N vibration is treated fully harmonically. Furthermore, rotational degrees of freedom of free N_2 are treated using the rigid rotor approximation.[85] The translational degrees of freedom of free adsorbate particles (isolated Ar and N_2) were treated using the ideal gas approximation,[85] whereby a reference pressure of 101.325 kPa has been considered. Moreover, the concentration of the adsorbed state is one adsorbate per unit cell (corresponding to one adsorbate per Brønsted acid site), where adsorbate-adsorbate interactions over a distance of around 9 Å are negligible. The computed total free energies (including electronic energies) of the adsorbate particles are −0.286 eV (Ar) and −16.873 eV (N_2). The Gibbs free energy of adsorption is computed from eq. (3.61) with the anharmonic Helmholtz free energies of the adsorbate in the substrate ($A_{1,\mathrm{A@S}}$) and the clean substrate ($A_{1,\mathrm{S}}$) as well as the Gibbs free energy of the gas phase adsorbate ($G_{0,\mathrm{A}}$) determined using the harmonic oscillator, rigid rotor, and ideal gas approximations.

$$\Delta_{\mathrm{ads}}G = A_{1,\mathrm{A@S}} - A_{1,\mathrm{S}} - G_{0,\mathrm{A}} \tag{3.61}$$

We note that the cancellation of the pV terms for the A@S and S systems is a reasonable assumption in eq. (3.61) since the substrate is a crystalline material and the concentration of adsorbate in the A@S system is relatively low – therefore volume changes of the substrate due to adsorption should be small.

For our λ-TI calculations the internal coordinates of the clean substrate were chosen as described in Section 3.1.2.5, whereby the proton position was described by four special internal coordinates: the OH bond length, the bond angles H–O–Al and H–O–Si, and the initially smallest torsion H–O–Al–O (the atoms involved in definitions of these coordinates are shown in Figure 3.2). To avoid numerical problems in simulations with $\lambda = 0$, all eigenvalues of the Hessian matrix have been increased to 1 eV Å^{-2} if they were lower than this limit (see Section 3.1.2.4). The harmonic free energy contribution $A_{0,\boldsymbol{x}}$ computed using the modified Hessian matrix ($\underline{\boldsymbol{H}}^{\boldsymbol{x}}$) is −286.968 eV, which is to be compared to −287.077 eV determined for the unmodified $\underline{\boldsymbol{H}}^{\boldsymbol{x}}$. The length of all MD simulations was 500 ps. The integration step used in the inexpensive force field to force field calculations was 0.5 fs while a larger step of 1 fs was used in the DFT calculations. As expected, the computed value of $\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}}$ is very small ((−5 ± 0) meV), which can be attributed to the low temperature (indeed, $\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}}$ tends to zero with decreasing temperature) and the limited structural variation allowed by the framework structure with all atoms being connected by strong covalent bonds. The term $\Delta A_{0,\boldsymbol{q}\to 1}$, on the other hand, takes a significantly greater value of (−99 ± 0) meV. Plugging all free energy contributions into eq. (3.6), the computed free energy of the substrate is $A_1 = -287.072$ eV. Comparing this value with the harmonic free energy determined using the original unmodified $\underline{\boldsymbol{H}}^{\boldsymbol{x}}$ (−287.077 eV) we recognize that the anharmonicity of the substrate at our target temperature is essentially negligible.

For the interacting system of Ar and zeolite, the set of internal coordinates used for the clean substrate (see above) was extended by the CN coordinates (eq. (3.29)) defined for all Ar–Si and Ar–Al pairs. The eigenvalues of $\underline{\boldsymbol{H}}^{\boldsymbol{x}}$ for this system have been increased to 2 eV Å^{-2} if they were lower than this limit (see Section 3.1.2.4). The harmonic free energy contribution $A_{0,\boldsymbol{x}}$ computed using the modified $\underline{\boldsymbol{H}}^{\boldsymbol{x}}$ is −287.102 eV, which is to be compared to −287.337 eV determined for the unmodified $\underline{\boldsymbol{H}}^{\boldsymbol{x}}$. The setting of the MD simulations for the interacting system was identical to that used for the clean substrate. An extra integration point $\lambda = 0.99$ was used in the calculation of $\Delta A_{0,\boldsymbol{q}\to 1}$. As in the case of the clean zeolite, the computed value of $\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}}$ is very small and barely statistically significant ((−2 ± 0) meV), while the term $\Delta A_{0,\boldsymbol{q}\to 1}$ is as large as (−271 ± 1) meV. Combining the free energy contributions, eq. (3.6) yields $A_1 = -287.375$ eV, which is to be compared with the harmonic free energy of −287.337 eV determined using the original unmodified $\underline{\boldsymbol{H}}^{\boldsymbol{x}}$. Hence, in line with our expectation, the anharmonicity lowers the free energy of the adsorbed state, which originates from the transformation of a part of the vibrations involving Ar into hindered translations, thus increasing the entropy of Ar adsorbed in the zeolite. Using eq. (3.61) with the free energies determined for the clean substrate, free Ar in gas phase and the adsorbed state (see Table 3.2), we arrive at $\Delta_{\text{ads}}G = (-17 \pm 2)$ meV, which is by 43 meV lower compared to the prediction made using the harmonic approximation (see Table 3.3). Remarkably, our prediction using λ-TI is very close to the experimental value reported by Barrer and Davies[182] (although we admit that this level of agreement is fortuitous to some extent). The internal coordinates chosen for the interacting system of N_2 and zeolite consisted of those used for the clean substrate (see above) extended by the N–N distance and the CN coordinates (eq. (3.29)) defined for all N–Si and N–Al pairs. All computed contributions to A_1 are listed in Table 3.2. The Hessian matrix eigenspectrum has been modified so that no force constant was lower than 2 eV Å^{-2} (see Section 3.1.2.4). For this modified harmonic reference model, a harmonic free energy contribution of −303.736 eV has been determined, while the value of $A_{0,\boldsymbol{x}}$ for the unmodified $\underline{\boldsymbol{H}}^{\boldsymbol{x}}$ is −303.974 eV. Owing to a relatively large variation of the term $\langle V_1 - V_{0,\boldsymbol{q}}\rangle_\lambda$ with λ, three extra integration points ($\lambda = 0.990$, 0.995 and 0.999) have been used in the $\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}}$ calculation. As evident from Figure 3.8b, the coordinate transformation in the harmonic model causes significant changes in the behavior of adsorbed N_2. Consequently, the computed $\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}}$ of (−22 ± 1) meV

Table 3.2.: Contributions to the free energies A_1 in eV at 200 K and at a reference pressure of 101.325 kPa. $A_{0,\boldsymbol{x}}$ is the free energy of the harmonic reference system where we added contributions from the ideal gas and rigid rotor approximations when applicable. The mass of tritium was used for H atoms in all calculations. Reprinted with permission from [17]. Copyright 2021 American Chemical Society.

System	$A_{0,\boldsymbol{x}}$	+	$\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}}$	+	$\Delta A_{0,\boldsymbol{q}\to 1}$	=	A_1
Ar	-0.286^a	+	0^b	+	0	=	-0.286
N_2	-16.873^c	+	0^b	+	0^d	=	-16.873
chabazite	-286.968	+	-0.005 ± 0.000	+	-0.099 ± 0.001	=	-287.072 ± 0.004
Ar@chabazite	-287.102	+	-0.002 ± 0.000	+	-0.271 ± 0.001	=	-287.375 ± 0.002
N_2@chabazite	-303.736	+	-0.022 ± 0.001	+	-0.245 ± 0.005	=	-304.003 ± 0.005

[a] ideal gas approximation. [b] $A_{0,\boldsymbol{x}} = A_{0,\boldsymbol{q}}$. [c] harmonic oscillator/ideal gas/linear rigid rotor approximations. [d] negligible.

Table 3.3.: Comparison of simulation results to experimental values for adsorption Gibbs free energies at 200 K in meV. Our calculations employing the λ-TI approach and conventional harmonic approximation were performed on the DFT level of theory (PBE-D2). Reprinted with permission from [17]. Copyright 2021 American Chemical Society.

Adsorbate	Experiment[182]	λ-TI approach	harmonic approx.
Ar	-24	-17 ± 2	$+26$
N_2	-45	-58 ± 5	-24

represents a significant contribution to the anharmonic part of the adsorption free energy. Indeed, a transformation to a suitable set of coordinates can allow to capture a part of anharmonicity even within a harmonic model. Finally, the computed value of $\Delta A_{0,\boldsymbol{q}\to 1}$ is (-245 ± 5) meV and the total free energy of the adsorbate is (-304.003 ± 0.005) eV. Comparing the latter result with the value of -303.974 eV predicted by the unmodified harmonic model it is obvious that the temporary collapse of the adsorption complex observed in an anharmonically interacting model leads to a decrease (via increased entropy) of free energy by 29 meV. Consequently, the Gibbs free energy of adsorption computed using λ-TI ((-58 ± 5) meV) is lower than that obtained using the harmonic approximation. The λ-TI value is also slightly closer to the experimental value of -45 meV although the improvement with respect to the harmonic approximation is less significant than for the adsorption of Ar. Clearly, a careful choice of the density functional approximation and dispersion correction method would be needed to further improve the accuracy of adsorption free energy calculations.

3.2.2.4. Classical compared to Quantum Harmonic Approximation for Adsorption Problems

At very low temperatures, the free energies of matter are dominated by quantum effects which are not considered in classical theory. Deviations between quantum and classical descriptions vanish with increasing temperature as they approach the same value. In this work we use classical DFT-based MD and λ-TI from the classical harmonic approximation to the interacting system to determine classical anharmonic corrections. While the anharmonic correction from classical MD is always classical, we are free to choose between classical and quantum harmonic approximation. For conceptual clarity we decided to use the classical harmonic approximation.

We now show that for the adsorption free energies at the temperature considered in our adsorption studies (200 K), there is essentially no significant difference between the free energy contributions of classical and quantum harmonic approximation. Quantum harmonic vibrational contributions to the

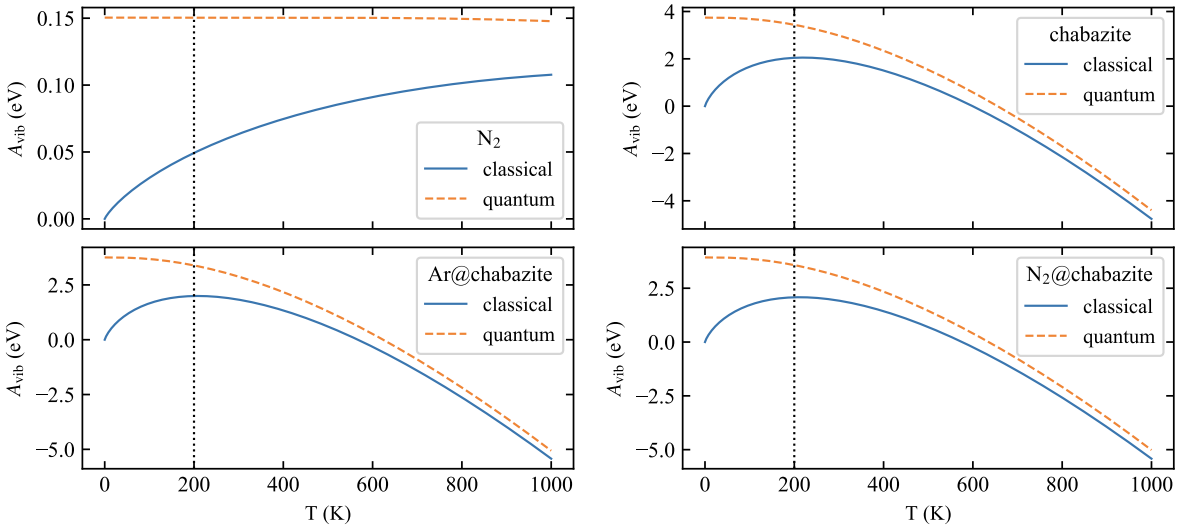


Figure 3.9.: Harmonic vibrational free energy contribution to individual components (molecular N_2, empty chabazite, adsorbed Ar, adsorbed N_2) needed for the calculation of adsorption free energies. Consistent with our MD simulations at 200 K, the mass of tritium was used for H atoms. Reprinted with permission from [17]. Copyright 2021 American Chemical Society.

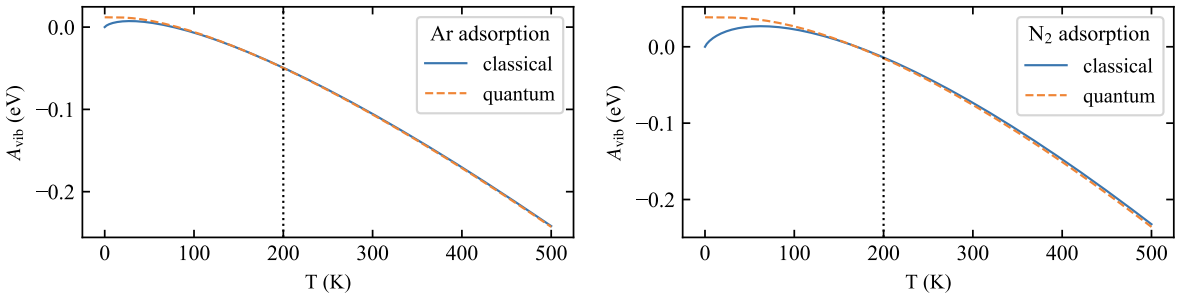


Figure 3.10.: Harmonic free energy contribution to the adsorption free energy of Ar and N_2 in chabazite zeolite. Consistent with our MD simulations at 200 K, the mass of tritium was used for H atoms. Reprinted with permission from [17]. Copyright 2021 American Chemical Society.

free energy were computed using the thermochemistry routines of the atomic simulation environment (ASE)[209] based on eqs. of Chapter 10 in Cramer (2004).[85] The classical vibrational contribution to the free energy of a system with N_{vib} degrees of freedom (number of harmonic frequencies) is computed by eq. (3.62), also compare eq. (3.2).

$$A_{\text{vib}} = -k_{\text{B}}T \sum_{i=1}^{N_{\text{vib}}} \ln \frac{k_{\text{B}}T}{\hbar\omega_i} \tag{3.62}$$

The harmonic vibrational free energy contributions to the individual components of the adsorption free energy as a function of temperature are depicted in Figure 3.9. The harmonic vibrational free energy contribution to the adsorption free energy as a function of temperature is shown in Figure 3.10. Classical and quantum theory for individual components approach each other rather slowly with increasing temperature. However, the significant difference of classical and quantum description above 100 K converges to negligible values (<2 meV at 200 K) for the adsorption free energy itself. In consequence, the classical limit is adequate for the adsorption studies discussed in this work.

3.2.3. Intermediate Summary and Conclusions

We presented thermodynamic λ-path integration (λ-TI) as a tool to compute anharmonic corrections to the free energies of molecular and periodically extended systems. Using translationally and rotationally invariant curvilinear coordinates we overcame the significant limitation of the λ-TI approach traditionally formulated in Cartesian coordinates. This newly developed method was demonstrated in three case studies and then applied to study adsorption energies of N_2 and Ar in the acidic chabazite zeolite H-SSZ-13. Hindered rotations and translations (dissociation) which are poorly described by the harmonic approximation were identified to cause significant anharmonic contributions to the free energy (translation more than rotation) while the anharmonic contributions of vibrations of covalent bonds are rather negligible. In comparison with experimental data from literature we observed that the anharmonic correction improves the accuracy of predicted adsorption free energies.

In principle, other popular simulation methods, such as the Blue Moon ensemble approach by Ciccotti et al.[104,105] or Umbrella sampling[106–108] could be employed to determine the adsorption free energy. These methods, however, require the sampling of the configuration space over a continuous reaction coordinate, which is extremely impractical for adsorption problems. In particular, the structural model of the substrate would have to be large with a wide vacuum gap so as to allow both the simulation of the sorbate adsorbed in bulk (unperturbed by surface effects) and the simulation of the desorbed sorbate unperturbed by interactions with the substrate. Hence, the presented λ-TI method requiring simulations of only initial and final states without the necessity to sample over the adsorption path is perfectly suited to study adsorption free energies or individual rate determining steps in multistep reaction cascades and could thus become a very useful tool for computational catalysis.

Another MD based method that is used in the literature to access the free energy of molecules and extended systems is the quasi-harmonic approximation (QHA).[210] In the QHA the harmonic oscillator model is used with renormalized vibrational frequencies determined from MD via Fourier transformation of the velocity autocorrelation function (VAF). Such an approach, recently used for instance, to study ethanol adsorption in H-ZSM-5[173] solves a part but not all problems associated with the harmonic approximation. The frequency renormalization considers thermal effects related to the thermal expansion of bonds. Furthermore, the frequencies emerging from MD are anharmonic and therefore their use in free energy calculations can account for a part of the vibrational anharmonicity. Nevertheless, we note that the latter are used in connection with the expressions derived for the harmonic oscillator, which is not rigorously correct for a general anharmonic case. Even worse, the QHA method cannot account for non-vibrational aperiodic degrees of freedom like hindered translations with no or negligible contribution to the VAF. Finally, the need for accurate accounting of the time correlation imposes a serious constraint on the simulation setting. The stochastic thermostats such as Andersen[98] or Langevin[211,212] cannot be used since the stochastic collisions quickly destroy any time correlations between atomic positions. The deterministic thermostats such as Nosé-Hoover[213,214] or Nosé-Hoover chains[215] must be adjusted with great care because the contributions from fictitious degrees of freedom of the thermostat can bias the computed VDOS. The VAF is therefore often computed in the microcanonical regime, which however, imposes a limitation on the length of the trajectory because the temperature can drift away from the desired value. Also, the need for an accurate VDOS implies that a relatively small integration step must be used in a QHA calculation, which in turn leads to a less efficient exploration of remote parts of the accessible configurational space. The λ-TI technique presented in this work is free of all these limitations.

3.2.4. Intermediate Outlook

Considering the focus of recent studies on adsorption[172,216,217] and on the degree of anharmonicity in solid materials,[218] our contribution advances the field of computational catalysis tremendously: the necessity of accurate free energy calculations beyond the static approach has already been well documented for reaction energetics[109,111,113,219,220] and our simulation method will enable to achieve a similar level of accuracy also for adsorption problems, which are almost always important parts of catalytic cycles. Most interesting is also the application of the λ-TI method to transition states with constrained ensembles, which will be discussed in the following section. Several strategies potentially improving the quality and effectiveness of sampling or accuracy of calculations are collected in Section 3.3.4.

3.3. Anharmonic Corrections to Free Energy Barriers

Preface. For the calculation of anharmonic contributions to free energy barriers, thermodynamic λ-path integration (λ-TI) from harmonic reference force fields to density functional theory (DFT) is presented as alternative to an established ξ-TI method, that depends on the reaction coordinate ξ. Two strategies using λ-TI are compared and benchmarked against the ξ-TI method. We find good agreement between λ-TI and ξ-TI, also for a nucleophilic substitution reaction, with significant anharmonicity. Its independence from ξ lets the λ-TI method excel in calculations of arbitrary states in a free energy landscape.

3.3.1. Overview and Simulation Details

The accurate quantification of anharmonicity on the molecular scale, defined as deviation from the harmonic oscillator approximation, represents one of the grandest challenges in computational heterogeneous catalysis with particular relevance for supported metal nanoparticles, confinement effects in zeolites and solid-liquid interfaces.[12,217] Disappointingly, despite the known significance of anharmonicity – oftentimes estimated on the basis of density functional theory (DFT) and molecular dynamics (MD) – relatively few studies provide more than a rough estimate for entropy and free energy contributions. In comparison, anharmonicity is relatively well explored for solid materials.[140,218,221–225]

Yet, free energies for anharmonic systems of catalytic relevance can be computed by several procedures. Typically, these procedures either determine absolute free energies of a single state or free energy differences between different states. In the former case, many procedures aim to improve the accuracy for stationary states on the potential energy surface. Such static approaches are beneficial in that the result for the single state can be referenced to any other state. Approximations to the partition function of hindered translators and hindered internal rotors for molecules and adsorbates have been developed to this end.[75,174–176,226] Not only require such assumptions and simplifications prior knowledge about the mobility of the system components, their reliability also strongly depends on the thermodynamic conditions. For instance, a comparison of predictions by the harmonic, hindered and free translator approximation with the complete potential energy sampling approach has shown that none of the simple approximations is reliable for modeling adsorbates over a wide range of temperatures.[76] Other ideas improve the harmonic oscillator approximation by adapting the force constants to the temperature of interest using quasi-harmonic or effective harmonic approximations.[127,222,227–229] In a rather involved correction to the static approach, anharmonic vibrational partition functions have been determined by solving one-dimensional Schrödinger eqs. defined for the potential energies explicitly sampled

along individual vibrational eigenmodes expressed in terms of internal coordinates.[134,135,144,177,230] Thermodynamic integration (TI)[83,101,116] along temperature and Hamiltonian-based λ-paths[117–122] has been used to study crystalline systems[123,138,140,231] and phase transitions.[125–130] In particular, TI from Debye models to fully interacting crystalline systems[132,133] is conceptually similar to the method reported in this work, where we present a more generally applicable TI method from harmonic reference systems to fully interacting DFT-based systems. Methods for single states aside, the second type of procedures computes free energy differences between different states using TI or similar schemes based on multistage sampling. Methods of this kind, like Umbrella sampling[106–108] or the Blue Moon ensemble technique by Ciccotti et al.,[104,105,157] can be employed routinely to calculate free energy barriers under consideration of anharmonicity as simulated by DFT-based MD. While these methods are designed for the investigation of rare events, such as the event of passing through a TS of a chemical reaction or phase transition, they also depend on the reaction coordinate ξ connecting reactant and TS. Thereby obtained free energy differences can hardly be referenced to arbitrary other states. For example, the calculation of a free energy barrier for a reaction near the end of an extended reaction cascade relative to the state of the initial reactants would require integration along all reactions from the beginning to that barrier.

To calculate anharmonic corrections to the harmonic approximation for adsorption free energies, we have reported a method harnessing thermodynamic λ-path integration (λ-TI) of DFT-based MD using curvilinear coordinates, see Sections 3.1 and 3.2. Instead of integrating along the reaction coordinate ξ, λ-TI integrates along a progressive parameter λ that couples the harmonic reference force field with DFT. This independence from ξ makes the λ-TI method also interesting when applied to reaction barriers, for example, selected barriers in an expanded reaction network could be corrected by anharmonic contributions relative to arbitrary reactant states. To this end, we now demonstrate how our λ-TI method can be applied to the constraint ensemble of a TS. We furthermore present λ-TI as an alternative to an established ξ-TI method by Ciccotti et al.,[104,105,157] which integrates over ξ (multistage Blue Moon ensemble technique) and thus depends on the reaction path. For simple reactions where ξ is known precisely, the λ-TI and ξ-TI methods are directly compared.

3.3.1.1. Simulation Details

The simulations have been performed as described in Sections 3.2.1.1 and 3.2.1.2. Transition structures have been optimized using automated relaxed potential energy surface scans.[97] The eigenvalues of the Hessian matrices were increased to 2 eV Å^{-2} if they were lower than this limit.

3.3.2. Results and Discussion

For a selection of reactions, we compute free energy barriers including anharmonic contributions using our λ-TI method in its strategic variants λ-TI_a and λ-TI_b benchmarked against ξ-TI. The selection includes the degenerate unimolecular rotamerization of ethane, the second order degenerate nucleophilic substitution of MeCl by a Cl^- anion and a proton transfer in the acidic chabazite zeolite H-SSZ-13. A tabular summary is provided at the end of this section, in particular, all free energy contributions from λ-TI and the corresponding harmonic approximations are shown in Table 3.4. The free energy differences $\Delta A_{\xi_{\text{ref,R}}\rightarrow\xi^*}$ between the constrained reactant and TS computed by ξ-TI and λ-TI are compared in Table 3.5. The results for the term $\Delta A_{\text{free}_\text{R}\rightarrow\xi_{\text{ref,R}}}$ that connects the constrained and free reactant states are shown in Table 3.6. With the addition of $\Delta A_{\text{free}_\text{R}\rightarrow\xi_{\text{ref,R}}}$ from Table 3.6 to ξ-TI and λ-TI in Table 3.5 we arrive at predictions by the complete ξ-TI method and by the λ-TI_a strategy, respectively, for the free energy barriers shown in Table 3.7, where they are compared to the direct predictions by the λ-TI_b

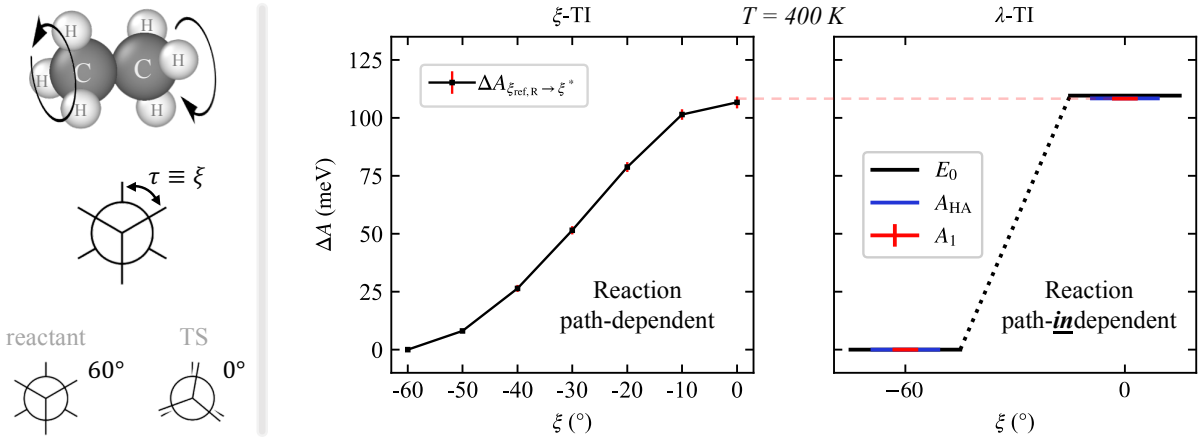


Figure 3.11.: The reaction coordinate ξ for the rotamerization of ethane (left) is described by the dihedral angle τ. The term $\Delta A_{\xi_{\mathrm{ref,R}} \to \xi^*} = A_1(0°) - A_1(-60°)$ at 400 K obtained by ξ-TI (middle) and λ-TI (right) is compared to the harm. approx. ΔA_{HA}.

strategy and the common harmonic approximation. Note that free energy contributions from the rigid rotor and free translator approximations have not been added, because they are independent of the method applied.

3.3.2.1. Internal rotation of ethane

The reaction coordinate ξ for the rotamerization of ethane can be described by the dihedral angle τ between two hydrogen atoms at opposite methyl groups of ethane (see Fischer projection in Figure 3.11). The term $\Delta A_{\xi_{\mathrm{ref,R}} \to \xi^*}$ at 400 K was determined to (0.1067 ± 0.0026) meV by ξ-TI from 7 Blue Moon ensembles (black dots in Figure 3.11) along the reaction path including the staggered reactant and ecliptic TS. The same quantity was determined to (0.1083 ± 0.0002) meV by λ-TI from Blue Moon ensembles for only the reactant and TS. Each Blue Moon ensemble was obtained from DFT-based MD production runs of 110 ps (ξ-TI) and 75 ps (λ-TI) carried out with an integration step of 0.5 fs. Using the same integration step, the numerical force field to force field transformation for the term $A_{0,\boldsymbol{x} \to 0,\boldsymbol{q}}$ was evaluated on production runs with a length of 250 ps. The term $\Delta A_{\mathrm{free_R} \to \xi_{\mathrm{ref,R}}}$ was determined to the value in Table 3.5 from an unconstrained MD simulation with a length of 75 ps with the probability distribution in Figure 3.14. To exclude the spontaneous reaction in the simulation, τ was restricted to values of $\tau \in [0°, 60°]$.

We see excellent agreement between ξ-TI and λ-TI in Figure 3.11. The constrained reactant ($\xi_{\mathrm{ref,R}}$) and TS (ξ^*) are barely anharmonic (see Table 3.4), therefore the harmonic approximation predicts a very similar value (0.1084 eV), see Table 3.5. A more significant but still little deviation from the harmonic approximation is observed for the free reactant ($\mathrm{free_R}$) with $A_1 - A_{\mathrm{HA}} = (-0.0060 \pm 0.0008)$ eV. When compared to the free energy barrier referenced to the free reactant state (without free translator and rigid rotor approximations), we see that the harmonic approximation predicts an increase of the barrier, while ξ-TI and λ-$\mathrm{TI_a}$ predict a decrease. The prediction by λ-$\mathrm{TI_b}$ is off by 0.0055 eV compared to the prediction by λ-$\mathrm{TI_a}$. This deviation can be explained by the worse convergence of the energy differences between harmonic and anharmonic system for the unconstrained reactant state, see Figure A.3 in the Appendix, which makes the λ-$\mathrm{TI_a}$ strategy much more accurate, faster and thus attractive in simulations. We conclude that anharmonicity is not important for the internal rotation of ethane, especially due to the similar interactions (covalent bonds) and degrees of freedom for reactant and TS.

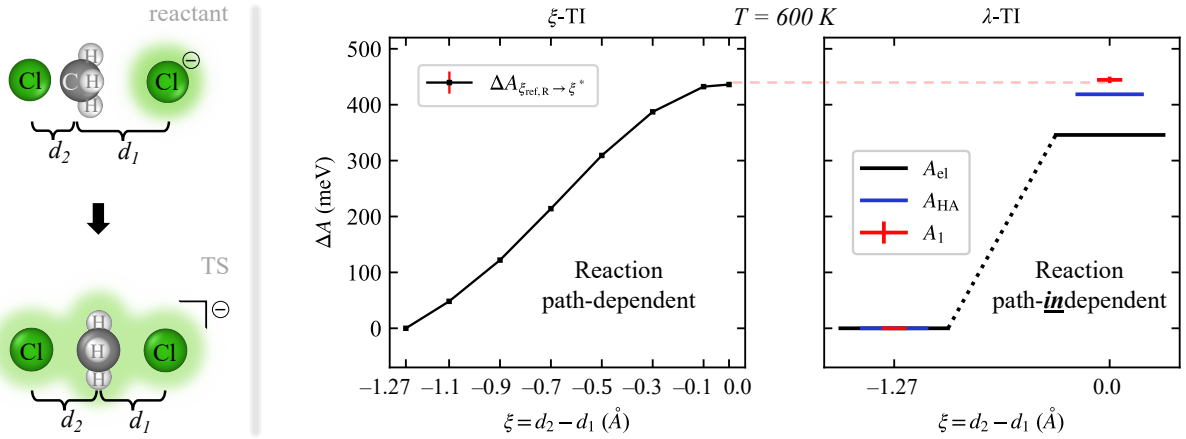


Figure 3.12.: The reaction coordinate ξ for the nucleophilic substitution of MeCl by Cl^- (left) is described by a linear combination of the interatomic distances d_1 and d_2. The term $\Delta A_{\xi_{ref,R} \to \xi^*} = A_1(0\,\text{Å}) - A_1(-1.27\,\text{Å})$ at 600 K obtained by ξ-TI (middle) and λ-TI (right) is compared to the harmonic approximation ΔA_{HA}.

3.3.2.2. Nucleophilic substitution of MeCl

The reaction coordinate ξ for the nucleophilic substitution of MeCl with a Cl^- anion can be described by a linear combination of the interatomic distances d_1 and d_2 involved in the formation and cleavage of bonds ($\xi = d_2 - d_1$), see Figure 3.12. The term $\Delta A_{\xi_{ref,R} \to \xi^*}$ at 600 K was determined to (0.4361 ± 0.0045) meV by ξ-TI from 8 Blue Moon ensembles (black dots in Figure 3.12) along the reaction path including the reactant and TS. The same quantity was determined to (0.4444 ± 0.0058) meV by λ-TI from Blue Moon ensembles for only the reactant and TS. Each Blue Moon ensemble was obtained from DFT-based MD production runs of 100 ps (ξ-TI) and 80 ps (λ-TI) carried out with an integration step of 1.0 fs and a constant background charge was set to neutralize the unit cell. Using the same integration step, the term $A_{0,\boldsymbol{x} \to 0,\boldsymbol{q}}$ was evaluated on the standard grid employing production runs with a length of 500 ps. The term $\Delta A_{\text{free}_R \to \xi_{ref,R}}$ was determined to the value in Table 3.5 from an unconstrained MD simulation with a length of 60 ps with the probability distribution in Figure 3.14.

We see excellent agreement between ξ-TI and λ-TI in Figure 3.12. In contrast to the barely anharmonic TS (ξ^*), anharmonic contributions are significant for the constrained reactant ($\xi_{ref,R}$) and even more for the free reactant (free_R), see Table 3.4. Thus, the energetic stabilization of the reactant state leads to the increase of the barrier by roughly 0.02 eV to 0.03 eV compared to the prediction by the harmonic approximation, see Table 3.5. When compared to the free energy barrier referenced to the free reactant state (without free translator and rigid rotor approximations), we see that that deviation from the harmonic approximation doubles (compare Tables 3.5 and 3.7). Again, we note that the prediction by λ-TI_b is a little off but its standard error still overlaps with that of λ-TI_a. We conclude that anharmonicity, caused by the weakly bound reactant state, is very important for this reaction.

3.3.2.3. Proton transfer in zeolite H-SSZ-13

The reaction coordinate ξ for a proton transfer in zeolite H-SSZ-13 can be described by a linear combination of the interatomic distances d_1 and d_2 involved in the formation and cleavage of bonds, similar to the substitution reaction in Section 3.3.2.2, see Figure 3.13. The term $\Delta A_{\xi_{ref,R} \to \xi^*}$ at 600 K was determined to (0.6723 ± 0.0111) meV by ξ-TI from 5 Blue Moon ensembles (black dots in Figure 3.13) along the reaction path including the reactant and TS. The same quantity was determined to (0.6788 ± 0.0173) meV by λ-TI from Blue Moon ensembles for only the reactant and TS. DFT-based MD production runs of

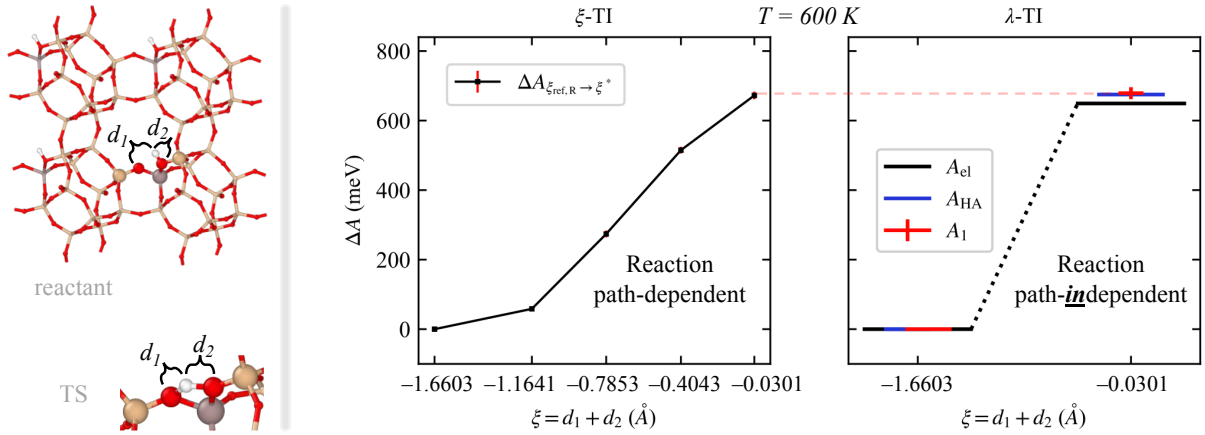


Figure 3.13.: The reaction coordinate ξ for a proton transfer in zeolite H-SSZ-13 (left) is described by a linear combination of the interatomic distances d_1 and d_2. The term $\Delta A_{\xi_{\mathrm{ref,R}} \to \xi^*} = A_1(-0.0301\,\text{Å}) - A_1(-1.6603\,\text{Å})$ at 600 K obtained by ξ-TI (middle) and λ-TI (right) is compared to the harmonic approximation ΔA_{HA}.

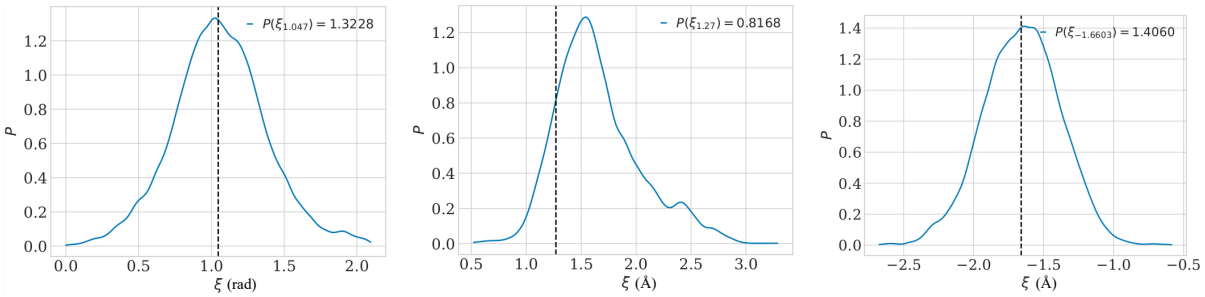


Figure 3.14.: Probability distribution (PD) of the unconstrained reactant as a function of the reaction coordinate ξ. Left: PD for ethane. Middle: PD for MeCl co-adsorbed with the Cl^- nucleophile. Right: PD for the proton transfer in zeolite H-SSZ-13.

50 ps (ξ-TI and λ-TI) were carried out with an integration step of 1.0 fs. Using the same integration step, the term $A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}}$ was evaluated on production runs with a length of 500 ps. The term $\Delta A_{\mathrm{free_R}\to\xi_{\mathrm{ref,R}}}$ was determined to the value in Table 3.5 from an unconstrained MD simulation with a length of 50 ps with the probability distribution in Figure 3.14.

We see excellent agreement between ξ-TI and λ-TI in Figure 3.13. In contrast to ethane, the individual anharmonic contributions are very high (see Table 3.4), but mostly cancel in the calculation of the free energy difference, see Tables 3.5 and 3.7. When compared to the free energy barrier referenced to the free reactant state (without free translator and rigid rotor approximations), we note that the prediction by λ-$\mathrm{TI_b}$ is again a little off, and although its standard error still overlaps with that of λ-$\mathrm{TI_a}$, the role of anharmonicity is not clear compared to the harmonic approximation due to a comparatively large statistical error of the λ-$\mathrm{TI_b}$ strategy. This same tendency observed in all cases discussed above – and most pronounced for the rotamerization of ethane – leads us to the conclusion that the convergence of the λ-$\mathrm{TI_b}$ strategy is not as good as that of the λ-$\mathrm{TI_a}$ strategy. When referenced to the free reactant state, the predictions by ξ-TI and λ-$\mathrm{TI_a}$ stay roughly the same ($\Delta A_{\mathrm{free_R}\to\xi_{\mathrm{ref,R}}}$ is less than 0.01 eV), while the harmonic approximation predicts a barrier increased by more than 0.02 eV. Anharmonicity is thus more important than it seems in Figure 3.13, because the harmonic approximation deviates more when referenced to the free reactant state. Due to anharmonic contributions the overall barrier is by 0.02 eV to 0.03 eV lower than predicted by the harmonic approximation. In this case, we attribute the most significant source of anharmonicity to the reactant state that is more anharmonic than the TS. This

conclusion is also in line with the observation that the proton is less mobile in the TS where it is trapped between the two oxygen atoms.

3.3.3. Intermediate Summary and Conclusions

We presented thermodynamic λ-path integration (λ-TI) as a tool to compute anharmonic corrections to free energy barriers. Using translationally and rotationally invariant curvilinear internal coordinates under periodic boundary conditions the approach was demonstrated in three examples comprising a simple unimolecular reaction, a bimolecular substitution reaction and a proton transfer inside a zeolite. Excellent agreement was observed in a comparison of our λ-TI method to the established ξ-TI method by Ciccotti et al.,[104,105,157] showing that under identical conditions (all reactants and products with constraints in one simulation box) both methods yield the same result within statistical accuracy. Among the two strategic variants λ-TI$_a$ and λ-TI$_b$, we recommend the first one, because it converges faster with generally smaller statistical errors. Furthermore, we quantified anharmonicity relative to the common harmonic approximation. The rotamerization of ethane with prevailing covalent bonds showed little anharmonicity, which was to be expected from previous insight into the relevance of anharmonic contributions for different interactions and degrees of freedom (dissociation contributes more than hindered rotation, vibrations of covalent bonds contribute the least).[17] Significant anharmonic contributions were observed for the nucleophilic substitution reaction of MeCl, which we attributed to the weakly bound nature of the reactant state. Anharmonic contributions to the barrier of the proton transfer in zeolite H-SSZ-13 are comparatively small, because the individually significant anharmonic contributions to reactant and TS largely cancel themselves. Nevertheless, a small anharmonic contribution to the barrier can be attributed to the difference between constrained and free reactant state, where the harmonic approximation deviates more for the latter. Henceforth, the λ-TI$_a$ method can be seen as a valuable computational tool to compute anharmonic corrections to free energy barriers independent of the reaction coordinate ξ.

3.3.4. Intermediate Outlook

Several strategies potentially improving the quality and effectiveness of sampling and the accuracy of the calculations will be addressed in future work. First, a more sophisticated choice of the reference Hessian matrix renormalized so as to effectively describe as large part of the anharmonicity of the system as possible would minimize the $\Delta A_{0,\boldsymbol{q}\to 1}$ term computed via λ-TI, which would consequently result in faster convergence of the $\langle V_1 - V_{0,\boldsymbol{q}}\rangle_\lambda$ term. In this respect, the ideas beyond the quasi-harmonic approach[227] appear to be very promising. Second, although internal coordinates are quite practical and often represent chemical degrees of freedom quite well, they always are very specific to the studied system. The performance of the λ-TI method could possibly benefit from the use of sophisticated, more universally applicable coordinates such as descriptors commonly used in the machine learning (ML) community. A short communication on the properties of universal coordinates like smooth overlap of atomic positions (SOAP),[166] and Coulomb-type descriptors[167] is given in Section 3.4, however, the use of these coordinates should be explored in detail with regard to their applicability and sampling efficiency. Third, the simulation times necessary to obtain well converged integrands $\langle V_1 - V_{0,\boldsymbol{q}}\rangle_\lambda$ are very long even for the relatively simple adsorption and reaction problems discussed in this work. This problem can be alleviated by combining the MD simulations with ML algorithms. Among the most promising approaches in this regard is the seamless ML method of Jinnouchi et al.[232] available in VASP as of version 6.3, which is able to learn the underlying model on-the-fly during a DFT-based MD run whereby the ML predictions gradually take control over the simulation while maintaining a

controllable error estimation of the ML model. Such a treatment has been shown to accelerate the MD by a large factor. Finally, Bennett's acceptance ratio[233,234] should also be explored as an alternative way of analyzing the data generated for λ-TI.

Table 3.4.: Contributions to the free energy A_1 in eV obtained using λ-TI compared to the classical standard harmonic approximation A_{HA}.

System	$A_{0,\boldsymbol{x}}$	$\Delta A_{0,\boldsymbol{x}\to 0,\boldsymbol{q}}$	$\Delta A_{0,\boldsymbol{q}\to 1}$	A_1	A_{HA}
ethane ($\mathrm{free_R}$)	−39.7684	−0.0003 ± 0.0000	−0.0371 ± 0.0008	−39.8058 ± 0.0008	−39.7998
ethane ($\xi_{\mathrm{ref,R}}$)	−39.8029	−0.0004 ± 0.0000	−0.0022 ± 0.0001	−39.8055 ± 0.0001	−39.8029
ethane (ξ^*)	−39.6921	−0.0004 ± 0.0000	−0.0047 ± 0.0001	−39.6972 ± 0.0001	−39.6945
$MeCl_2^-$ ($\mathrm{free_R}$)	−27.8951	−0.0019 ± 0.0000	−0.2189 ± 0.0106	−28.1159 ± 0.0106	−28.0492
$MeCl_2^-$ ($\xi_{\mathrm{ref,R}}$)	−27.8871	−0.0016 ± 0.0000	−0.1586 ± 0.0044	−28.0473 ± 0.0044	−28.0287
$MeCl_2^-$ (ξ^*)	−27.5840	−0.0018 ± 0.0000	−0.0171 ± 0.0014	−27.6029 ± 0.0014	−27.6101
H^+ transfer ($\mathrm{free_R}$)	−288.4145	−0.0175 ± 0.0003	−0.5811 ± 0.0121	−289.0131 ± 0.0124	−288.9314
H^+ transfer ($\xi_{\mathrm{ref,R}}$)	−288.4140	−0.0160 ± 0.0002	−0.5551 ± 0.0085	−288.9851 ± 0.0087	−288.9085
H^+ transfer (ξ^*)	−287.7650	−0.0148 ± 0.0002	−0.5265 ± 0.0084	−288.3063 ± 0.0086	−288.2336

Table 3.5.: Free energy difference $\Delta A_{\xi_{\mathrm{ref,R}}\to\xi^*}$ between the constrained reactant and TS computed by ξ-TI and λ-TI compared to the corresponding harmonic approximation ΔA_{HA}.

System	T / K	ξ-TI / eV	λ-TI / eV	ΔA_{HA} / eV
ethane	400	0.1067 ± 0.0026	0.1083 ± 0.0002	0.1084
$MeCl_2^-$	600	0.4361 ± 0.0045	0.4444 ± 0.0058	0.4186
H^+ transfer	600	0.6723 ± 0.0111	0.6788 ± 0.0173	0.6749

Table 3.6.: Data for the calculation of $\Delta A_{\mathrm{free_R}\to\xi_{\mathrm{ref,R}}}$ according to eq. (3.22).

System	T / K	$\left\langle\lvert\dot{\xi}^*\rvert\right\rangle$ / $\mathrm{s^{-1}}$	$P(\xi_{\mathrm{ref,R}})$	$\Delta A_{\mathrm{free_R}\to\xi_{\mathrm{ref,R}}}$ / eV
ethane	400	$1.4653 \cdot 10^{13}$	1.3228	−0.0052
$MeCl_2^-$	600	$1.1058 \cdot 10^{13}$	0.8168	0.0526
H^+ transfer	600	$1.9883 \cdot 10^{13}$	1.4060	−0.0058

Table 3.7.: Free energy barriers $\Delta A_{\mathrm{free_R}\to\xi^*}$ computed by the TI methods compared to the corresponding harmonic approximation ΔA_{HA}. Free energy contributions from the rigid rotor and free translator approximations have not been added.

System	T / K	ξ-TI / eV	λ-$\mathrm{TI_a}$ / eV	λ-$\mathrm{TI_b}$ / eV	ΔA_{HA} / eV
ethane	400	0.1015 ± 0.0026	0.1031 ± 0.0002	0.1086 ± 0.0009	0.1053
$MeCl_2^-$	600	0.4887 ± 0.0045	0.4970 ± 0.0058	0.5130 ± 0.0120	0.4391
H^+ transfer	600	0.6665 ± 0.0111	0.6730 ± 0.0173	0.7068 ± 0.0210	0.6978

3.4. Outlook on Universal Coordinates

Although we decided to use sets of redundant internal coordinates for the thermodynamic λ-path integration (λ-TI) in this work, any other type of geometric descriptor for the coordinates $\boldsymbol{q}$ should work as well, as long as the coordinate space meets certain criteria. Most importantly, a descriptor must be invariant with respect to transformations that preserve the target property, that is energy in our case. Moreover, the descriptor should be continuous, unique and non-degenerate. Ideally, the descriptor should also be efficient in the calculation of itself and its Cartesian derivatives. Under consideration of Occam's razor[235] and Solomonoff's general theory of inductive inference[236] the descriptor should also follow the minimum descriptor length principle[237–239] to avoid the curse of dimensionality (empty space phenomenon[240]).[168] Lastly, the generated phase space after transformation of the Hessian matrix from Cartesian coordinates to the new descriptor should not contradict chemical intuition. In other words, the new phase space should have similarities to either the phase space of the harmonic approximation or the anharmonically interacting system.

Studying the adsorption of N_2 in H-SSZ-13, we have noticed that the transformation of the harmonic approximation from Cartesian to well-chosen internal coordinates captures a part of the anharmonicity at basically no computational cost (see Figure 3.8). In principle, the harmonic approximation can be generalized to any coordinate system that passes the requirements mentioned above. Recent advances in the development of geometric descriptors for machine learning (ML) applications have led to a rich variety of readily available coordinate spaces.[168] Although the perfect universal descriptor has not been found yet, for the application in our λ-TI method there exist promising candidates among two classes of descriptors – Coulomb-type descriptors and descriptors based on spherical harmonics. The herein briefly tested descriptors, namely Coulomb, Sine, Ewald and SOAP, were generated by the software library DScribe (v1.2).[241] An algorithm for the harmonic force field to force field transformation was contributed to the open-source atomic simulation environment (ASE).[209] Cartesian derivatives were obtained numerically using a central finite difference scheme. The Coulomb matrix defined in eq. (3.63) is a simple global descriptor which mimics the electrostatic interaction between nuclei.[167]

$$M_{ij}^{\text{Coulomb}} = \begin{cases} 0.5 Z_i^{2.4} & \text{for} \quad i = j \\ \frac{Z_i Z_j}{R_{ij}} & \text{for} \quad i \neq j \end{cases} \tag{3.63}$$

Diagonal elements can be seen as the interaction of an atom with itself and are essentially a polynomial fit of the atomic energies to the nuclear charge Z_i, while off-diagonal elements represent the Coulomb repulsion (at a distance R_{ij}) between nuclei i and j.[241] Unlike the Coulomb matrix, which is not defined for systems in periodic boundary conditions, the Sine matrix given by eq. (3.64) and the Ewald sum matrix in eq. (3.65) have been designed as periodic counterparts to the Coulomb matrix.[242]

$$M_{ij}^{\text{Sine}} = \begin{cases} 0.5 Z_i^{2.4} & \text{for} \quad i = j \\ \frac{Z_i Z_j}{\left| \mathbf{B} \sum_{k=x,y,z} \hat{\boldsymbol{e}}_k \sin^2\left(\pi \mathbf{B}^{-1} (R_i - R_j)\right) \right|} & \text{for} \quad i \neq j \end{cases} \tag{3.64}$$

The matrix $\mathbf{B}$ is formed by the lattice vectors and $\hat{\boldsymbol{e}}_k$ are the Cartesian unit vectors. Although this functional form has no physical interpretation, it captures some of the properties of the Coulomb interaction, such as the periodicity of the crystal lattice and infinite energy for overlapping atoms.[241,242] Figure 3.15 comprises the DFT-based PES of the HF molecule as well as the typical harmonic approximation in Cartesian coordinates transformed to the other descriptor spaces. Due to their linear dependence in this simple example, we observe that the harmonic approximation in Cartesian coordinates is identical when transformed to internal coordinates, namely the interatomic distance. The Coulomb-type descriptors Coulomb, Sine and Ewald form a group with similar properties. In fact, the Ewald descriptor, which is

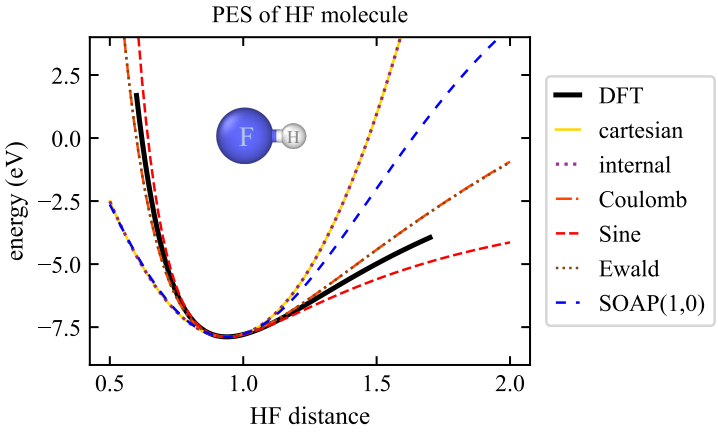


Figure 3.15.: Comparison of the DFT-based potential energy surface for the HF molecule to the harmonic approximation in Cartesian coordinates and the potentials obtained by transformation to a variety of descriptor spaces including the Coulomb, Sine and Ewald sum matrices as well as smooth overlap of atomic positions (SOAP) with $n_{\max} = 1$ and $l_{\max} = 0$.

the logical extension of the Coulomb matrix for periodic boundary conditions, overlaps exactly with the Coulomb matrix in the shown subinterval of the cubic unit cell with an edge length of 15 Å. The Ewald summation technique is used for ϕ_{ij}^{self} with a uniform neutralizing background charge ϕ_{ij}^{bg}, where V is the unit cell volume, α is the screening parameter for converging the sum and N is the number of atoms in the unit cell.[241,242] Note that the used definition by Himanen et al.[241] differs from the original work by Faber et al.[242]

$$\phi_{ij}^{\text{self}} + \phi_{ij}^{\text{bg}} = -\frac{\pi}{2V\alpha^2} Z_i Z_j \quad \forall\, i \neq j \quad \text{with} \quad \alpha = \sqrt{\pi}\left(\frac{N}{V^2}\right)^{1/6} \tag{3.65}$$

The SOAP descriptor (smooth overlap of atomic positions)[166,243] is a more complex function based on spherical harmonics, see eq. (3.66), where $\rho^Z(r) = \sum_i^{|Z_i|} \exp\left(-0.5\sigma^{-2}|r - R_i|^2\right)$ is the Gaussian smoothened atomic density for atoms with atomic number Z.

$$p_{nn'l}^{Z_1 Z_2} = \pi\sqrt{\frac{8}{2l+1}} \sum_m c_{nlm}^{Z_1} c_{n'lm}^{Z_2} \quad \text{with} \quad c_{nlm}^{Z} = \int dV g_n(r) Y_{lm}(\theta,\phi) \rho^Z(r) \tag{3.66}$$

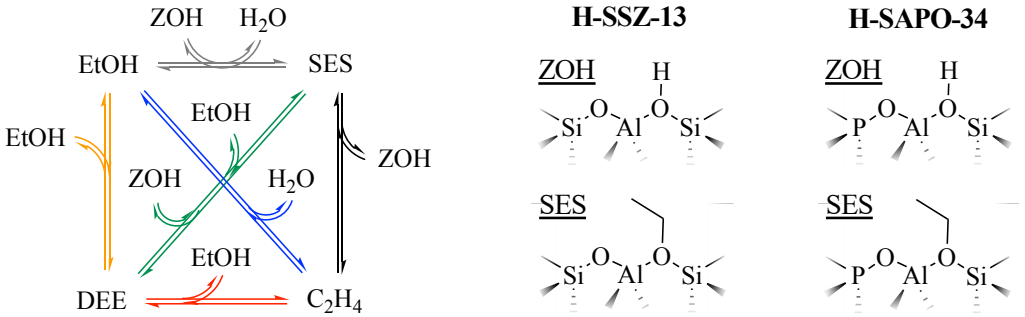
Moreover, the SOAP descriptor requires further hyper-parameters such as the number of (Gaussian-type) radial basis functions g_n with the indices n and n' for the different radial basis functions up to $n_{\max}$, the angular degree l of spherical harmonics Y_{lm} up to $l_{\max}$ or the specification of a cutoff radius. Remarkably, a simple transformation of the coordinate system transforms the parabola of the Cartesian harmonic approximation into curves which fascinatingly resemble to overall trends of the DFT-based PES. Unfortunately, the SOAP descriptor tends to dramatically inflate feature space with higher choices of hyper-parameters thereby likely violating the minimum descriptor length principle. Furthermore, SOAP is not only invariant with respect to translation and rotation but also to the permutation of atoms of the same element. If not properly controlled by the eigenvalue spectrum of the Cartesian Hessian, this characteristic can lead to unintuitive phase spaces where molecules may unphysically juggle with atoms of the same element if they come close enough together. Obviously, these challenging complications need to undergo critical scrutiny until the capabilities and usefulness of universal descriptors for the λ-TI method has been (dis-)proven. Nevertheless, the exploratory and preliminary results in this section are meant to inspire and to convey the concept of using force fields obtained by transformation of the harmonic approximation from Cartesian coordinates to recently developed coordinate spaces. Future investigations will tell whether these universal descriptors can compete with the system specific internal coordinates used so far.

4. Olefin Production with Chabazite Catalysts

Scope. This chapter presents computational studies on ethanol (EtOH) dehydration as well as on the Ethanol-to-Olefins (ETO) and Methanol-to-Olefins (MTO) processes based on a bottom-up multiscale approach passing parameters from quantum mechanics up to ideal reactor models. Section 4.1 compares zeolite H-SSZ-13 to zeotype H-SAPO-34 for EtOH dehydration. All simulations related to H-SAPO-34 were tailored to contribute to a joint research venture with an experimental group. Corresponding parts are adapted and reprinted with permission from [244]. Copyright 2023 American Chemical Society. Section 4.2 explores reaction pathways for a significant part of the ETO olefin cycle in zeolite H-SSZ-13, including olefin ethylations up to hexene isomers and the corresponding cracking reactions. The corresponding contents including EtOH dehydration in zeolite H-SSZ-13 have been published and are adapted and reprinted with permission from [245]. Copyright 2022 The Royal Society of Chemistry (CC-BY). The work was published with the co-authors Philipp N. Plessow and Felix Studt in advisor roles, and Sarah Bernart, who located most transition structures as part of her master thesis. Section 4.3 investigates the effect of omnipresent impurities on the MTO initiation in a kinetic study using zeolite H-SSZ-13 as the catalyst. The corresponding contents have been published and are adapted and reprinted with permission from [246]. Copyright 2021 Springer Nature (CC-BY). The work was published with the co-authors Philipp N. Plessow and Felix Studt in advisor roles.

4.1. Ethanol Dehydration – Zeolite H-SSZ-13 versus Zeotype H-SAPO-34

Preface. In a comparison of zeolite H-SSZ-13 to zeotype H-SAPO-34, the formation of ethene and diethyl ether (DEE) from ethanol (EtOH) is studied in kinetic models with *ab initio* computed reaction barriers. Free energies are computed using density functional theory (DFT) and post-Hartree-Fock methods with a complete basis set extrapolation applied to a hierarchy of periodic and cluster models – entropy contributions are estimated using the rigid-rotor harmonic-oscillator approximation. The studied mechanism depicted in Scheme 4.1 is employed in batch and plug flow reactor simulations to



Scheme 4.1.: Mechanism for ethanol (EtOH) dehydration to diethyl ether (DEE) and ethene including the surface ethoxy species (SES). Despite different compositions, zeolite H-SSZ-13 and zeotype H-SAPO-34 appear in the same chabazite framework structure. Adapted and reprinted with permission from [245]. Copyright 2022 The Royal Society of Chemistry (CC-BY).

investigate EtOH conversion as well as the selectivity of DEE and ethene for different temperatures and as a function of the weight hourly space velocity (WHSV) in the plug flow reactor model.

4.1.1. Overview and Simulation Details

The zeolite-catalyzed dehydration of EtOH allows to produce ethene from potentially renewable feedstocks. Several studies have compared different zeolites and have demonstrated excellent activity in the dehydration reaction of EtOH to ethene,[247] and also further reactions to a complex mixture of short-chain hydrocarbons.[248,249] In the experimentally determined thermodynamic equilibrium between EtOH, diethyl ether (DEE) and ethene at atmospheric pressure the ratio of DEE to ethene is 1:1 at roughly 370 K, while above 450 K no significant amount of DEE is observed.[250] However, in many experimental studies the actually observed product ratio greatly varies with the overall reaction conditions,[251] such as ethanol pressure,[252] catalyst contact time, overall conversion and temperature. A particularly high sensitivity of the ethene selectivity to the catalyst contact time and temperature has been found in H-SAPO-34 in the narrow temperature interval from 458 K to 503 K.[253] Many other experimental works have been reported.[251] Tong used a diluted EtOH solution (20 %) over H-ZSM-5 and achieved 99 % conversion with 80 % ethene selectivity at 672 K, and 42 % conversion with 72 % ethene selectivity at 571 K.[254] Le Van Mao et al. observed 96 % conversion of EtOH solution (15 %) over H-ZSM-5 at 673 K with 49 % ethene selectivity.[255] Phillips and Datta applied EtOH in very low pressures and discovered that temperatures below 473 K are too low for EtOH dehydration leading to catalyst deactivation.[256] Pan and Li used modified H-ZSM-5 at 533 K and obtained 98 % conversion with 98 % to 99 % ethene selectivity.[257,258] Wang et al. fed 0.2 MPa EtOH, water and Ar carrier gas to H-SAPO-34 at 493 K to 593 K where the observed EtOH conversion was 90 % with an ethene selectivity of 99 % above 533 K, whereas DEE was generated at lower temperatures.[259] Zhou et al. used 0.1 MPa EtOH at 513 K over H-SAPO-11/H-ZSM-5 to achieve a conversion of 99 % with 99 % ethene selectivity.[260] Xin et al. used H-ZSM-5 at 1 atm and 473 K with an EtOH partial pressure of 19.8 kPa diluted in He. The observed conversion was around 70 % to 76 % with ethene selectivities ranging from 5 % to 17 % depending on the degree of dealumination and desilication of the catalyst tuning its acidity for controlled selectivity.[261] Phung et al. compared FER, MFI, MOR, BEA, Y, and USY catalysts using a feed of 7.9 % EtOH in N_2 carrier gas to observe over 70 % DEE at 473 K with MFI and BEA, 99.9 % ethene selectivity at 573 K with FER and FAU/USY (competition with DEE was not relevant under practical conditions). Confinement was concluded to explain the observations best. EtOH conversion occurred only above 400 K with almost full conversion at temperatures above 600 K. DEE was the main product at low temperatures and low conversion but decreased at 500 K.[262] Zhang et al. found that 90 % EtOH conversion and 90 % ethene selectivity can be achieved with zeolite catalysts or catalytically active Al_2O_3 but the latter requires higher temperatures by at least 100 K. Among zeolites, the required temperature for similar conversion and ethene selectivity can differ by up to 70 K.[247] In tests with heteropoly acids it was concluded that DEE and ethene are probably produced mostly through parallel routes rather than following a consecutive reaction scheme.[263,264] In a theoretical study using H-ZSM-5, however, Alexopoulos et al. identified the consecutive dimer-mediated etherification followed by ether decomposition to be energetically most favorable for ethanol dehydration to ethene at 500 K – closely followed by the ethoxide-mediated mechanism.[252] EtOH conversion and ethene selectivity thus strongly depend on the employed catalyst and the reaction conditions.[248]

Density functional theory (DFT) and (micro-)kinetic models are nowadays routinely used to complement experimental studies with insights that are otherwise hardly observable. In general, a careful choice of the simulation setup is advisable. For instance, Alexopoulos et al. have shown that a reaction path analysis of EtOH dehydration from free energy profiles should include the effect of reaction

conditions, feed concentrations and conversion levels. Among theoretical insights on the effect of the zeolite framework[265] on the EtOH dehydration in zeolite H-ZSM-5[252] and on the stabilization of dehydration transition states,[266] a simple scheme to predict dehydration barriers has emerged in the literature.[267]

Below we investigate the dehydration of EtOH in zeolite H-SSZ-13 compared to zeotype H-SAPO-34 using a hierarchical theoretical approach employing plane-wave DFT (PBE-D3) calculations to periodic structural models corrected by highly accurate post-Hartree-Fock (DLPNO-CCSD(T)) calculations on cluster models of the materials.[61] This way, the periodic DFT calculations take the steric implications of the entire zeolite framework into account while the high-level post-Hartree-Fock method yields accurate electronic energies.[268–270] The calculated reaction barriers are used to derive microkinetic parameters for batch and plug flow reactor models.

4.1.1.1. Simulation Details

A hexagonal unit cell with pre-optimized parameters and one Brønsted acid site per unit cell was used for zeolite H-SSZ-13 (a = 13.625 Å, c = 15.067 Å) and zeotype H-SAPO-34 (a = 13.875 Å, c = 15.017 Å). Structures were optimized employing periodic density functional theory (DFT) calculations with the dispersion-corrected PBE-D3[60,96] density functional (zero damping) using the VASP code with the standard PAWs[196–198] and an energy cutoff of 400 eV. A convergence criterion of 0.01 eV Å^{-1} was applied to geometry optimizations. The Brillouin zone was sampled at the Γ-point only[200] using Gaussian smearing with a width of 0.1 eV. Coordinates for all optimized structures of zeolite H-SSZ-13, and also the locations of bulkier transition structures within the zeolite cavities, can be found in Amsler et al.[245] Transition structures were optimized using automated relaxed potential energy surface scans[97] and the existence of one imaginary mode connecting the correct minima of the reaction was confirmed. Entropic contributions to the free energy barriers have been calculated using the harmonic approximation at the respective temperatures and at a reference pressure of 1 bar. Vibrational frequencies were derived from the partial Hessian matrix computed using a central finite difference scheme including only the adsorbate, the acid site and its adjacent Al- and Si-atom. Moreover, vibrational frequencies below 12 cm^{-1} were raised to this value because they can lead to inaccurate entropies otherwise.[21] To obtain accurate reaction barriers,[61] corrections of the electronic energies according to eqs. (4.1) and (4.2) were applied to H-SSZ-13 using a hierarchical cluster approach[68–73] comprising periodic models and 46T cluster models saturated by hydrogen atoms,[268] using highly accurate domain-based local pair natural orbital coupled cluster (DLPNO-CCSD(T))[90,271,272] calculations combined with a complete basis set (CBS) extrapolation based on DLPNO-MP2 calculations.

$$E = E^{\mathrm{PBC}}_{\mathrm{PBE\text{-}D3}} - E^{\mathrm{46T}}_{\mathrm{PBE\text{-}D3}} + E^{\mathrm{46T}}_{\mathrm{DLPNO\text{-}CCSD(T)/DZ}} + \Delta E^{\mathrm{46T}}_{\mathrm{MP2/CBS}} \tag{4.1}$$

$$\text{with} \quad \Delta E^{\mathrm{46T}}_{\mathrm{MP2/CBS}} = E^{\mathrm{46T}}_{\mathrm{DLPNO\text{-}MP2/CBS}} - E^{\mathrm{46T}}_{\mathrm{DLPNO\text{-}MP2/DZ}} \tag{4.2}$$

In this approach, the cc-pVDZ basis set was used for $E^{\mathrm{46T}}_{\mathrm{DLPNO\text{-}CCSD(T)/DZ}}$ of the 46T cluster model, the def2-TZVPP basis set was used for $E^{\mathrm{46T}}_{\mathrm{PBE\text{-}D3}}$, and $\Delta E^{\mathrm{46T}}_{\mathrm{MP2/CBS}}$ stands for the difference between MP2-based CBS extrapolation ($E^{\mathrm{46T}}_{\mathrm{DLPNO\text{-}MP2/CBS}}$) and MP2/cc-pVDZ calculations for 46T clusters ($E^{\mathrm{46T}}_{\mathrm{DLPNO\text{-}MP2/DZ}}$). The CBS extrapolations of Hartree-Fock energies were carried out with the three-point exponential formula[273] with cc-pVXZ (X=D, T, Q). For MP2-correlation, the two point l^{-3} formula[274] was used with cc-pVXZ (X=D, T). The non-periodic calculations were performed on 46T cluster models using the ORCA[91,275] and TURBOMOLE[276] program packages. PBE-D3 calculations on the 46T cluster models were performed using TURBOMOLE along with the def2-TZVPP basis set.[277,278] ORCA was used to perform DLPNO-CCSD(T),[90] DLPNO-MP2[279–281] and restricted Hartree-Fock (RHF) calculations for

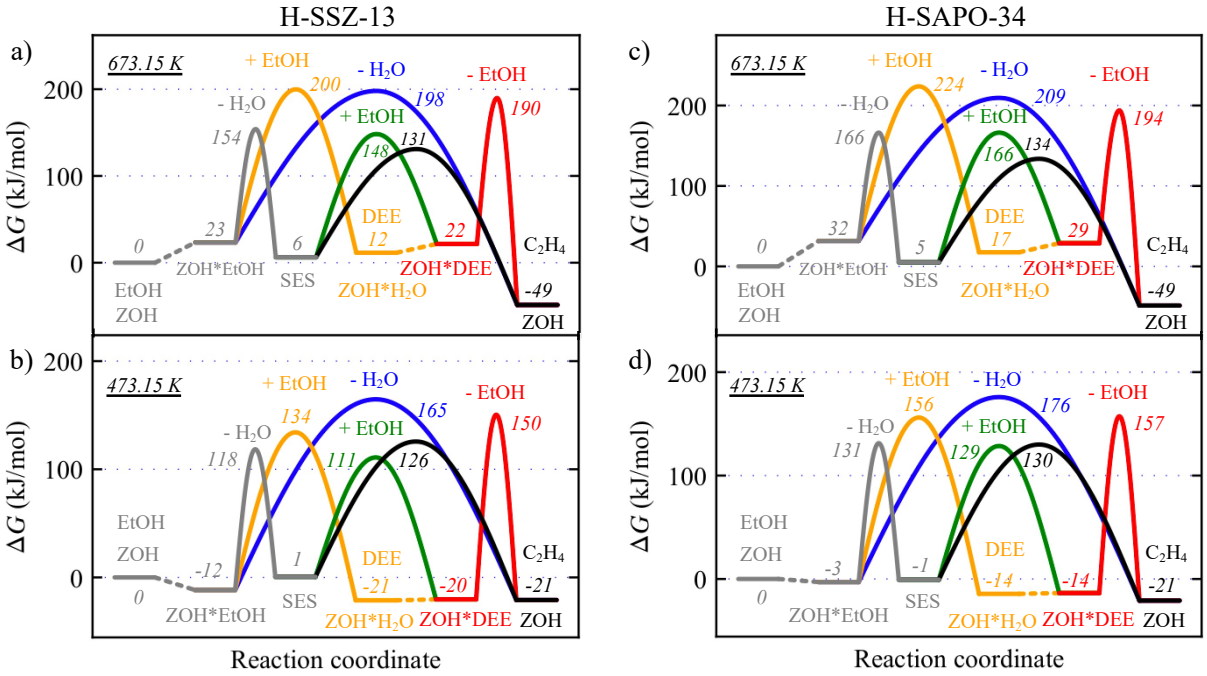


Figure 4.1.: Calculated Gibbs free energy diagrams for the conversion of ethanol (EtOH) to diethyl ether (DEE) and ethene at a reference pressure of 1 bar in H-SSZ-13 (a, b) and H-SAPO-34 (c, d) for a temperature of 673.15 K (a, c) and 473.15 K (b, d). Reprinted (a, b) with permission from [245]. Copyright 2022 The Royal Society of Chemistry (CC-BY).

cc-pVXZ (X=D, T, Q)[282] in the DLPNO approximation.[90,283–285] The RIJCOSX (resolution of identity for Coulomb integrals and semi-numerical chain-of-sphere integration for Hartree-Fock exchange integrals) approximation[286] with GridX6 was used in RHF calculations. Justified by known correlations between H-SSZ-13 and H-SAPO-34,[24] the periodic models of the latter were corrected by the post-Hartree-Fock results of the former. Parts of the calculations were operated using the atomic simulation environment (ASE).[209]

The kinetic simulations with the batch reactor model were carried out based on the mean field approximation using simple Euler integration with a maximum time step smaller than $5 \cdot 10^{-7}$ s as implemented in the in-house code used in earlier work by Plessow et al.[80,246] Diffusion limitations are neglected and we assume instantaneous equilibration of adsorption processes. For numerical reasons, the lower threshold of 80 kJ mol^{-1} for adsorption barriers was set when deriving thermodynamically consistent kinetic parameters from the barriers. It was verified that lowering this threshold has no impact on the outcome of the simulations, see Amsler et al.[245] The reference concentration of the active site (ZOH) was set to 17.9 mol m^{-3} analogous to earlier work by Plessow et al.[80] where further details on the kinetic model can be found. Isothermal plug flow simulations under steady-state conditions were performed with the PLUG code of the program package DETCHEM$^{\text{PLUG}}$,[102] see also Section 2.4.2. The length of the cylindrical catalyst bed was set to 40 mm with a radius of 2 mm.

4.1.2. Results and Discussion

We computed stepwise and concerted reaction pathways for the conversion of EtOH to DEE and ethene as depicted in Scheme 4.1. The Gibbs free energy diagrams for the catalysts H-SSZ-13 and H-SAPO-34 are shown in Figure 4.1. EtOH dehydration to DEE or ethene starts by adsorption of EtOH at the acid site. For zeolite H-SSZ-13 we computed a Gibbs free energy of adsorption of −12 kJ mol^{-1} and

23 kJ mol^{-1} for the temperatures of 473.15 K and 673.15 K, respectively. Starting from adsorbed EtOH (ZOH*EtOH), the reaction network for the formation of DEE comprises both the stepwise (gray and green lines) and concerted (orange line) mechanisms (compare Scheme 4.1). Due to a high barrier for direct DEE decomposition on H-SSZ-13 (red line), the stepwise dehydration mechanism has lower overall free energy barriers, already at 473.15 K. At 673.15 K, the stepwise barriers for zeolite H-SSZ-13 are 131 kJ mol^{-1} and 125 kJ mol^{-1} when measured from adsorbed EtOH. Ethene formation can occur via three different mechanisms, namely via (1) direct dehydration of adsorbed EtOH (blue line), (2) the dehydration of adsorbed DEE (red line) and (3) from surface ethoxy species (SES), whereby the acid site is reestablished again (black line). At 673.15 K in zeolite H-SSZ-13, these barriers have values of 175 kJ mol^{-1}, 168 kJ mol^{-1}, and 125 kJ mol^{-1}, when referenced to adsorbed EtOH, adsorbed DEE and SES, respectively. Judging from the calculated Gibbs free energy diagram for H-SSZ-13 in Figure 4.1a at 673.15 K (and likewise for H-SAPO-34 in Figure 4.1c), the stepwise mechanism (formation of SES and subsequent elimination to ethene (gray and black lines)) has the overall lowest barriers for ethene formation (154 kJ mol^{-1} and 131 kJ mol^{-1} in zeolite H-SSZ-13 when referenced to EtOH in the gas phase and adsorbed at the acid site, respectively). We observe, however, that the ranking of the barriers for DEE and ethene formation (green and black lines) changes with temperature. In consequence, DEE formation at 473.15 K is kinetically favored even though ethene is thermodynamically slightly more stable (the chemical equilibrium between EtOH, DEE and ethene depends strongly on the temperature, see Amsler et al.[245] and Garbarino et al.[250]). We can use this finding to explain why the conversion-dependent formation of DEE decreases with increasing temperature. The barriers for H-SAPO-34 in Figure 4.1c and d reveal the exact same trends and effects of temperature as discussed for H-SSZ-13, however, the barriers of H-SAPO-34 are in general slightly shifted to higher energies. In particular, the crucial barrier for SES formation increases by 12 kJ mol^{-1} at 673.15 K. The observed trends are no surprise because H-SSZ-13 and H-SAPO-34 exhibit the same chabazite topology with a pore size of 3.7 Å and virtually identical van-der-Waals (vdW) interactions – the slightly lower barriers (higher rates) for H-SSZ-13 are caused by its higher acidity.[24]

To shed more light on the formation of DEE and ethene and to identify reaction pathways in catalyst H-SSZ-13, we also performed simulations of the reaction kinetics using a batch reactor model with the reaction scheme shown in Scheme 4.1. The barriers at 673.15 K for the kinetic model derived from Figure 4.1 are summarized in Table 4.1 and the data at 473.15 K and 573.15 K are listed in Amsler et al.[245] Figure 4.2 shows the outcome of the kinetic simulations using a feed of 1 bar EtOH and three different temperatures. The upper panel of Figure 4.2 shows simulated partial pressures, while the coverages of all possible intermediates are given in the middle row. Note that due to our choice of initial settings, surface coverage corresponds equally to partial pressure (a DEE coverage of 0.15 corresponds to 0.15 bar DEE when desorbed). As expected, surface coverages decrease with increasing temperatures where adsorbed water dominates among the surface species. At 473.15 K, DEE is the dominating product during the first minute where half of the amount of DEE is adsorbed at the reactive sites. Our simulations predict a theoretical initial DEE selectivity (S) of 90 % at 0.2 % conversion which decreases monotonically with increasing EtOH conversion (X). After 50 s at 473.15 K we observe a peak DEE yield with a DEE selectivity of 61 % and an EtOH conversion of 79 %. At 673.15 K on the other hand, ethene is simulated to form within hundreds of ms, while DEE is not present in large amounts throughout the simulation time. This behavior depending on temperature and conversion can be understood by the temperature-dependent ranking of the barriers for DEE and ethene formation, which was discussed above. We also note that the absence of DEE formation at elevated temperatures is in line with experimental observations.[250,253,257–260,262,287] Besides the temperature dependence, we also observe conversion to be significant for DEE and ethene selectivity. After one minute at 473.15 K, 81 % EtOH have been converted into DEE (56 %) and ethene (25 %) on a per carbon basis. At higher conversion and for longer reaction times, ethene becomes the dominant product since the

Table 4.1.: Adsorption free energies of water, EtOH and DEE as well as forward and reverse free energy barriers in catalyst H-SSZ-13 at 673.15 K and a reference pressure of 1 bar for all elementary reaction steps shown in Figure 4.1. Reprinted with permission from [245]. Copyright 2022 The Royal Society of Chemistry (CC-BY).

No.	Elementary reaction		$\Delta G^{\ddagger}\left(\frac{\text{kJ}}{\text{mol}}\right)$
1	ZOH + H_2O	$\rightleftarrows$ ZOH*H_2O	25[a]
2	ZOH + EtOH	$\rightleftarrows$ ZOH*EtOH	23[a]
3	ZOH + DEE	$\rightleftarrows$ ZOH*DEE	35[a]
4	ZOH*EtOH	$\rightleftarrows$ SES + H_2O	131, 148
5	SES + EtOH	$\rightleftarrows$ ZOH*DEE	142, 127
6	ZOH*EtOH + EtOH	$\rightleftarrows$ ZOH*H_2O + DEE	176, 188
7	SES	$\rightleftarrows$ ZOH + ethene	125, 180
8	ZOH*EtOH	$\rightleftarrows$ ZOH*H_2O + ethene	175, 222
9	ZOH*DEE	$\rightleftarrows$ ZOH*EtOH + ethene	168, 215

[a] adsorption free energies; low thermodynamically consistent barriers have been used for numerical reasons with no rate-limiting effect to the simulation results, see Amsler et al.[245]

thermodynamic equilibrium favors ethene even at 473.15 K (see Amsler et al.[245]). A high variation of the ethene selectivity within a narrow temperature range and depending on the catalyst contact time was also reported for EtOH dehydration in H-SAPO-34 by Potter et al.[253] At 473.15 K and a contact time of 22 min they observed 40 mol % DEE and 10 mol % ethene. Moreover, ethene increased to 20 mol % when the contact time increased to 90 min. At 503.15 K, their measured ethene selectivity went up to 45 mol % and 80 mol % for the same contact times of 22 min and 90 min, respectively. As discussed above, the catalysts H-SSZ-13 and H-SAPO-34 exhibit the same chabazite topology and we observe similar effects of temperature on the dehydration of EtOH. Nevertheless, the barriers tend to be slightly lower for catalyst H-SSZ-13 which can be explained by its higher acidity. In MTO experiments, H-SSZ-13 displayed a higher conversion capacity at lower temperatures.[288] Our kinetic batch reactor simulations thus show that DEE selectivity varies greatly with temperature, conversion and catalyst contact time, with DEE being mainly observed at lower temperatures and lower conversion.

Experimental studies on ethanol dehydration by our collaborators are typically carried out with zeotype H-SAPO-34 in a continuously operated tubular reactor. Therefore, we performed plug flow simulations for this catalyst in a catalyst bed as described in Section 4.1.1.1. The catalyst mass was set to $m_{\text{cat}} = 0.3\,\text{g}$ with an active site concentration of $1300\,\text{mol}\,\text{m}^{-3}$, which translates to a surface concentration parameter of $1.307\,\text{mol}\,\text{m}^{-2}$ in the software package DETCHEM$^{\text{PLUG}}$. The feed at a total pressure of 1 bar consists of 0.1 bar EtOH diluted in nitrogen as inert gas. The resulting mole fractions of the species in the gas phase and on the catalyst surface at the steady state at 428.15 K as a function of the length of the catalyst bed are shown in Figure 4.3. The flux velocity u_0 at the inlet was calculated according to eq. (4.3), with the weight hourly space velocity (WHSV) defined as mass flux of ethanol per mass of catalyst per hour, the molar mass of ethanol M_{EtOH}, the ideal gas constant R, the temperature T, the partial pressure p_{EtOH} of ethanol and the cross-sectional area a of the cylindrical catalyst bed.

$$u_0 = WHSV * \frac{m_{\text{cat}}}{M_{\text{EtOH}}} * \frac{RT}{p_{\text{EtOH}}} * \frac{1}{a} \tag{4.3}$$

For a WHSV in the range from $0.01\,\text{h}^{-1}$ to $0.05\,\text{h}^{-1}$ we observe a fast formation of ethene and DEE in the first half of the catalyst bed, where adsorbed EtOH dominates among the surface species. Slow conversion of DEE to ethene can be seen in the second half of the catalyst bed. Interestingly, the presence of SES correlates with the dynamics of the gas phase composition – as long as there is a

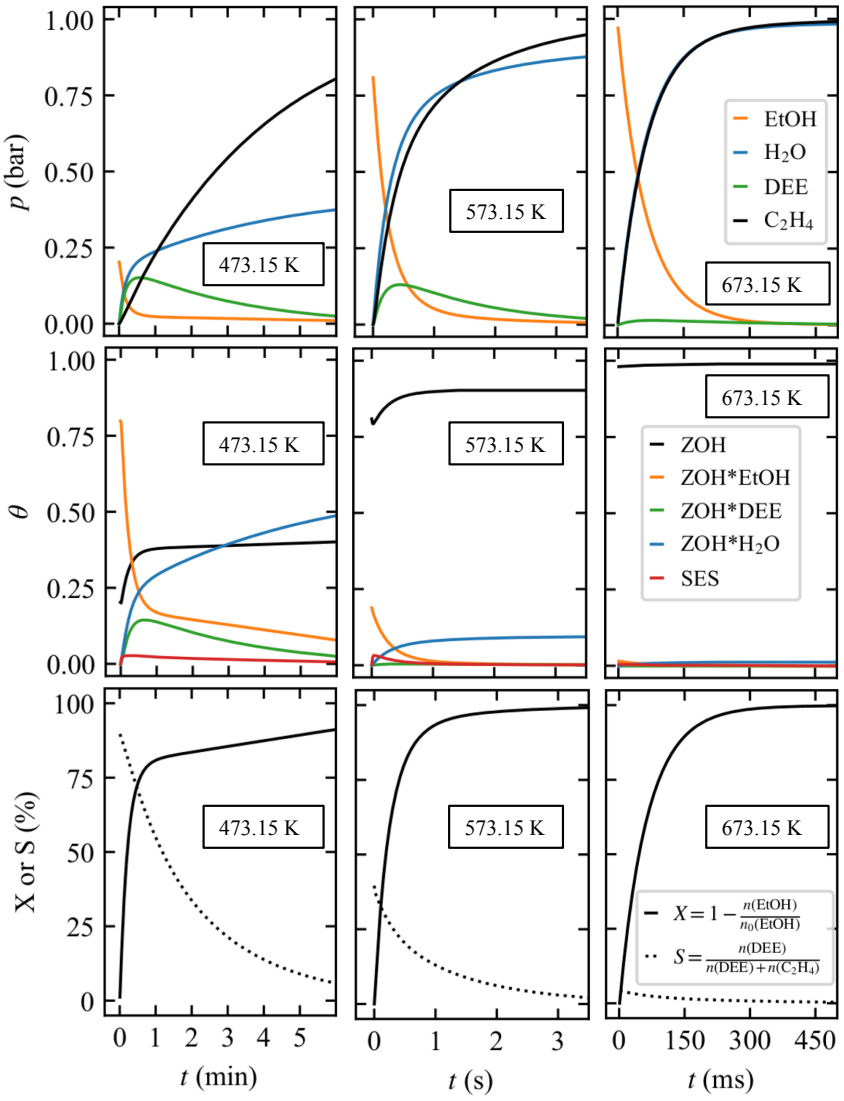


Figure 4.2.: Simulated partial pressures of gas phase species (top panel), surface coverages (middle row) as well as EtOH conversion (X) and selectivity (S) of DEE (bottom row) in a batch reactor with catalyst H-SSZ-13 at three different temperatures as a function of time using the reaction mechanism shown in Scheme 4.1 and an initial pressure of 1 bar EtOH. Note that the time scale differs significantly from minutes at 473.15 K, over seconds at 573.15 K to milliseconds at 673.15 K. Reprinted with permission from [245]. Copyright 2022 The Royal Society of Chemistry (CC-BY).

noticeable amount of SES, the gas phase composition changes rapidly. Furthermore, the ratio between DEE and ethene at the end of the catalyst bed strongly depends on the WHSV, while EtOH is nearly completely consumed independently of the WHSV. In the plug flow reactor simulations, we define the selectivity S of any species by its mole fractions n normalized on a per carbon basis according to eq. (4.4) – note that DEE contains twice the amount of carbon, thus the theoretical peak selectivity cannot be greater than 0.5 according to eq. (4.4).

$$S(\text{species}) = \frac{n_{\text{species}}}{2 \cdot n_{\text{DEE}} + n_{\text{ethene}} + n_{\text{EtOH}}} \tag{4.4}$$

The plug flow reactor simulations were carried out for a range of temperatures and the gas phase composition at the end of the catalyst bed was monitored. Figure 4.4 shows the results as a function of the WHSV and of the temperature in a format that could easily be compared to experiments by our collaborators. The mole fraction of EtOH at the end of the catalyst (see Figure 4.4a) is marginal and

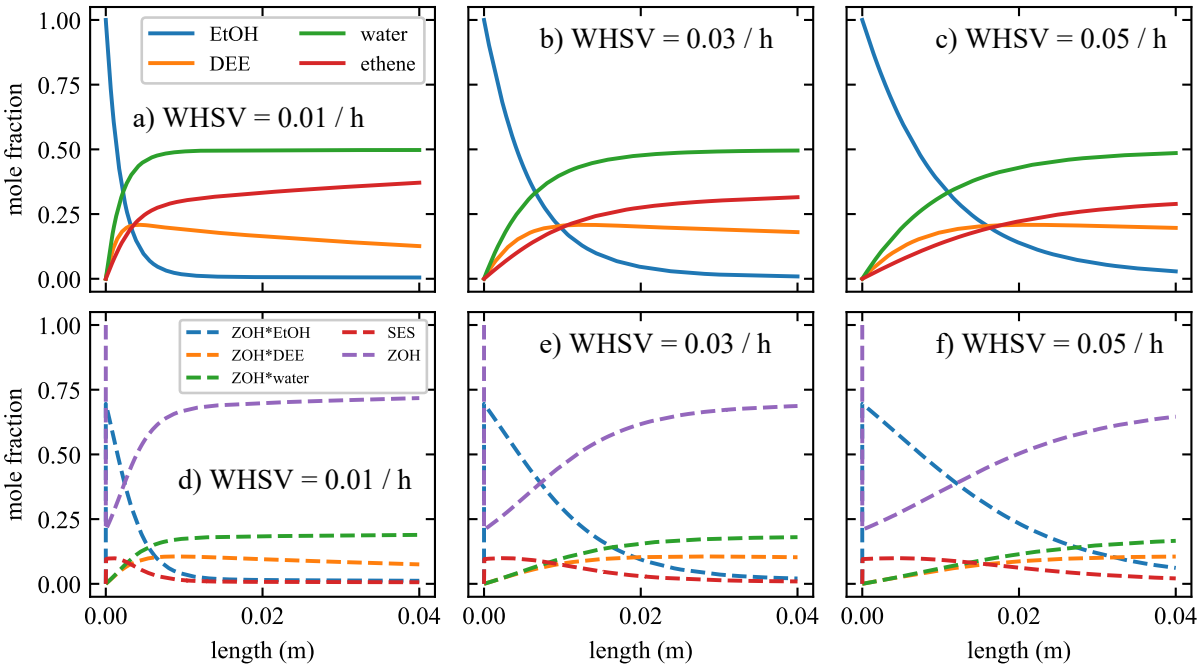


Figure 4.3.: Simulated mole fractions of gas phase species (a, b, c) and surface coverages (d, e, f) for catalyst H-SAPO-34 in the catalyst bed of a plug flow reactor at 428.15 K and for a WHSV of 0.01 h^{-1} (a, d), 0.03 h^{-1} and 0.05 h^{-1} as a function of the catalyst bed length. Employing the reaction mechanism in Scheme 4.1, the EtOH partial pressure in the feed was set to 0.1 bar.

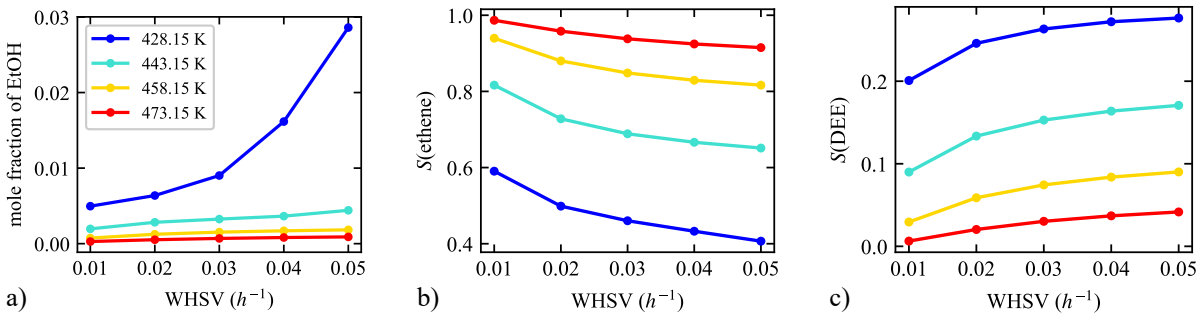


Figure 4.4.: Predictions for the gas phase composition at the end of the catalyst bed at different temperatures as a function of the WHSV. (a) Mole fraction of EtOH. (b) Selectivity of ethene and (c) selectivity of DEE as defined in eq. (4.4).

independent of the WHSV at all temperatures but at 428.15 K, where the still small but exponentially increasing mole fraction of EtOH is caused by incomplete conversion due to shorter contact times at higher values of the WHSV. The selectivity of ethene, as defined by eq. (4.4), decreases with increasing WHSV, while the selectivity for DEE increases, see Figures 4.4b and 4.4c. At 473.15 K, the ethene selectivity is near 1 for a WHSV of 0.01 h^{-1}. Here, a decrease in temperature strongly effects the selectivity, partially due to the shifting thermodynamic equilibrium, but mostly due to the further decrease of the barrier for DEE formation from SES relative to the barrier for competing ethene formation (compare Figure 4.1).

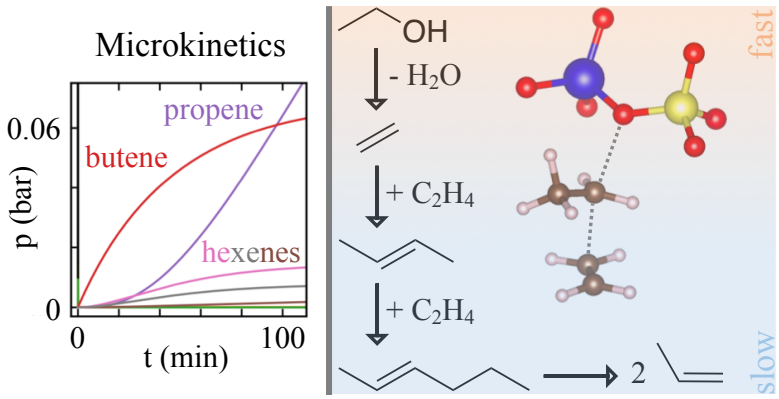
4.1.3. Intermediate Summary and Conclusions

The formation of DEE and ethene from EtOH was studied using kinetic simulations in batch and plug flow reactor models with *ab initio* calculated rate constants for zeolite catalyst H-SSZ-13 compared

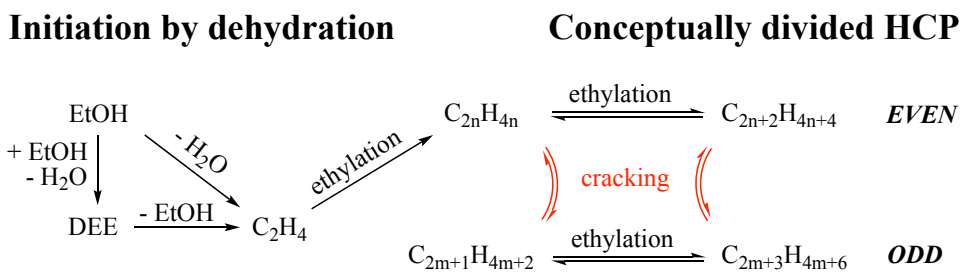
to zeotype catalyst H-SAPO-34. Similar trends for different reaction pathways and similar effects of temperature were observed for both catalysts. While H-SSZ-13 and H-SAPO-34 exhibit the same chabazite topology with a pore size of 3.7 Å and virtually identical van-der-Waals (vdW) interactions, the slightly lower barriers (higher rates) for H-SSZ-13 were explained by its higher acidity. For instance, the crucial barrier for SES formation is decreased by 12 kJ mol^{-1} at 673.15 K in H-SSZ-13 compared to H-SAPO-34. With increasing temperature, the ranking of the barriers for DEE and ethene formation changes, which explains the temperature dependence of the conversion-dependent ratio between DEE and ethene. Essentially, the free energy landscape suggests that DEE formation can be kinetically favored at 473.15 K before the thermodynamically slightly more stable ethene eventually dominates the gas phase composition upon higher conversions. Indeed, temporary DEE formation with initially high DEE selectivity was observed in the batch reactor simulations for catalyst H-SSZ-13 within the first couple of minutes or seconds, at 473.15 K or 573.15 K, respectively, as long as the conversion of EtOH was relatively low. This is due to the fact that at these temperatures the dehydration pathway to DEE exhibits lower reaction barriers compared to the dehydration pathway to ethene. Ethene is formed rapidly at temperatures above 473.15 K with the main reaction pathway going through SES. While the batch reactor simulations with catalyst H-SSZ-13 showed that the selectivity depends on the conversion, the plug flow simulations with catalyst H-SAPO-34 showed that the selectivity also varies with the WHSV in a regime where the conversion is nearly complete. Highly dynamic changes in the gas phase composition along the catalyst bed were observed to correlate with the presence of SES, which further strengthens our confidence for its presumed key role in the reaction mechanism. Our study thus provides theoretically derived kinetic parameters for the dehydration of EtOH as well as reactor simulations tailored to a typical setup of our experimental collaborators. The effect of the EtOH partial pressure on the DEE selectivity could be an interesting subject of future research.

4.2. Ethanol-to-Olefins – Exploration of Reaction Pathways

Preface. The formation of the hydrocarbon pool (HCP) in the ethanol-to-olefins (ETO) process catalyzed by H-SSZ-13 is now studied in a kinetic model with *ab initio* computed reaction barriers. Free energy barriers are computed using density functional theory (DFT) and post-Hartree-Fock methods with a complete basis set extrapolation applied to a hierarchy of periodic and cluster models – entropy contributions are estimated using the rigid-rotor harmonic-oscillator approximation. The kinetic model includes ethanol (EtOH) dehydration to ethene as well as olefin ethylations up to hexene isomers and the



Scheme 4.2.: Left: Simulated ethanol (EtOH) dehydration and formation of olefins in a batch reactor model at 673.15 K from 1 bar EtOH with parameters derived from *ab initio* calculations. Right: Simplified version of the mechanism used in the simulations. Reprinted with permission from [245]. Copyright 2022 The Royal Society of Chemistry (CC-BY).



Scheme 4.3.: Dehydration of EtOH leads to the HCP, that can be conceptualized as being divided into the even and the odd-numbered olefins self-contained through ethylations but partially interconnected by cracking reactions ($n, m \in \mathbb{N}$). Adapted and reprinted with permission from [245]. Copyright 2022 The Royal Society of Chemistry (CC-BY).

corresponding cracking reactions, see the overview in Scheme 4.2. Ethylation of ethene and of products thereof leads only to even-numbered olefins, while cracking can lead to propene and thus initiate the formation of olefins with an odd number of carbon atoms. At 673.15 K we find that EtOH dehydration occurs much faster than ethylation of the formed ethene, which takes considerably longer due to higher free energy barriers. Hexene isomers form on the same time scale as butene, where branched isomers are favored with 2-methyl-pentene isomers contributing most to the formation of propene through cracking. As in the methanol-to-olefins (MTO) process, the most relevant alkylation pathway is the stepwise mechanism via surface alkoxy species (SAS) on the zeolite catalyst. A comparison of ethylation with methylation barriers of up to heptene isomers forming nonene and octene isomers, respectively, shows that ethylation barriers are lower by around 11 kJ mol^{-1} on average.

4.2.1. Overview and Simulation Details

The zeolite-catalyzed ethanol-to-olefins (ETO) process and its closely related methanol-to-olefins (MTO) process are promising key drivers in the value-added chain from biomass to fuels and raw materials for the petrochemical industry partaking in a sustainable energy economy.[25,26] Aided by crystalline microporous zeolite catalysts at temperatures around 673 K to 823 K and intermediate pressures around 1 bar to 5 bar, ethanol (EtOH) and methanol (MeOH) are converted into light olefins in these processes.[28] The current understanding of the mechanistic details revolves around the concept of the hydrocarbon pool (HCP) comprised of a variety of (unsaturated) hydrocarbons that continuously undergo oligomerization and cracking reactions,[27,30–32,34–36] see Section 1.2.2. While the initial formation of C–C bonds in the multistep initiation of the MTO process from MeOH has been attributed to either impurities[80,246,289,290] or direct initiation,[291–294] the ETO process initiates via EtOH dehydration to ethene.[295] In a majority of experimental studies, the ETO conversion was performed on (modified) H-ZSM-5 where the first stage of EtOH dehydration proceeds at a high rate and is virtually independent of the selectivity of the catalyst towards higher olefins.[25] The most important acid-catalyzed reactions (dehydration, alkylation/oligomerization, isomerization, cyclization, aromatization, hydrogen transfers and cracking) in the HCP are promoted by the acidic Brønsted sites and governed by the structural properties (framework, pore size) as well as the acidity of the catalyst.[296]

Chowdhury et al. have elucidated the homologation reaction network originating from ethene with homologated and non-homologated products.[36] In the absence of methylation reactions, the ETO reaction network can thus be thought of two separate networks for the even and odd-numbered olefins, interconnected by cracking reactions, see Scheme 4.3. This concept of even and odd-numbered carbon units is further supported by Ingram and Lancashire who found that propene formation mainly originates from cracking of hexene isomers.[297]

Quantum chemical calculations, typically based on density functional theory (DFT) are nowadays routinely used to study the reaction mechanisms associated with, for example, the HCP. These methods can predict reaction barriers thus shedding light on reaction mechanisms and determining possible rate determining steps. Often, the obtained data also serves as input for (micro-)kinetic models of the corresponding processes.[23,37,44,219,298–303] While DFT and *ab initio* post-Hartree-Fock calculations have been used extensively for MTO and related processes,[16,21,55,80,219,246,270,299,302,304,305] theoretical insight into the analogous conversion of EtOH in acidic zeolites has so far mostly been limited to EtOH dehydration to ethene and DEE[252,265–267] and is rather scarce for subsequent propagation steps.

Herein we investigate the reaction pathways in the ETO process with zeolite H-SSZ-13 using the same theoretical approach as previously in Section 4.1 on EtOH dehydration. The calculated reaction barriers serve as input for a microkinetic model comprising ethanol dehydration in competition with diethyl ether (DEE) formation as well as subsequent ethylation and cracking reactions up to hexene isomers. We furthermore compare stepwise ethylation barriers with corresponding methylation barriers in the MTO process.

4.2.1.1. Simulation Details

Simulations were performed as described in Section 4.1.1.1. Since we only consider the initial part of the ETO reaction mechanism we neglect diffusion limitations in the reactor simulations, thus assuming perfect mixing, as diffusion of smaller olefins has been shown to be rather fast.[57] Olefin isomerization barriers that only shift the double bond are expected to be small compared to ethylation and cracking barriers[56] and are therefore treated implicitly; for instance, 1-butene and 2-butene are lumped together as butene, 2-methyl-2-hexene is simply called 2-methyl-hexene, and so on.

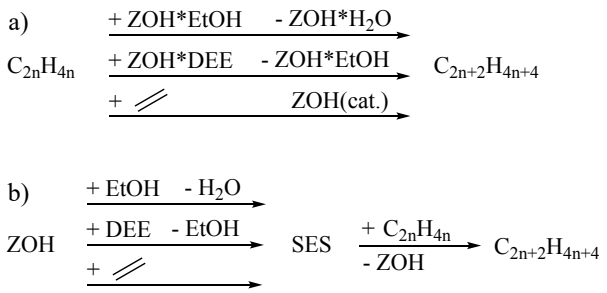
4.2.2. Results and Discussion

This study is divided into two parts. We start with our results regarding the ethylation of ethene to butene and further formation of hexene as well as the selectivities among hexene isomers. Then we discuss similarities and differences of EtOH conversion with the well-established MTO process.

4.2.2.1. Ethylation of Olefins

In this section, we focus on a small subset of the ETO reaction network comprising only olefins with an even number of carbon atoms as well as their cracking products (including propene). The alkylation pathways shown in Scheme 4.4 are inspired by the methylation pathways in the MTO process.[21] We consider stepwise and concerted mechanisms in zeolite H-SSZ-13, in other words, the formation of surface ethoxy species (SES) followed by a reaction with an olefin and the direct reaction of the olefin with EtOH, DEE, or ethene at the acid site of the catalyst.

Ethylation of ethene to butene. After ethene formation from EtOH, which was investigated in Section 4.1, the route to higher olefins necessarily proceeds via ethylation of ethene.[295] As mentioned above, we always consider ethylation pathways through the three ethylation agents (EtOH, DEE and SES) and potentially further ethylation pathways by olefin dimerization. Butene formation from ethene can thus proceed via the four reactions listed in Table 4.2. From the listed intrinsic free energy barriers, the lowest barrier of 173 kJ mol^{-1} in zeolite H-SSZ-13 is found for ethylation by SES formed from



Scheme 4.4.: Depiction of (a) concerted and (b) stepwise ethylation paths. Adapted and reprinted with permission from [245]. Copyright 2022 The Royal Society of Chemistry (CC-BY).

Table 4.2.: Elementary reaction steps for the formation and cracking of butene with forward and reverse free energy barriers at 673.15 K and at a reference pressure of 1 bar. Reprinted with permission from [245]. Copyright 2022 The Royal Society of Chemistry (CC-BY).

No.	Elementary reaction	$\Delta G^{\ddagger}\left(\frac{\text{kJ}}{\text{mol}}\right)$
10	ZOH*EtOH + ethene $\rightleftarrows$ ZOH*H_2O + butene	205, 274
11	SES + ethene $\rightleftarrows$ ZOH + butene	173, 250
12	ZOH*DEE + ethene $\rightleftarrows$ ZOH*EtOH + butene	199, 267
13	ZOH + 2 ethene $\rightleftarrows$ ZOH + butene	250, 272

ethene (at 673.15 K). Note that the lowest ethylation barrier is still significantly higher than the lowest barrier for ethene formation from SES with 125 kJ mol^{-1}. We thus conclude that the conversion of EtOH to ethene is much faster than the subsequent butene formation. Unsurprisingly, the intrinsic formation/cracking barrier for the short-chained butene (reaction no. 13) is very high in both directions (250 kJ mol^{-1} and 272 kJ mol^{-1}).[56,306] The Gibbs free energy diagrams corresponding to stepwise and concerted butene formation and cracking at 673.15 K and at a reference pressure of 1 bar along with optimized transition structures are shown in Figure 4.5. Inspection of Figure 4.5c shows that the intrinsic barrier for the SES-mediated dimerization of ethene is higher in comparison to the reactions of ethene with EtOH (Figure 4.5a) or DEE (Figure 4.5b). Thermodynamically, however, ethene is considerably lower in free energy than DEE or EtOH (see Figure 4.1 and Amsler et al.[245]), such that the ethene dimerization to butene in Figure 4.5c is actually the most favorable pathway from a kinetic point of view. In addition, since ethene is formed rapidly by EtOH dehydration, see Figure 4.2, the partial pressure of ethene will be much higher than that of both EtOH or DEE. We therefore conclude, that the overall lowest ethylation barrier is found for the stepwise ethylation pathway via SES.

Formation of hexene isomers. Ethylation from butene onwards can lead to a variety of hexene isomers. The Gibbs free energy diagrams for the corresponding ethylation, cracking and methyl shift reactions are shown in Figure 4.6. Analogous to the ethylation of ethene, the overall lowest ethylation barrier of butene is found for the stepwise ethylation by SES formed from ethene (black barriers in Figure 4.6c), followed by ethene dimerization (red barrier in Figure 4.6c), direct ethylation through DEE (orange barrier in Figure 4.6b) and direct ethylation with EtOH (blue barrier in Figure 4.6a). Note that ethene in lower in energy than DEE, which is lower in energy than EtOH. Hence, our calculations indicate that the preferred alkylation pathway in the ETO process is similar to that identified for the MTO process, that is alkylation via surface alkoxy species (SAS) in the stepwise pathway. A comparison of the ethylation pathways towards linear and branched hexene isomers shows that the formation of branched products is favored with lower barriers and more favorable reaction energies which principally

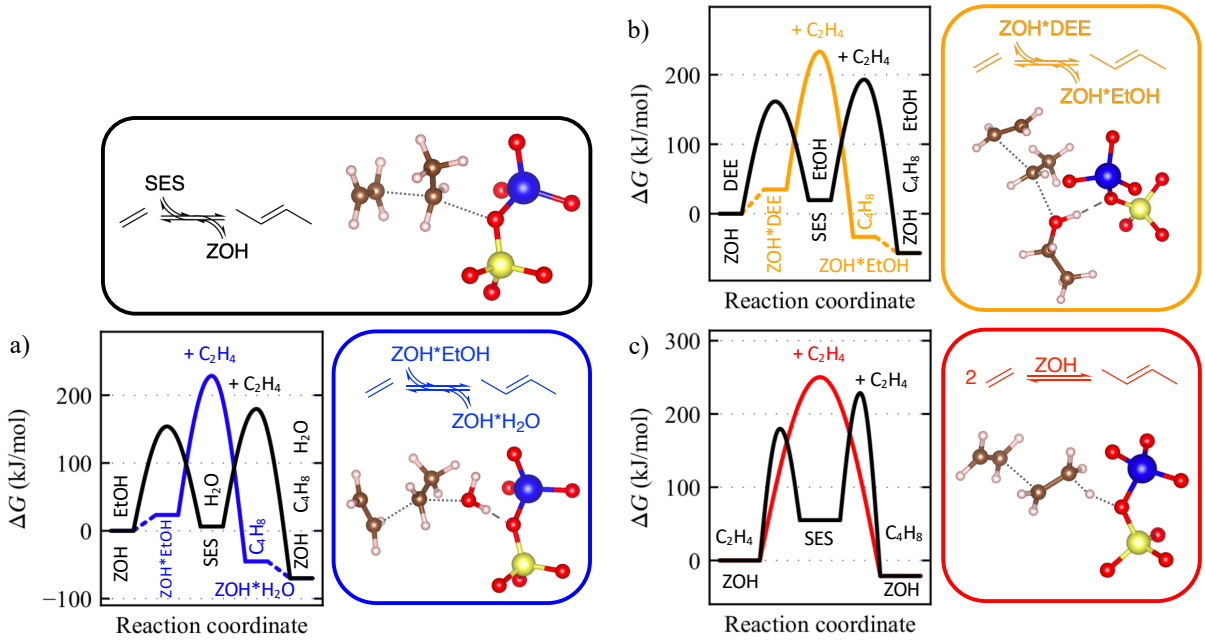


Figure 4.5.: Gibbs free energy diagrams and transition structures for the direct ethylation of ethene at 673.15 K and at a reference pressure of 1 bar through (a) EtOH (blue), (b) DEE (orange) and (c) ethene (red). Corresponding stepwise ethylations are shown in black. Adapted and reprinted with permission from [245]. Copyright 2022 The Royal Society of Chemistry (CC-BY).

obeys the idea of the Brønsted-Evans-Polanyi (BEP) principle. For analogous stepwise methylations of olefins by MeOH this BEP principle, which relates energy barriers with reaction energies, has been applied in the literature to estimate barriers from reaction energies.[64,65] The preference of branched products over linear products can also generally be observed, as we will see later in Section 4.2.2.2. In particular, the stepwise ethylation barrier for the formation of 3-methyl-pentene is by 19 kJ mol^{-1} lower than the analogous formation of hexene. Relative to ethene in the gas phase the free energy barriers for the methyl shift of 3-methyl-pentene to 2-methyl-pentene and subsequent cracking to propene are slightly lower than the formation of 3-methyl-pentene from ethene and butene (see Figure A.4 in the Appendix and Amsler et al.[245]). Once longer olefins are formed, the formation of propene and subsequent formation of other odd-numbered olefins is hence kinetically favorable. Using our calculated free energy barriers in Table 4.3 together with the data in Tables 4.1 and 4.2, we extended our kinetic model to simulate the entire ETO reaction network. The results of this kinetic simulation are shown in Figure 4.7 for a temperature of 673.15 K with a feed of 1 bar of EtOH. As can be seen, the kinetic regimes can roughly be separated into 3 time domains. At 673.15 K, ethene is formed from EtOH within milliseconds (Figure 4.7a). Ethene then dimerizes mediated by SES to form butene on the scale of minutes (Figure 4.7b), with the olefin partial pressures eventually approaching thermodynamic equilibrium through dimerization, ethylation and cracking (propene formation, Figure 4.7c).

Note that branched hexene isomers are observed in small amounts while the partial pressure of linear hexene stays noticeably lower, which is due to the lower barriers for the formation of branched olefins (see Figures 4.6 and 4.7). Due to the high temperatures, the equilibrium partial pressures of hexene isomers are always low and these species present intermediates towards propene. The propene formation accompanying the formation of hexene isomers is also within our expectations based on the computed barriers shown in Figure 4.6. In this regard, the methyl shift of 3-methyl-pentene to 2-methyl-pentene and subsequent cracking to propene contributes most significantly to propene formation. From

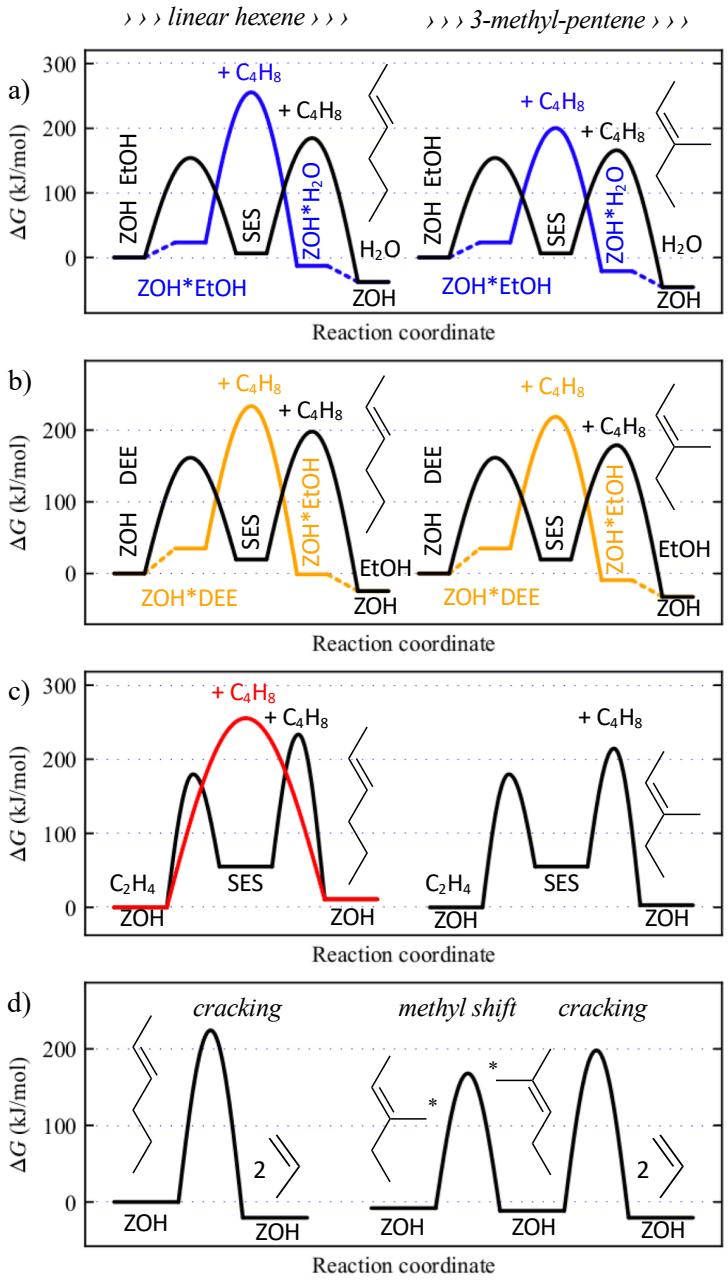


Figure 4.6.: Gibbs free energy diagrams for the direct ethylation of butene at 673.15 K and at a reference pressure of 1 bar through (a) EtOH (blue), (b) DEE (orange) and (c) ethene (red). Corresponding stepwise ethylations are shown in black. The formation of linear (left) and branched (right) hexene isomers is compared. The barriers for a methyl shift and for propene formation through cracking are shown in (d). Reprinted with permission from [245]. Copyright 2022 The Royal Society of Chemistry (CC-BY).

the Gibbs free energy diagrams, it is also evident that the ethylation barriers for butene are lower than the ethylation barrier for ethene, similar to the trends of methylation barriers in the MTO process.[56]

Since odd-numbered olefins are eventually formed through cracking, the ETO process shows fairly similar product distributions compared to that observed for MTO.[34] Our findings hence support the work of Ingram and Lancashire who demonstrated that propene originates from cracking of presumably 2-methyl-pentene rather than from intermediate steps of butene formation.[297]

Table 4.3.: Elementary reaction steps for the formation, cracking and methyl shift reactions of hexene isomers with forward and reverse free energy barriers at 673.15 K and at a reference pressure of 1 bar. Reprinted with permission from [245]. Copyright 2022 The Royal Society of Chemistry (CC-BY).

No.	Elementary reaction	$\Delta G^{\ddagger}\left(\frac{\text{kJ}}{\text{mol}}\right)$
14	ZOH*EtOH + butene ⇄ ZOH*H_2O + hexene	232, 268
15	ZOH*EtOH + butene ⇄ ZOH*H_2O + 3-Me-C5[a]	177, 221
16	SES + butene ⇄ ZOH + hexene	178, 223
17	SES + butene ⇄ ZOH + 3-Me-C5[a]	159, 212
18	ZOH*DEE + butene ⇄ ZOH*EtOH + hexene	199, 235
19	ZOH*DEE + butene ⇄ ZOH*EtOH + 3-Me-C5[a]	184, 228
20	ZOH + hexene ⇄ ZOH + ethene + butene	245, 255
21	ZOH + hexene ⇄ ZOH + 2 propene	224, 245
22	ZOH + 2-Me-C5[b] ⇄ ZOH + 2 propene	210, 219
23	ZOH + 2-Me-C5[b] ⇄ ZOH + 3-Me-C5[a]	179, 176

[a] 3-methyl-pentene. [b] 2-methyl-pentene.

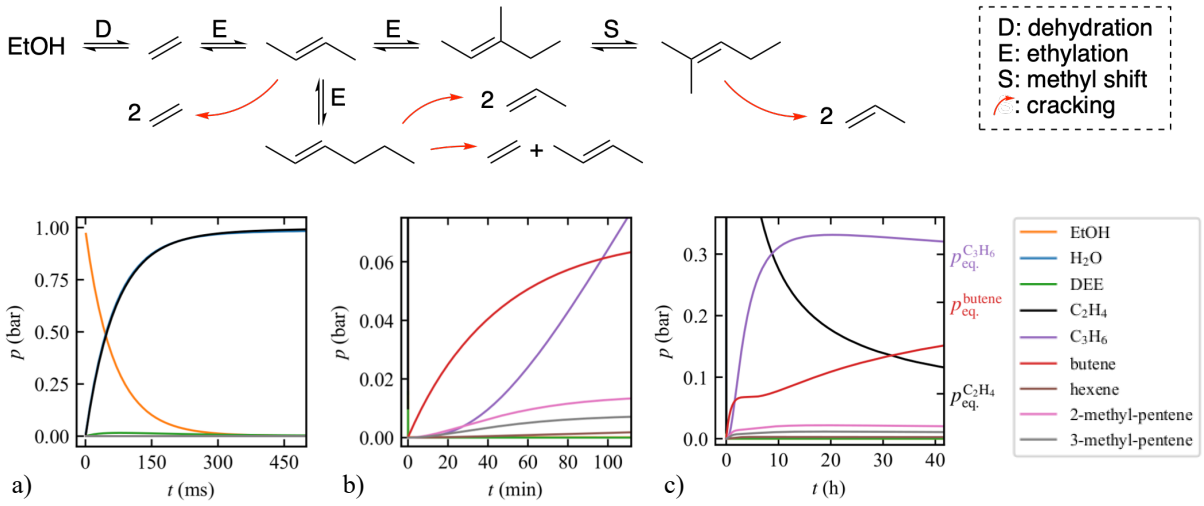


Figure 4.7.: The upper mechanism is a subset of the ETO reaction network with even-numbered olefins (ethene, butene and hexene isomers). The even-numbered olefins are self-sustained by ethylations (E) unless hexene isomers are cracked to propene leading to the existence of odd-numbered olefins. The three diagrams depict the evolution of partial pressures as a function of time in a batch reactor simulation with the kinetic model combined from Tables 4.1, 4.2 and 4.3 at 673.15 K with an initial pressure of 1 bar EtOH. The crucial events happen on different time scales, namely (a) the conversion of EtOH to ethene within 300 ms, (b) the formation of higher olefins with propene partial pressure surpassing butene after 97 min and (c) the slow approach towards the equilibrium that is not completed even after 40 h (equilibrium pressures are indicated). Note that further reactions of propene such as the formation of pentene through ethylation or the formation of heptene from propene and butene are not included in our simulation. Surface coverages are rather low at 673.15 K and not depicted here. Reprinted with permission from [245]. Copyright 2022 The Royal Society of Chemistry (CC-BY).

At lower temperatures, the formation of olefins higher than ethene quickly becomes unfeasible due to the high overall ethylation barriers. This is in line with experimental studies on zeolite H-ZSM-5, where the selectivity for ethene formation was 95.5 % at 543.15 K with almost no formation of higher olefins.[307] Only at temperatures above 573.15 K, significant amounts of higher olefins were observed, leading to gradually increased concentrations of butene and propene through cracking.[307] Although our calculations were performed for H-SSZ-13 and not H-ZSM-5, it is known that significant differences

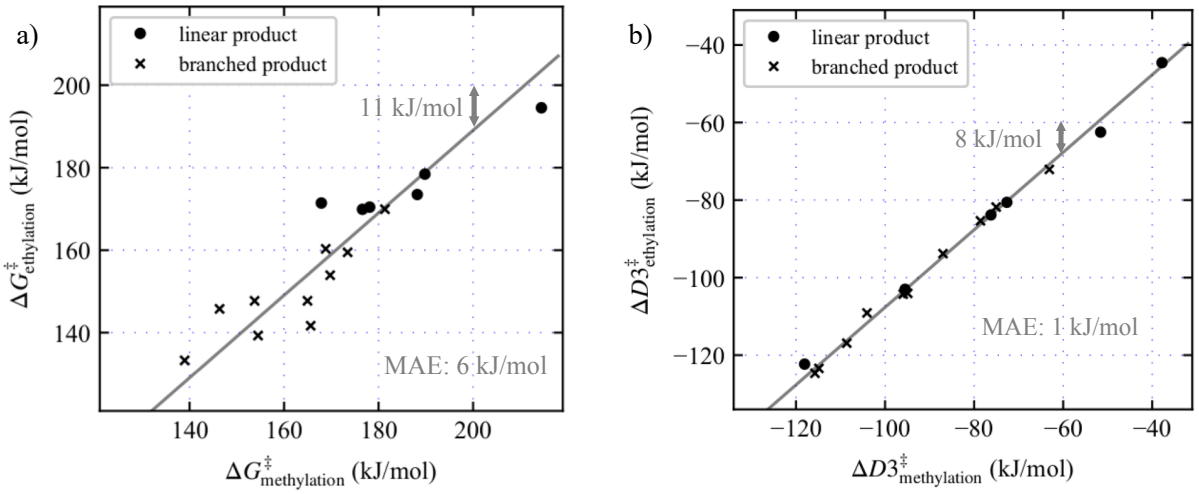


Figure 4.8.: Correlation of (a) free energy barriers and (b) their vdW (D3) contributions, referenced to the gas phase, for the second step of the stepwise alkylation between ETO and MTO processes at 673.15 K and at a reference pressure of 1 bar. The formation of linear and branched products is distinguished. The barriers for ethylation scale linearly with the barriers for methylation with a slope of one and an offset of 11 kJ mol^{-1} with a mean absolute error (MAE) of 6 kJ mol^{-1}. The corresponding D3 contributions scale linearly with a slope of one and an offset of 8 kJ mol^{-1} with a MAE of 1 kJ mol^{-1}. Reprinted with permission from [245]. Copyright 2022 The Royal Society of Chemistry (CC-BY).

in activation barriers are typically transferable to other zeotype frameworks,[24,308] which we have also observed in Section 4.1 where zeolite H-SSZ-13 was compared to zeotype H-SAPO-34.

An in-depth reaction flow analysis shows that the largest contributions come from reactions no. 2, 4, 7, 11, 17, 22 and 23 in Tables 4.1, 4.2 and 4.3.[245] Notably, the majority of relevant reactions proceed via SES. In this key aspect the ETO process is comparable to the MTO process where most reactions of the olefin (and aromatic) cycle proceed via SMS under typical reaction conditions.

4.2.2.2. Correlations between Ethylation and Methylation

Above we have noticed similarities between ETO and MTO in that the predominant alkylation pathways are mediated by surface alkoxy species (SAS). In the following, we compare free energy barriers in zeolite H-SSZ-13 for the second step of the stepwise alkylation, referenced to the gas phase, for a few olefin methylations and ethylations where we differentiate between linear and branched products. This correlation of intrinsic alkylation barriers for ETO and MTO at 673.15 K is presented in Figure 4.8a. We see a correlation between methylation and ethylation barriers, with ethylation barriers being lower by about 11 kJ mol^{-1}.

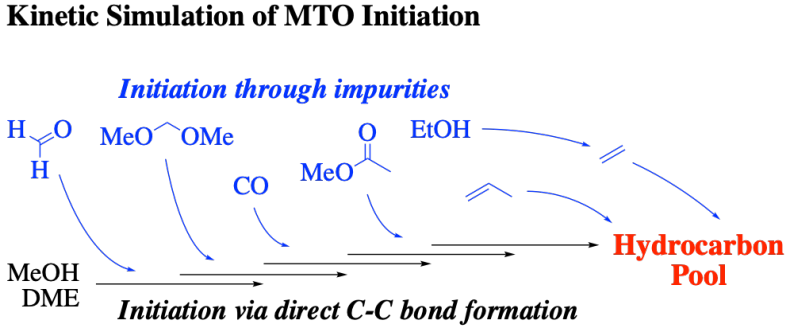
In earlier work on MTO by Plessow and Studt,[24] it was observed that the van-der-Waals (vdW) contribution to reaction barriers increases with decreasing zeolite pore size and that for H-SSZ-13 a contribution of roughly −5.8 kJ mol^{-1} per adsorbate atom can be expected. With the additional CH_2 fragment in ethylations, we would therefore expect a vdW contribution to the energies of the barriers of about −17.4 kJ mol^{-1}. We indeed observe that the major part of the energetic differences of the reaction barriers (−8 kJ mol^{-1}) originates from the difference in vdW contributions (see Figure 4.8b), albeit it is only half as strong as expected from earlier work. Since vdW interaction generally depends on the specific fitting of the adsorbate and the microporous structure, different zeolites can be expected to differ in their vdW interactions with adsorbates, although on average, interaction increases with the size of the adsorbate.[24,308]

4.2.3. Intermediate Summary and Conclusions

The formation of the HCP in the ETO process was studied based on batch reactor kinetics using *ab initio* calculated rate constants for zeolite catalyst H-SSZ-13. Ethene ethylation to butene and hexene isomers was observed to be significantly slower than ethene formation by EtOH dehydration. Because butene ethylation is faster than ethene ethylation at 673.15 K, the ethylation of ethene to butene leads to the formation of hexene isomers opening cracking pathways to the formation of odd-numbered olefins (propene). Although the importance of cracking increases for higher olefins,[56] our simulations predict that propene is mainly formed through cracking of 2-methyl-pentene, a finding that has also been proposed based on experiments.[297] Furthermore, our study revealed a mechanistic similarity between ETO and MTO processes, that is the alkylation mediated by SAS being favored over direct alkylation with the alkanol or alkylation through the dialkyl ether. A correlation of Gibbs free energy barriers for alkylations mediated by SAS shows that ethylation barriers are generally lower than their methylation counterparts by roughly 11 kJ mol^{-1}, which we assign to differences in vdW interactions. Our study thus provides a theoretically derived reaction mechanism for the homologation and cracking of olefins in the ETO process.

4.3. Methanol-to-Olefins – The Role of Impurities

Preface. The initiation of the Methanol-to-Olefins (MTO) process from methanol (MeOH) and its condensation product dimethyl ether (DME) is not as straightforward as the initiation of the Ethanol-to-Olefins (ETO) process from ethanol (EtOH) and diethyl ether (DEE) presented in the previous section. During MTO initiation the first C–C bond needs to be formed in absence of autocatalytically active olefins. Here we quantify the relevance of a selection of organic impurities for the initiation of the MTO process using a kinetic model that comprises 107 elementary steps with *ab initio* computed reaction barriers. This kinetic model adopted from earlier work[56,80,268] includes a representative part of the autocatalytic olefin cycle as well as a direct initiation mechanism starting from MeOH through CO-mediated direct C–C bond formation. We find that the effect of different impurities on the olefin evolution varies with the type of impurity and with their partial pressures. The considered impurities, see Scheme 4.5, increasingly enhance the initiation in the order formaldehyde < dimethoxy methane < CO < methyl acetate < EtOH < ethene < propene. In the kinetic model, already extremely low quantities of impurities such as EtOH lead to faster initiation than through direct C–C bond formation which only matters in complete absence of impurities.



Scheme 4.5.: Simplified scheme for the initiation of the hydrocarbon pool from MeOH and DME with entry points of several impurities considered in this work. Reprinted with permission from [246]. Copyright 2021 Springer Nature (CC-BY).

4.3.1. Overview and Simulation Details

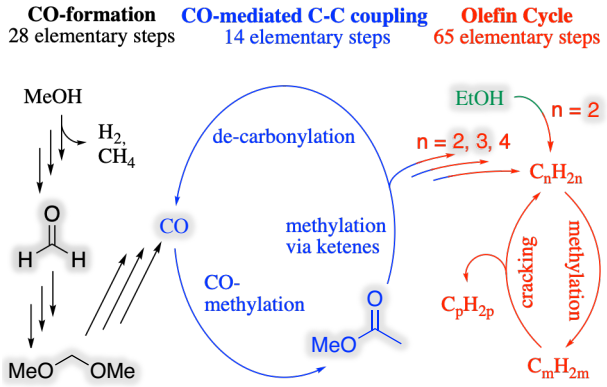
In a sustainable methanol economy the methanol-to-olefins (MTO) process enables the production of hydrocarbons through the formation of C–C bonds from potentially renewable feed stocks.[26] The industrially applied process is operated with porous acidic zeolite catalysts at temperatures between 350 °C to 400 °C.[27] Essentially, the acidic sites of the catalyst transform methanol (MeOH) and its condensation product dimethyl ether (DME) into a variety of olefins. The product distribution is governed both by the intrinsic catalyst properties such as acid strength, acid site density and catalyst topology (framework and shape confinement) as well as by the overall process conditions.[16,42–45]

Advances in the understanding of the mechanistic details of the process lead to the establishment of the hydrocarbon pool (HCP) concept.[30–33,35] Extensive experimental and theoretical studies have described a dual-cycle concept comprising the olefin and the aromatic cycle, see Section 1.2.2. In the former, olefins are formed by repeated alkene methylation[38–40] and cracking,[16,27] whereas the latter comprises the methylation of aromatics producing olefins via either the side chain or pairing mechanism.

Density functional theory (DFT) calculations are increasingly employed to gain insight into the reactivity of zeolites,[64,219,299,309–311] for methanol dehydration[51,55,266,305] as well as for subsequent steps in the MTO process involving the reaction of olefins and aromatics.[23,37,269,298,300–303,312] Kinetic models have been developed to study initiation, stepwise methylation via surface methoxy species (SMS), concerted methylation, cracking and deactivation through coking revealing many features of the process.[21,313–316] Perhaps the most vividly debated issue relates to the early stages of the MTO process and the initial formation of the HCP, that is the formation of the first C–C bond in absence of any olefins.[16,33,44,290–294,299,304,317–337] One of the major challenges for the experimental investigation of the initiation mechanism lies in the avoidance of organic impurities that could be present in the reaction feed as well as during the preparation of the zeolite, as they have been proposed to be responsible for initiating the HCP.[44,289,332]

Theoretical studies have been employed to calculate free energy barriers of possible initiation and C–C coupling steps to shed light on plausible reaction mechanisms.[24,80,268,290,291,293,320–328] Using H-SSZ-13 as the catalyst, a viable mechanism of the initiation reaction was recently proposed based on DFT calculations that have been corrected using second order Møller Plesset perturbation theory (MP2)[56] thus comprising highly accurate reaction barriers.[61] This is crucial since non-hybrid DFT calculations, particularly those using PBE-D3,[60,96] give rise to significantly underestimated activation barriers, effectively making reaction mechanisms appear more favorable than they actually are.[61,268] The identified key steps for the initiation reaction starting from MeOH or DME are (1) the oxidation of MeOH to formaldehyde (FA) and subsequent oxidation to carbon monoxide (CO) and (2) the methylation of CO to methyl acetate (MA) forming the first C–C bond (see Scheme 4.6).[268] Within this reaction sequence, the highest free energy barrier at a reaction temperature of 673 K has been identified for the oxidation of FA to CO ranging from 250 to 275 kJ mol^{-1}, depending on the zeolite employed.[24] In previous studies, the CO methylation pathway has been identified as the most favorable initiation pathway in absence of olefins.[268,293] Therefore, we only consider this pathway here and do not include other mechanisms such as the carbene, oxonium ylide or methane-formaldehyde mechanisms.

As noted above, advance theoretical methods are required to obtain accurate activation barriers. Furthermore, the interpretation of activation barriers is particularly complicated in the context of the initiation due to the autocatalytic nature of the olefin cycle. Therefore, the effect of the above described reaction energetics for H-SSZ-13 on the mechanism of the initiation reaction has been investigated in earlier work by Plessow et al.[80] using a multiscale modeling approach employing a kinetic batch



Scheme 4.6.: Key steps in the zeolite catalyzed initiation of the MTO initiation and subsequent olefin cycle.[56,268] Both stepwise and concerted methylations are considered. Potential impurities and co-feeds investigated in this work are highlighted in gray. Reprinted with permission from [246]. Copyright 2021 Springer Nature (CC-BY).

reactor model comprising both the initiation as well as a substantial part of the autocatalytic olefin cycle. The simulated kinetics showed that the proposed pathway (Scheme 4.6) is indeed viable and the most likely mechanism. The simulations also showed, however, that the initiation mechanism is only relevant in the very early stages with no to little olefins present, as the autocatalytic cycle becomes the dominant source of olefin formation already at very small pressures of olefins.

In this work we quantify the effect of a variety of impurities in the reaction feed as well as typical experimental co-feeds on the olefin formation light-off and mechanism by employing the existing kinetic batch reactor model[80] using *ab initio* rate constants for the initiation[268] and olefin cycle.[56] We also consider the dehydration of EtOH to ethene at the same level of theory (MP2:DFT).

4.3.1.1. Simulation Details

The rate constants were calculated using harmonic transition state theory according to eq. (4.5) with $\Delta G^{\ddagger}$ being extracted from previously computed data[56,268] which used a hierarchical cluster approach[70,71] employing periodic density functional theory (DFT) calculations using the dispersion-corrected PBE-D3[60,96] density functional and the VASP code with the standard PAWs and an energy cutoff of 400 eV.[196–198]

$$k = \frac{k_B T}{h} \cdot \exp\left(-\frac{\Delta G^{\ddagger}}{k_B T}\right) \tag{4.5}$$

Convergence criteria of $1 \cdot 10^{-8}$ eV and 0.01 eV Å^{-1} were applied to SCF cycles and geometry optimization. The Brillouin zone was sampled at the Γ-point only[200] using Gaussian smearing with a width of 0.1 eV. Transition structures were optimized using automated relaxed potential energy surface scans[97] and confirmed to contain one imaginary frequency whose transition vector leads to the correct minima of the reaction. The partial Hessian matrix was computed using a central finite difference scheme including only the adsorbate, the acid site and its adjacent Al- and Si-atom. Energy corrections have been applied with highly accurate non-periodic MP2/def2-TZVPP calculations[278,338–340] on 46T cluster models employing Turbomole.[276] Entropic contributions to the free energy barriers have been calculated using the harmonic approximation at a temperature of 673.15 K and a reference pressure of 1 bar where vibrational frequencies of the adsorbate below 10 cm^{-1} were raised to this value because they can lead to inaccurate entropies otherwise.[21] Additional rate constants for the EtOH dehydration to ethene were computed using the exact same procedure on the same unit cell of zeolite H-SSZ-13.

While the energy contribution to the free energy is quite accurate as it is computed with MP2 calculations on cluster models, entropic contributions are of limited accuracy as they are derived from the harmonic approximation based on DFT calculations. As known from Section 3.2, the harmonic approximation is particularly problematic for loosely bound adsorbates in the zeolite,[219,309] whereas molecules in the gas phase are expected to be described accurately. We therefore expect that activation energies and adsorption energies may in some cases be overestimated, when referenced to the gas phase. In future studies, we expect great improvements in this regard when using the thermodynamic integration procedure introduced in Chapter 3 combined with on-the-fly machine learning enabling accelerated and accurate MD simulations. Another approximation that we employ is to neglect diffusion limitations, which should generally slow down reactions, particular those involving relatively large molecules.[57] An additional limitation of our study is that the reaction network is incomplete, although it contains more than 100 elementary reactions. For example, aromatic molecules are not present. Although aromatic molecules cannot be formed before the first C–C bond during initiation, it cannot be excluded that other relevant reaction pathways are missing. Lastly, we focus on Brønsted acid sites, as they are present in the pores of the zeolite. However, other active site motifs such as surface Brønsted and Lewis acid sites,[341–344] as well as extra-framework aluminum sites[325] may also play an important role.

The kinetic model comprises stepwise and direct methylation with DME and MeOH as well as cracking up to 2-nonene. Olefin isomerization barriers that only shift the double bond (for example 1-butene to 2-butene) are expected to be small compared to methylation and cracking barriers[56] and are therefore treated implicitly. The kinetic simulations with the batch reactor model were carried out based on the mean field approximation using simple Euler integration with a maximum time step smaller than $5 \cdot 10^{-8}$ s as implemented in the in-house code used in earlier work by Plessow et al.[80] Consistently with that earlier work, the reference concentration of the active site (ZOH) in zeolite H-SSZ-13 was set to 17.9 mol m^{-3}. Since we only consider the initial part of the MTO reaction mechanism we neglect diffusion limitations in the reactor simulations, thus assuming perfect mixing. Further details on the kinetic model can be found in the earlier work by Plessow et al.[80]

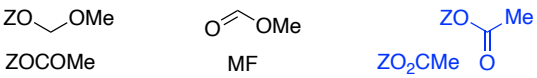
4.3.2. Results and Discussion

Methanol (MeOH) synthesized via $Cu/Zn/Al_2O_3$ catalysts from syngas (CO, CO_2, and H_2O) contains trace amounts of dimethyl ether (DME), methane, ethanol (EtOH), ethane and propane.[345] Commercially available MeOH[346] contains ≤0.05 % EtOH, ≤0.05 % water, ≤0.001 % acetone, ≤0.001 % acetaldehyde and ≤0.001 % formaldehyde (FA). Impurities that are commonly considered to play an important role for the formation of the C–C bond during the initiation phase or within the autocatalytic olefin cycle and that we consider in the present study are: propene (C_3H_6), ethene (C_2H_4), EtOH, methyl acetate (MA), carbon monoxide (CO), dimethoxy methane (DMM) and FA. In total, our kinetic batch reactor model comprises rate constants of 107 (*ab initio*) computed elementary reaction steps for the CO-formation, the CO-mediated C–C coupling and the olefin cycle as shown in Scheme 4.6. We consider both stepwise methylation mediated by surface methoxy species (SMS) as well as concerted methylation via MeOH or DME up to 2-nonene. The free energy barriers of the most important reaction steps that involve the considered impurities are listed in Table 4.4. A complete list of all reaction barriers can be found in Amsler et al.[246] The computational procedure to obtain these barriers at the MP2:DFT level of theory is described in Section 4.3.1.1.

All kinetic simulations were carried out in zeolite H-SSZ-13 starting from a mixture of MeOH, DME and H_2O that was equilibrated from 1 bar of pure DME at 673.15 K. Adsorption processes and SMS formation were included in the equilibration, see Amsler et al.[246] We stress that our focus lies solely on the olefin cycle of the HCP without considering aromatics formation and the aromatics cycle, as we

Table 4.4.: Selection of the most important reaction steps directly linked with impurities considered in this work. Free energy barriers in zeolite H-SSZ-13 are listed for the forward reaction. A complete collection of all free energy barriers can be found in Amsler et al.[246] Potential impurities and co-feeds investigated in this work are highlighted in gray. The reaction free energies for EtOH dehydration to ethene (steps no. 43 and 44) have been calculated in this work using the identical computational procedure and setups as described in,[268] all other free energy barriers have been taken from earlier work by Plessow and Studt.[56,268] Reprinted with permission from [246]. Copyright 2021 Springer Nature (CC-BY).

No.	Elementary reaction		$\Delta G^{\ddagger}$ $\left(\frac{\text{kJ}}{\text{mol}}\right)$
3	ZOH*MeOH	⇄ SMS + H_2O	169
5	SMS + MeOH	⇄ ZOH*DME	151
11	ZOH*MeOH	⇄ ZOH + H_2 + FA	226
12	ZOH*MeOH + DME	⇄ ZOH*MeOH + CH_4 + FA	228
13	ZOH*MeOH + MeOH	⇄ ZOH*H2O + CH_4 + FA	275
14	SMS + MeOH	⇄ ZOH + CH_4 + FA	237
15	ZOH + FA	⇄ ZOH + CO + H_2	251
16	SMS + FA	⇄ ZOCOMe	173
19	ZOCOMe + MeOH	⇄ ZOH + DMM	119
28	ZOH + MF	⇄ ZOH*MeOH + CO	135
29	SMS + CO	⇄ ZO_2CMe	190
36	ZO_2CMe + MeOH	⇄ ZOH + MA	97
43	ZOH + EtOH	⇄ SES + H_2O	183
44	SES	⇄ ZOH + ethene	163
45	SMS + ethene	⇄ ZOH + propene	176
65	ZOH*MeOH + ethene	⇄ ZOH*H_2O + propene	208



expect the olefin cycle to be the most relevant post-initiation mechanism during and right after the initiation phase.

The outcome of our kinetic simulations is shown in Figure 4.9 where the partial pressures are displayed as a function of time. We simulated a feed without impurities (Figure 4.9a) as well as with 2 mbar of FA (Figure 4.9b), EtOH (Figure 4.9c) and propene (Figure 4.9d).

Our kinetics show the general trends observed experimentally. Olefin production has an initiation phase after which production lights off, for instance, after about 3 s in the case where there are no impurities in the feed gas. We also observe that DME is consumed faster than MeOH, which is due to the fact that methylation of olefins is faster with DME compared to with MeOH.[21,38] Methylation with DME produces MeOH and we therefore observe an increase in MeOH pressure shortly after the light-off, which is also observed experimentally.[27] While FA has little effect on the light-off behavior, the presence of both EtOH and propene greatly reduces the time before olefin formation occurs. EtOH is dehydrated to ethene through which the olefin cycle is started directly. The same analysis with 2 mbar of ethene impurity yields a light-off time of 0.43 s. Therefore, the difference in light-off time between EtOH and propene (0.89 s) is composed of the times needed to convert sufficient amounts of EtOH to ethene (0.58 s) and for subsequent methylation (0.31 s).

We note that the definition of the olefin light-off criteria is somewhat arbitrary. Here we quantify the effect of impurities on the olefin light-off using two measures: the time t_{lightoff} after which the

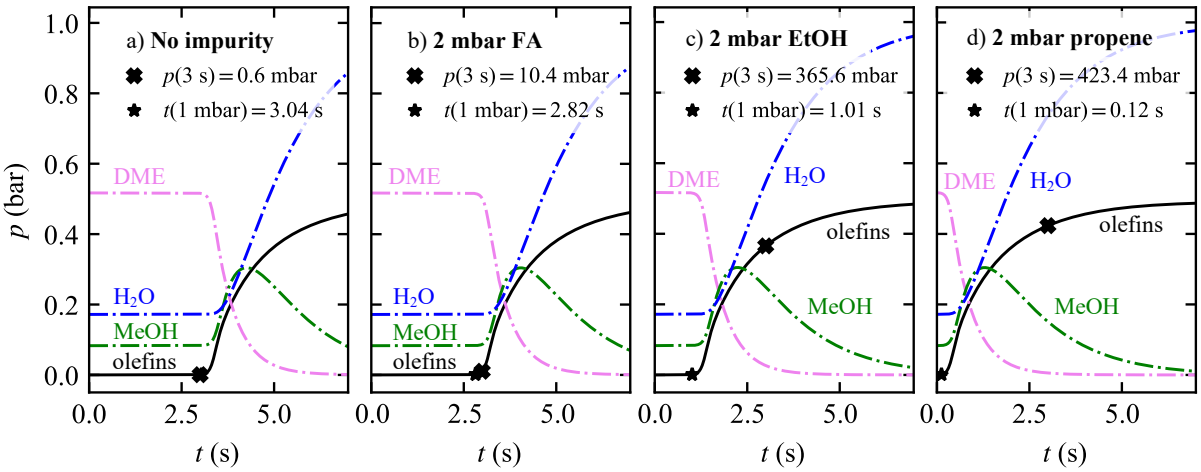


Figure 4.9.: Results of the kinetic batch reactor model: Partial pressure evolution of an equilibrated DME feed with (a) no impurities, (b) 2 mbar of formaldehyde, (c) 2 mbar of ethanol and (d) 2 mbar of propene. The black line shows the total partial pressure of the produced olefins. Two light-off criteria for olefin production are highlighted: the time after which the produced olefins surpass 1 mbar and their total partial pressure after 3 s. Adapted and reprinted with permission from [246]. Copyright 2021 Springer Nature (CC-BY).

total pressure of the produced olefins surpasses 1 mbar (eq. (4.6)) and the olefin production after 3 s (eq. (4.7)).

$$p(t_{\text{lightoff}}) = \sum_{i}^{\text{olefins}} p_i(t_{\text{lightoff}}) - p_i(t_0 = 0\,\text{s}) = 1\,\text{mbar} \tag{4.6}$$

$$p(t = 3\,\text{s}) = \sum_{i}^{\text{olefins}} p_i(t = 3\,\text{s}) - p_i(t_0 = 0\,\text{s}) \tag{4.7}$$

The relation between the amount of impurity and its effect on the olefin production is shown in Figure 4.10. Figure 4.10a shows that all considered impurities shorten the initiation period t_{lightoff} to some extent but vary largely in the observed effect. Interestingly, over a wide pressure range we observe a linear dependence of the light-off time on the logarithm of the initial partial pressure of the impurity. The dependence of our second light-off measure on the initial partial pressure of the impurity is depicted in Figure 4.10b. The dashed horizontal line in Figure 4.10a and b at 3.04 s and $6.32 \cdot 10^{-4}$ bar, respectively, represents the limit for a pure feed without impurities where all initial olefin contributions arise from the initiation reaction. The dotted lines crossing the dashed line show the simulation with excluded initiation mechanism that entirely depends on autocatalysis through impurities. Both light-off measures allow for a ranking of impurity effectiveness. The onset of the olefin formation benefits from these impurities in descending order: C_3H_6 > C_2H_4 > EtOH > MA > CO > DMM > FA. This ranking is plausible for several reasons. Propene and ethene are olefins and thus part of the HCP which means autocatalysis is active from the beginning. In general, cracking barriers decrease for higher alkenes. Propene is one methylation step closer to these higher alkenes than ethene and thus more effective. Dehydration of EtOH leads to ethene thereby circumventing the need for initial C–C bond formation. The ranking of the other impurities can be explained in terms of how advanced they are in the initiation mechanism. In particular, MA and CO are one oxidation step ahead of DMM and FA.

We now discuss the cause of the linear decrease of t_{lightoff} with increasing logarithm of p_{impurity} and provide an explanation with two model reactions for initiation and autocatalysis of first (eq. (4.8)) and second (eq. (4.9)) order, respectively.

$$\text{A} \longrightarrow \text{B} \tag{4.8}$$

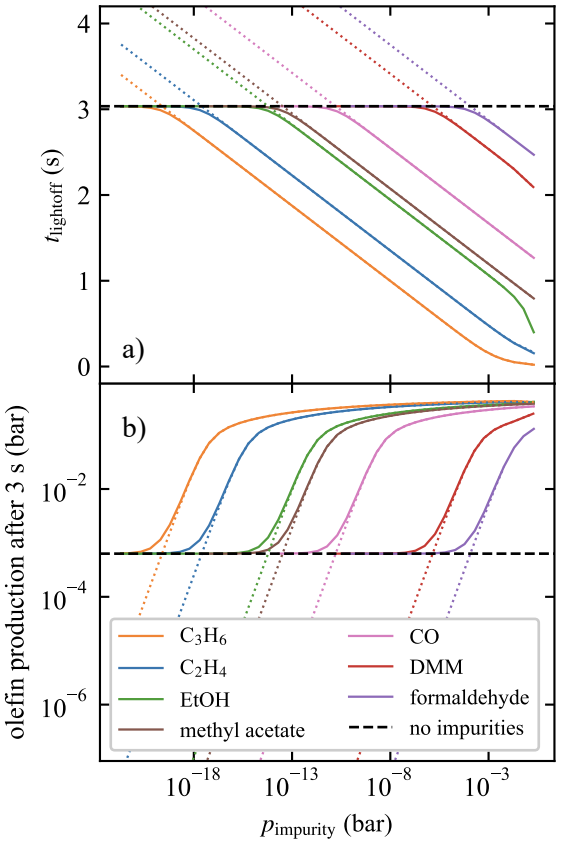


Figure 4.10.: Simulated effect of impurities on olefin formation with (solid lines) and without (dotted lines) consideration of the initiation mechanism using a kinetic batch reactor model employing H-SSZ-13 as a function of the corresponding impurities. The dashed line marks the limit for a pure reactant mixture with the initiation mechanism included. (a) The light-off time is defined as the time when the produced olefins surpass a pressure of 1 mbar. (b) The light-off pressure is defined as the olefin pressure produced after 3 s of simulation time. Reprinted with permission from [246]. Copyright 2021 Springer Nature (CC-BY).

$$\mathrm{A} + \mathrm{B} \longrightarrow 2\,\mathrm{B} \tag{4.9}$$

Here, A represents a C_1 species (MeOH, DME or SMS) and B represents an autocatalytically active species from the HCP which also includes the impurity. Assuming that the rate of autocatalysis (eq. (4.9)) is significantly faster than that of initiation (eq. (4.8)) the latter can be neglected and the integrated rate expression of formal kinetics for a second order autocatalytic reaction[347] leads to eq. (4.10).

$$\frac{1}{p_0^{\mathrm{A}} + p_0^{\mathrm{B}}} \ln\left(\frac{p_0^{\mathrm{B}} p^{\mathrm{A}}}{p_0^{\mathrm{A}} p^{\mathrm{B}}}\right) = -k_{\text{rate}} \cdot t \tag{4.10}$$

Assuming $p_0^{\mathrm{A}} \gg p_0^{\mathrm{B}}$ and $p_0^{\mathrm{A}} \approx p^{\mathrm{A}}$ in eq. (4.10), the second order autocatalytic reaction in eq. (4.9) is of pseudo first order, which is expressed in eq. (4.11).

$$\frac{1}{p_0^{\mathrm{A}}} \ln\left(\frac{p_0^{\mathrm{B}}}{p^{\mathrm{B}}}\right) = -k_{\text{rate}} \cdot t \tag{4.11}$$

Recalling the definition of t_{lightoff} being the time when $p^{\mathrm{B}}(t_{\text{lightoff}}) = p_0^{\mathrm{B}} + 1\,\text{mbar}$ and assuming $p_0^{\mathrm{B}} \ll 1\,\text{mbar}$ leads to eq. (4.12).

$$\frac{1}{p_0^{\mathrm{A}}} \ln\left(\frac{p_0^{\mathrm{B}}}{1\,\text{mbar}}\right) = -k_{\text{rate}} \cdot t_{\text{lightoff}} \tag{4.12}$$

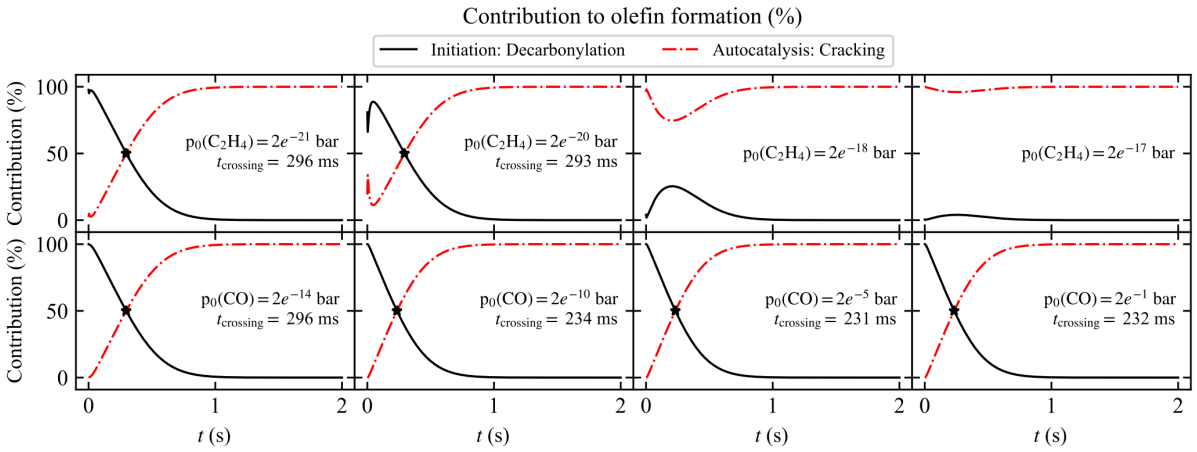


Figure 4.11.: The relative contribution of initiation and autocatalysis to olefin production is shown as a function of reaction time for varying impurity levels of ethene (upper panel) and CO (lower panel). Reprinted with permission from [246]. Copyright 2021 Springer Nature (CC-BY).

Eventually, assuming the impurity to be equivalent to the autocatalytically active species ($p_0^{\mathrm{B}} = p_0^{\mathrm{impurity}}$) leads to $t_{\mathrm{lightoff}} \propto -\log\left(p_0^{\mathrm{impurity}}/\mathrm{mbar}\right)$. Under these assumptions, we can divide the dependence on the impurity partial pressure into three regimes: (i) the regime, where the initiation reaction (4.8) dominates and the solid curves in Figure 4.10 approach the dashed line representing the limit light-off for a pure mixture without impurities, (ii) the regime where autocatalysis dominates and the assumptions of pseudo first order hold and (iii) the regime where autocatalysis dominates but the linear relationship is lost because the assumptions do not hold anymore at higher amounts of HCP species.

The contribution of initial C–C bond formation to olefin formation is investigated in Figure 4.11. The relative contribution of the initiation to the olefin production becomes negligible at extremely low concentrations of ethene. Even with ethene impurities as low as $2 \cdot 10^{-18}$ bar the olefin formation is dominated by autocatalysis from the beginning. This can be understood when comparing the calculated free energy barriers of the olefin methylation with that of the highest step of the oxidation of MeOH to CO. Stepwise olefin methylation has free energy barriers on the order of 130 kJ mol^{-1} to 170 kJ mol^{-1}.[246] In stark contrast the highest free energy barrier for the formation of CO from MeOH is 251 kJ mol^{-1} (see reaction no. 15 in Table 4.4). These differences in the order of 100 kJ mol^{-1} translate to differences in the rate constants of about seven to eight orders of magnitude at 673.15 K. In addition, olefin methylation is autocatalytic in itself, increasing the overall rate with which olefins are formed.

Figure 4.11 also illustrates that, although the effect of CO on olefin production is clearly large (see Figure 4.10), this is not due to the direct formation of olefins from CO, but due to the faster initiation of the autocatalytic olefin cycle. It may appear puzzling at first that the time after which the rates for olefin production from initiation and autocatalysis cross over hardly changes in Figure 4.11, when the CO pressure is increased from $2 \cdot 10^{-10}$ bar, over $2 \cdot 10^{-5}$ bar to 0.2 bar. However, the absolute value of the rates at this crossover does increase dramatically, leading to the reduced light-off times as shown in Figure 4.10.

4.3.3. Intermediate Summary and Conclusions

We have analyzed the effect of various impurities on the initiation of the MTO process based on batch reactor kinetics using *ab initio* calculated rate constants. Initiation of the autocatalytic olefin cycle

through these impurities competes with initiation through direct C–C bond formation starting from MeOH. Direct C–C bond formation proceeds through sequential oxidation of the carbon in MeOH (oxidation state -II), over FA (oxidation state 0) to CO (oxidation state +II). CO can then be methylated to form the first C–C bond and subsequently olefins.

The effect of impurities can be directly related to their role in either initiation or olefin cycle. The olefins propene and ethene have clearly the largest effect, since they can directly initiate the olefin cycle. MA and CO are somewhat less active since a few more reaction steps are still required to convert them into olefins. The effect of EtOH, which enters the HCP via dehydration to ethene, is comparable to that of MA. FA and DMM are least active, which is due to the fact that an oxidation step is required to form CO.

We conclude the more active impurities such as EtOH or MA suffice to be present in sub ppb amounts to allow the initiation of the olefin cycle, without the need for a direct C–C bond forming initiation mechanism.

Part III.

Concluding Remarks

5. Summary and Outlook

This work presented a new method to quantify anharmonicity on the molecular scale and simulations for reactor models with kinetic parameters derived from fundamental calculations on atomistic models. As illustrated in Figure 5.1, several computational methods ranging from density functional theory (DFT) up to ideal reactor models were employed to illuminate the catalytic properties of zeolite H-SSZ-13 – and comparisons to other zeotype catalysts, in particular H-SAPO-34, were discussed.

Based on molecular dynamics (MD) simulation, thermodynamic λ-path integration (λ-TI) was presented as a tool to compute anharmonic corrections to the harmonic oscillator approximation for molecular and periodically extended systems. The significant limitation of the λ-TI approach traditionally formulated in Cartesian coordinates was overcome by using rotationally and translationally invariant curvilinear coordinates. The resulting procedure was demonstrated in three case studies with semi-analytical solutions, where hindered rotations and translations (dissociation) were identified to cause significant anharmonic contributions to the free energy (translation more than rotation) while the anharmonic contributions to vibrations of covalent bonds seemed rather negligible. In general, the poorer a system is described by the harmonic oscillator approximation, the larger anharmonicity is to be expected. Moreover, the more the interactions and degrees of freedom in two compared states differ, the more anharmonicity affects the free energy difference between them. This is particularly relevant

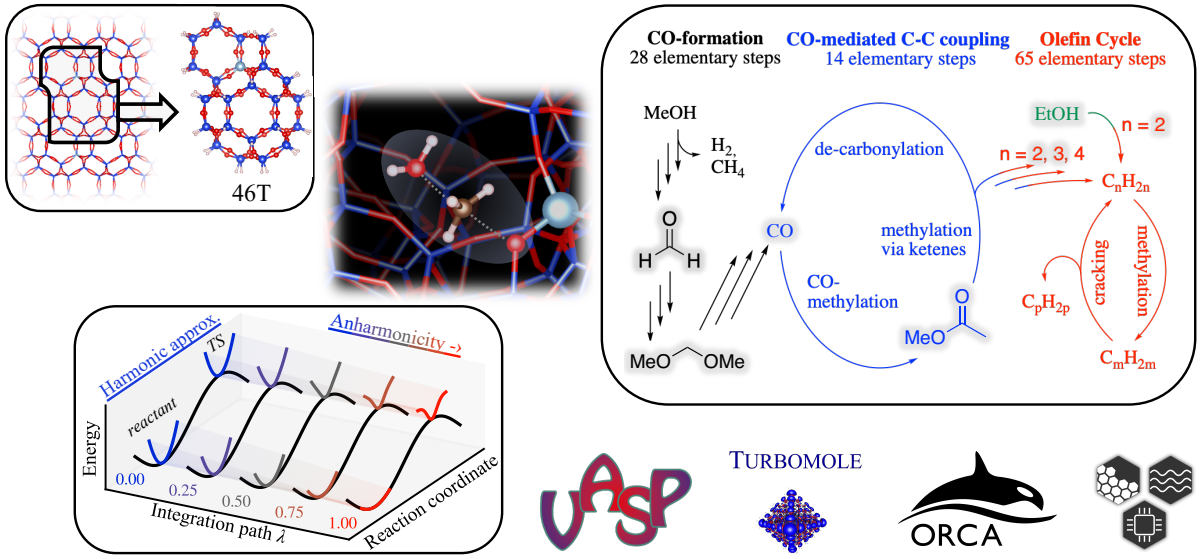


Figure 5.1.: The center image depicts the transition structure for the formation of surface methoxy species at the acidic site of zeolite H-SSZ-13. A 46T cluster model, as used for some of the quantum mechanical calculations, is shown in the upper left box. Representing the reaction networks in this work, the box on the right contains the initiation mechanism and olefin cycle of the MTO process. Particular interest was paid to the role of molecules highlighted in gray. The extent of anharmonicity in adsorption processes and reactions was investigated with the newly developed λ-TI method summarized in the lower left box. The row on the lower right displays the logos of the licensed simulation software packages employed, namely VASP, Turbomole, ORCA and DETCHEM. Adapted compilation of images with permissions from [17, 246]. Copyright 2021 American Chemical Society and 2021 Springer Nature (CC-BY), respectively.

for adsorption processes due to the prevailing weak interactions for which the estimate of entropy contributions by means of the harmonic approximation is prone to be inaccurate. Therefore, the λ-TI method was then applied to study adsorption free energies of N_2 and Ar in zeolite H-SSZ-13 at 200 K. Remarkably, the prediction of experimental data from literature was improved by the anharmonic corrections from the λ-TI method performed on the DFT level of theory with solely a dispersion-corrected GGA functional (PBE-D2). The adsorption study clearly benefited from the advantage that the λ-TI method is independent of any reaction coordinate ξ, thus requiring simulations of only initial and final states without the necessity to sample over ξ. In contrast, popular but ξ-dependent methods such as Umbrella sampling would turn out extremely impractical for such endeavors since the structural models of the adsorbent would need to incorporate a continuous desorption path up to the desorbed state unperturbed by the surface. With regards to anharmonicity of transition states, the λ-TI method excels as comparable alternative to a ξ-dependent procedure. For these reasons, the presented λ-TI method is perfectly suited to study adsorption free energies or individual rate determining steps in multistep reaction cascades and could thus become a very useful tool for computational catalysis in general. All in all, the λ-TI method integrates effortlessly into existing multiscale workflows, bringing accurate entropy estimations within reach and expanding the frontiers of what is currently possible in computational studies.

Several strategies for improving the λ-TI method in terms of quality, accuracy and effectiveness of sampling were outlined, some with preliminary tests. The ideas include a pre-optimization of the reference Hessian matrix to effectively describe as large part of the anharmonicity of the system as possible aiming to minimize the anharmonic correction by λ-TI in favor of faster convergence. Furthermore, universally applicable coordinates were briefly explored as an alternative to the highly system specific internal coordinates. Two classes of coordinate descriptors – Coulomb-types and descriptors based on spherical harmonics – commonly used in the machine learning (ML) community showed promising transformation properties, that warrant further investigation regarding their general applicability and sampling efficiency, also for more complex systems. On-the-fly ML with built-in error estimation was outlined to be able to alleviate the typically high computational costs. In the future, Bennett's acceptance ratio should also be explored as an alternative to the multistage λ-TI.

The second part of this work reported a detailed investigation of the dehydration of ethanol as well as of the ETO and MTO processes based on a parameter passing multiscale approach from quantum mechanics up to reactor models. Highly accurate quantum methods were applied to 46T cluster models of the zeolite framework optimized by plane-wave DFT in periodic boundary conditions to arrive at microkinetic parameters for batch and plug flow reactor models. Free energy barriers for ethanol dehydration in zeotype H-SAPO-34 were slightly higher compared to zeolite H-SSZ-13. Temporary diethyl ether formation was observed in batch reactor simulations with catalyst H-SSZ-13 at lower temperatures (473.15 K) with decreasing selectivity for increasing conversion. The temperature dependence of this kinetic selectivity was explained by a change in the ranking of free energy barriers. While the batch reactor simulations showed that the selectivity depends on the conversion of ethanol, the plug flow simulations with catalyst H-SAPO-34 showed that the selectivity also varies with the weight hourly space velocity in a regime where the conversion is practically complete. Highly dynamic changes in gas phase composition along the catalyst bed were observed to correlate with the presence of surface ethoxy species, which further strengthens our confidence for its presumed key role in the reaction mechanism. Our study thus provides theoretically derived kinetic parameters for the dehydration of ethanol as well as reactor simulations tailored to the setup of our experimental collaborators. Considering a significant part of the autocatalytic olefin cycle in ETO, the formation of the hydrocarbon pool was explored and studied. Due to lower free energy barriers, ethanol dehydration occurred much more rapidly than ethylation of the formed ethene. Ethylations of ethene and products thereof lead to only even-numbered olefins. Hexene isomers formed on the same time scale as butene, with branched

isomers being favored and 2-methyl-pentene isomers contributing most to the formation of propene through cracking. Surface alkoxy species were found to enable the most relevant (stepwise) alkylation mechanism under the studied reaction conditions in zeolite H-SSZ-13 for both ETO and MTO. In direct comparison, ethylation barriers tend to be slightly lower than the corresponding methylations, which was attributed mainly to differences in van-der-Waals interactions between the zeolite pores and the alkenes. Furthermore, the relevance of a selection of common organic impurities for the initiation of the MTO process was quantified in a kinetic model comprising 107 elementary steps. The model includes a representative part of the olefin cycle and direct initiation from methanol through CO-mediated C–C bond formation. An enhanced formation of olefins was observed depending on both impurity type and amount, in direct relation to their role in either initiation or olefin cycle. The effect increased in the order formaldehyde < dimethoxy methane < CO < methyl acetate < ethanol < ethene < propene. In the kinetic model, already extremely low quantities of impurities lead to faster initiation than direct C–C bond formation, which only matters in complete absence of impurities. In conclusion, already sub ppb amounts of the more active impurities such as ethanol or methyl acetate drive the initiation of the olefin cycle without the need for a direct C–C bond forming initiation mechanism.

Despite most studies in this work were carried out with zeolite H-SSZ-13, the results are largely transferable to other zeotypes under consideration of known trends for changes in acidity and shape confinement known from literature – which was explicitly confirmed for zeotype H-SAPO-34 for ethanol dehydration. In future studies, the now available λ-TI method should be applied to crucial steps in the ETO and MTO mechanisms, for instance, to the formation of the surface alkoxy species. The next logical course of action, after presenting the theoretically derived reaction mechanism for ETO and MTO, would be a continuation of predictions for experiments closer to the desired (industrial) application. For the comparison to such experiments, a careful selection of the reactor model for the simulation is advisable, for instance, experiments with a reactor tube should be simulated with a plug flow reactor rather than a batch reactor. Further considerations could include transport effects such as diffusion limitation or non-isothermal reaction control.

Altogether, this work responded to the introductory call for catalysts with a new computational method and some experimentally hardly observable insights into the hydrocarbon pool of the ETO and MTO processes. With the foundations in this work, I can envision a workflow for computational zeolite catalysis where highly accurate quantum chemical methods are sequentially applied to hierarchical cluster models of zeolite frameworks optimized by DFT – while anharmonic corrections to entropic contributions are considered simultaneously by λ-TI using ML-accelerated DFT-based MD. Noteworthy, this workflow would be generally applicable and interesting for other (catalytic) systems, for instance, for catalysis on metallic surfaces. More accurate free energies also mean more accurate rate constants, when recalling the gargantuan impact of small errors in energy on the reaction rate (see p. 7) every $\mathrm{kJ\,mol^{-1}}$ counts. It remains open, how computational zeolite catalysis is shaped by further insight into structure-performance relationships and by new ways of searching correlations among catalytic systems as well as how these findings impact the chemical and energy industry of the future. In other words, as the logician and philosopher Sir Karl Popper (1902 - 1994) once said,[1] “It is remarkable to what extent nature has been made calculable. There is only one thing we cannot know in principle, namely what we will know in the future. Because otherwise we would already know.”

Bibliography

[1] J. P. Gerling, F.-W. Wellmer, *ChiuZ* **2005**, *39*, 236–245.

[2] *Aktuelle Wirtschaftsanalysen* **1995**, *25*, (Ed.: Deutsche Shell Aktiengesellschaft).

[3] BP p.l.c., Ed., Energy Outlook 2020 Edition, **2020**.

[4] A. Behr, D. W. Agar, J. Jörissen, *Einführung in die technische Chemie*, Spektrum Akademischer Verlag, Heidelberg, **2010**.

[5] M. S. Dresselhaus, I. L. Thomas, *Nature* **2001**, *414*, 332–337.

[6] M. Meinshausen, N. Meinshausen, W. Hare, S. C. B. Raper, K. Frieler, R. Knutti, D. J. Frame, M. R. Allen, *Nature* **2009**, *458*, 1158–1162.

[7] K. Hiller, *Erdöl Erdgas Kohle* **1999**, *115*, 50–54.

[8] P. R. Odell, *Energy Exploration & Exploitation* **2000**, *18*, 131–145.

[9] J. D. Edwards, *AAPG Memoir* **2001**, 21–34.

[10] C. J. Campbell, *ASPO Newsletter* **2002**, *13*.

[11] I. Chorkendorff, J. W. Niemantsverdriet, *Concepts of Modern Catalysis and Kinetics*, First, Wiley, **2003**.

[12] F. Studt, *Front. Catal.* **2021**, *1*, 658965.

[13] *Handbook of Heterogeneous Catalysis*, 2nd, completely revised and enlarged ed, (Eds.: Ertl, Gerhard, Knözinger, Helmut, Schüth, Ferdi, Weitkamp, Jens), Wiley-VCH, Weinheim, **2008**.

[14] X. Chen, D. Xi, Q. Sun, N. Wang, Z. Dai, D. Fan, V. Valtchev, J. Yu, *Microporous Mesoporous Mater.* **2016**, *234*, 401–408.

[15] R. M. Barrer in *Surface Organometallic Chemistry: Molecular Approaches to Surface Catalysis*, (Eds.: J.-M. Basset, B. C. Gates, J.-P. Candy, A. Choplin, M. Leconte, F. Quignard, C. Santini), Springer Netherlands, Dordrecht, **1988**, pp. 221–244.

[16] I. Yarulina, A. D. Chowdhury, F. Meirer, B. M. Weckhuysen, J. Gascon, *Nat. Catal.* **2018**, *1*, 398–411.

[17] J. Amsler, P. N. Plessow, F. Studt, T. Bučko, *J. Chem. Theory Comput.* **2021**, *17*, 1155–1169.

[18] R. A. van Santen, G. J. Kramer, *Chem. Rev.* **1995**, *95*, 637–660.

[19] J. Macht, R. T. Carr, E. Iglesia, *J. Am. Chem. Soc.* **2009**, *131*, 6554–6565.

[20] Y. Chu, B. Han, A. Zheng, F. Deng, *J. Phys. Chem. C* **2012**, *116*, 12687–12695.

[21] R. Y. Brogaard, R. Henry, Y. Schuurman, A. J. Medford, P. G. Moses, P. Beato, S. Svelle, J. K. Nørskov, U. Olsbye, *J. Catal.* **2014**, *314*, 159–169.

[22] C.-M. Wang, R. Y. Brogaard, B. M. Weckhuysen, J. K. Nørskov, F. Studt, *J. Phys. Chem. Lett.* **2014**, *5*, 1516–1521.

[23] C.-M. Wang, Y.-D. Wang, Y.-J. Du, G. Yang, Z.-K. Xie, *Catal. Sci. Technol.* **2015**, *5*, 4354–4364.

[24] P. N. Plessow, F. Studt, *Catal. Lett.* **2018**, *148*, 1246–1253.

[25] V. F. Tret'yakov, Y. I. Makarfi, K. V. Tret'yakov, N. A. Frantsuzova, R. M. Talyshinskii, *Catal. Ind.* **2010**, *2*, 402–420.

[26] G. A. Olah, *Angew. Chem. Int. Ed.* **2013**, *52*, 104–107.

[27] U. Olsbye, S. Svelle, M. Bjørgen, P. Beato, T. V. W. Janssens, F. Joensen, S. Bordiga, K. P. Lillerud, *Angew. Chem. Int. Ed.* **2012**, *51*, 5810–5831.

[28] M. R. Gogate, *Pet. Sci. Technol.* **2019**, *37*, 559–565.

[29] N. M. Eagan, M. D. Kumbhalkar, J. S. Buchanan, J. A. Dumesic, G. W. Huber, *Nat. Rev. Chem.* **2019**, *3*, 223–249.

[30] I. M. Dahl, S. Kolboe, *Catal. Lett.* **1993**, *20*, 329–336.

[31] I. Dahl, S. Kolboe, *J. Catal.* **1994**, *149*, 458–464.

[32] I. M. Dahl, S. Kolboe, *J. Catal.* **1996**, *161*, 304–309.

[33] W. Wang, Y. Jiang, M. Hunger, *Catal. Today* **2006**, *113*, 102–114.

[34] R. Johansson, S. L. Hruby, J. Rass-Hansen, C. H. Christensen, *Catal. Lett.* **2009**, *127*, 1–6.

[35] D. S. Wragg, M. G. O'Brien, F. L. Bleken, M. Di Michiel, U. Olsbye, H. Fjellvåg, *Angew. Chem. Int. Ed.* **2012**, *51*, 7956–7959.

[36] A. D. Chowdhury, A. Lucini Paioni, G. T. Whiting, D. Fu, M. Baldus, B. M. Weckhuysen, *Angew. Chem. Int. Ed.* **2019**, *58*, 3908–3912.

[37] C.-M. Wang, Y.-D. Wang, Z.-K. Xie, *Catal. Sci. Technol.* **2014**, *4*, 2631–2638.

[38] S. Svelle, S. Kolboe, O. Swang, U. Olsbye, *J. Phys. Chem. B* **2005**, *109*, 12874–12878.

[39] F. Bleken, S. Svelle, K. P. Lillerud, U. Olsbye, B. Arstad, O. Swang, *J. Phys. Chem. A* **2010**, *114*, 7391–7397.

[40] E. A. Redekop, A. Lazzarini, S. Bordiga, U. Olsbye, *J. Catal.* **2020**, *385*, 300–312.

[41] D. Lesthaeghe, A. Horré, M. Waroquier, G. B. Marin, V. Van Speybroeck, *Chem. Eur. J.* **2009**, *15*, 10803–10808.

[42] S. Teketel, U. Olsbye, K.-P. Lillerud, P. Beato, S. Svelle, *Microporous Mesoporous Mater.* **2010**, *136*, 33–41.

[43] J. F. Haw, W. Song, D. M. Marcus, J. B. Nicholas, *Acc. Chem. Res.* **2003**, *36*, 317–326.

[44] U. Olsbye, S. Svelle, K. P. Lillerud, Z. H. Wei, Y. Y. Chen, J. F. Li, J. G. Wang, W. B. Fan, *Chem. Soc. Rev.* **2015**, *44*, 7155–7176.

[45] S. Müller, Y. Liu, F. M. Kirchberger, M. Tonigold, M. Sanchez-Sanchez, J. A. Lercher, *J. Am. Chem. Soc.* **2016**, *138*, 15994–16003.

[46] N. Chen, W. Reagan, *J. Catal.* **1979**, *59*, 123–129.

[47] Y. Ono, T. Mori, *J. Chem. Soc. Faraday Trans. 1* **1981**, *77*, 2209.

[48] C. D. Chang, *Catal. Rev.* **1983**, *25*, 1–118.

[49] E. Pidko, E. Hensen in *Zeolites and Zeolite-Like Materials*, Elsevier, **2016**, pp. 111–135.

[50] Y. Jeanvoine, J. G. Ángyán, G. Kresse, J. Hafner, *J. Phys. Chem. B* **1998**, *102*, 5573–5580.

[51] A. Ghorbanpour, J. D. Rimer, L. C. Grabow, *Catal. Commun.* **2014**, *52*, 98–102.

[52] A. T. Smith, P. N. Plessow, F. Studt, *J. Phys. Chem. C* **2021**, *125*, 16508–16515.

[53] A. T. Smith, P. N. Plessow, F. Studt, *J. Phys. Chem. C* **2021**, *125*, 20373–20379.

[54] F. Göltl, P. Sautet, *J. Chem. Phys.* **2014**, *140*, 154105.

[55] A. Ghorbanpour, J. D. Rimer, L. C. Grabow, *ACS Catal.* **2016**, *6*, 2287–2298.

[56] P. N. Plessow, F. Studt, *Catal. Sci. Technol.* **2018**, *8*, 4420–4429.

[57] A. T. Smith, P. N. Plessow, F. Studt, *Chem. Phys.* **2021**, *541*, 111033.

[58] S. Grimme, *J. Comput. Chem.* **2004**, *25*, 1463–1473.

[59] S. Grimme, *J. Comput. Chem.* **2006**, *27*, 1787–1799.

[60] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.

[61] T. J. Goncalves, P. N. Plessow, F. Studt, *ChemCatChem* **2019**, *11*, 4368–4376.

[62] P. N. Plessow, F. Studt, *J. Phys. Chem. Lett.* **2020**, *11*, 4305–4310.

[63] C.-M. Wang, R. Y. Brogaard, Z.-K. Xie, F. Studt, *Catal. Sci. Technol.* **2015**, *5*, 2814–2820.

[64] C. Liu, I. Tranca, R. A. van Santen, E. J. M. Hensen, E. A. Pidko, *J. Phys. Chem. C* **2017**, *121*, 23520–23530.

[65] C.-M. Wang, Y.-D. Wang, Z.-K. Xie, *Catal. Sci. Technol.* **2019**, *9*, 2245–2252.

[66] Á. Szécsényi, E. Khramenkova, I. Y. Chernyshov, G. Li, J. Gascon, E. A. Pidko, *ACS Catal.* **2019**, *9*, 9276–9284.

[67] Y. Chen, X. Zhao, Z. Qin, S. Wang, Z. Wei, J. Li, M. Dong, J. Wang, W. Fan, *J. Phys. Chem. C* **2020**, *124*, 13789–13798.

[68] C. Tuma, J. Sauer, *Phys. Chem. Chem. Phys.* **2006**, *8*, 3955–3965.

[69] C. Tuma, T. Kerber, J. Sauer, *Angew. Chem. Int. Ed.* **2010**, *49*, 4678–4680.

[70] S. Svelle, C. Tuma, X. Rozanska, T. Kerber, J. Sauer, *J. Am. Chem. Soc.* **2009**, *131*, 816–825.

[71] N. Hansen, T. Kerber, J. Sauer, A. T. Bell, F. J. Keil, *J. Am. Chem. Soc.* **2010**, *132*, 11525–11538.

[72] M. Rybicki, J. Sauer, *J. Am. Chem. Soc.* **2018**, *140*, 18151–18161.

[73] Q. Ren, M. Rybicki, J. Sauer, *J. Phys. Chem. C* **2020**, *124*, 10067–10078.

[74] A. Kundu, G. Piccini, K. Sillar, J. Sauer, *J. Am. Chem. Soc.* **2016**, *138*, 14047–14056.

[75] L. H. Sprowl, C. T. Campbell, L. Árnadóttir, *J. Phys. Chem. C* **2016**, *120*, 9719–9731.

[76] M. Jørgensen, H. Grönbeck, *J. Phys. Chem. C* **2017**, *121*, 7199–7207.

[77] H. Li, C. Paolucci, W. F. Schneider, *J. Chem. Theory Comput.* **2018**, *14*, 929–938.

[78] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, A. Aspuru-Guzik, *Chem. Rev.* **2019**, *119*, 10856–10915.

[79] S. Kozuch, S. Shaik, *Acc. Chem. Res.* **2011**, *44*, 101–110.

[80] P. N. Plessow, A. Smith, S. Tischer, F. Studt, *J. Am. Chem. Soc.* **2019**, *141*, 5908–5915.

[81] R. Schlögl, *ChemCatChem* **2017**, *9*, 533–541.

[82] A. Bruix, J. T. Margraf, M. Andersen, K. Reuter, *Nat. Catal.* **2019**, *2*, 659–670.

[83] D. Frenkel, B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed, Academic Press, San Diego, **2002**.

[84] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods*, Cambridge University Press, Cambridge, UK ; New York, **2004**.

[85] C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, 2nd ed, Wiley, Chichester, **2004**.

[86] C. Kittel, *Introduction to Solid State Physics*, 8th ed, Wiley, Hoboken, NJ, **2005**.

[87] F. Jensen, *Introduction to Computational Chemistry*, 2nd ed, John Wiley & Sons, Chichester, England, **2007**.

[88] G. H. Findenegg, T. Hellweg, *Statistische Thermodynamik*, Springer Berlin Heidelberg, Berlin, Heidelberg, **2015**.

[89] B. Bryson, *A Short History of Nearly Everything*, 1st ed, Black Swan, London, **2004**.

[90] C. Riplinger, P. Pinski, U. Becker, E. F. Valeev, F. Neese, *J. Chem. Phys.* **2016**, *144*, 024109.

[91] F. Neese, *WIREs Comput. Mol. Sci.* **2018**, *8*, e1327.

[92] Y. Guo, C. Riplinger, U. Becker, D. G. Liakos, Y. Minenkov, L. Cavallo, F. Neese, *J. Chem. Phys.* **2018**, *148*, 011101.

[93] E. Paulechka, A. Kazakov, *J. Phys. Chem. A* **2017**, *121*, 4379–4387.

[94] D. G. Liakos, M. Sparta, M. K. Kesharwani, J. M. L. Martin, F. Neese, *J. Chem. Theory Comput.* **2015**, *11*, 1525–1539.

[95] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, *Chem. Phys. Lett.* **1995**, *242*, 652–660.

[96] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[97] P. N. Plessow, *J. Chem. Theory Comput.* **2018**, *14*, 981–990.

[98] H. C. Andersen, *J. Chem. Phys.* **1980**, *72*, 2384–2393.

[99] H. Flyvbjerg, H. G. Petersen, *J. Chem. Phys.* **1989**, *91*, 461–466.

[100] *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms: Winter School, 25 February - 1 March 2002, Rolduc Conference Centre, Kerkrade, the Netherlands ; Lecture Notes*, (Eds.: J. Grotendorst, John von Neumann-Institut für Computing), NIC-Secretariat, Jülich, **2002**.

[101] *Free Energy Calculations: Theory and Applications in Chemistry and Biology*, (Eds.: C. Chipot, A. Pohorille), Springer, Berlin, **2007**.

[102] Deutschmann, O. and Tischer, S. and Kleditzsch S. and Janardhanan, V. and Correa, C. and Chatterjee, D. and Mladenov, N. and Minh, H. D. and Karadeniz, H. and Hettel, M. and Menon, V. and Banerjee, A. and Goßler, H. and Daymo E., DETCHEM Software Package, Karlsruhe, **2020**.

[103] J. Amsler, P. N. Plessow, F. Studt, T. Bučko, *J. Chem. Theory Comput.* **2023**, DOI `10.1021/acs.jctc.3c00169`.

[104] E. A. Carter, G. Ciccotti, J. T. Hynes, R. Kapral, *Chem. Phys. Lett.* **1989**, *156*, 472–477.

[105] G. Ciccotti, R. Kapral, E. Vanden-Eijnden, *ChemPhysChem* **2005**, *6*, 1809–1814.

[106] G. M. Torrie, J. P. Valleau, *Chem. Phys. Lett.* **1974**, *28*, 578–581.

[107] G. Torrie, J. Valleau, *J. Comput. Chem.* **1977**, *23*, 187–199.

[108] J. Kästner, *WIREs Comput. Mol. Sci.* **2011**, *1*, 932–942.

[109] T. Bučko, L. Benco, J. Hafner, J. G. Ángyán, *J. Catal.* **2011**, *279*, 220–228.

[110] J. Rey, A. Gomez, P. Raybaud, C. Chizallet, T. Bučko, *J. Catal.* **2019**, *373*, 361–373.

[111] J. Rey, P. Raybaud, C. Chizallet, T. Bučko, *ACS Catal.* **2019**, *9*, 9813–9828.

[112] P. Cnudde, K. De Wispelaere, J. Van der Mynsbrugge, M. Waroquier, V. Van Speybroeck, *J. Catal.* **2017**, *345*, 53–69.

[113] P. Cnudde, K. De Wispelaere, L. Vanduyfhuys, R. Demuynck, J. Van der Mynsbrugge, M. Waroquier, V. Van Speybroeck, *ACS Catal.* **2018**, *8*, 9579–9595.

[114] S. Bailleul, S. M. J. Rogge, L. Vanduyfhuys, V. Van Speybroeck, *ChemCatChem* **2019**, *11*, 3993–4010.

[115] S. Bailleul, K. Dedecker, P. Cnudde, L. Vanduyfhuys, M. Waroquier, V. Van Speybroeck, *J. Catal.* **2020**, *388*, 38–51.

[116] D. L. Beveridge, F. M. DiCapua, *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431–492.

[117] T. P. Straatsma, H. J. C. Berendsen, *J. Chem. Phys.* **1988**, *89*, 5876–5886.

[118] C. A. F. de Oliveira, D. Hamelberg, J. A. McCammon, *J. Chem. Theory Comput.* **2008**, *4*, 1516–1525.

[119] F. Leroy, D. J. V. A. dos Santos, F. Müller-Plathe, *Macromol. Rapid Commun.* **2009**, *30*, 864–870.

[120] A. Pohorille, C. Jarzynski, C. Chipot, *J. Phys. Chem. B* **2010**, *114*, 10235–10253.

[121] S. Bruckner, S. Boresch, *J. Comput. Chem.* **2011**, *32*, 1320–1333.

[122] A. de Ruiter, C. Oostenbrink, *J. Chem. Theory Comput.* **2012**, *8*, 3686–3695.

[123] S. G. Moustafa, A. J. Schultz, D. A. Kofke, *Phys. Rev. E* **2015**, *92*, 043303.

[124] S. G. Moustafa, A. J. Schultz, D. A. Kofke, *J. Chem. Theory Comput.* **2017**, *13*, 825–834.

[125] L. Vočadlo, D. Alfè, *Phys. Rev. B* **2002**, *65*, 214105.

[126] B. Grabowski, L. Ismer, T. Hickel, J. Neugebauer, *Phys. Rev. B* **2009**, *79*, 134106.

[127] O. Hellman, I. A. Abrikosov, S. I. Simak, *Phys. Rev. B* **2011**, *84*, 180301.

[128] A. Glensk, B. Grabowski, T. Hickel, J. Neugebauer, *Phys. Rev. X* **2014**, *4*, 011018.

[129] A. Glensk, B. Grabowski, T. Hickel, J. Neugebauer, *Phys. Rev. Lett.* **2015**, *114*, 195901.

[130] A. I. Duff, T. Davey, D. Korbmacher, A. Glensk, B. Grabowski, J. Neugebauer, M. W. Finnis, *Phys. Rev. B* **2015**, *91*, 214311.

[131] R. Jinnouchi, F. Karsai, G. Kresse, *Phys. Rev. B* **2020**, *101*, 060201.

[132] S. Habershon, D. E. Manolopoulos, *Phys. Chem. Chem. Phys.* **2011**, *13*, 19714.

[133] M. Rossi, P. Gasparotto, M. Ceriotti, *Phys. Rev. Lett.* **2016**, *117*, 115702.

[134] G. Piccini, J. Sauer, *J. Chem. Theory Comput.* **2014**, *10*, 2479–2487.

[135] G. Piccini, M. Alessio, J. Sauer, Y. Zhi, Y. Liu, R. Kolvenbach, A. Jentys, J. A. Lercher, *J. Phys. Chem. C* **2015**, *119*, 6128–6137.

[136] E. B. Wilson, J. C. Decius, P. C. Cross, *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*, Dover Publications, New York, **1980**.

[137] S. G. Moustafa, A. J. Schultz, D. A. Kofke, *J. Chem. Phys.* **2013**, *139*, 084105.

[138] S. G. Moustafa, A. J. Schultz, D. A. Kofke, *J. Chem. Phys.* **2018**, *149*, 124109.

[139] A. Purohit, A. J. Schultz, S. G. Moustafa, J. R. Errington, D. A. Kofke, *Mol. Phys.* **2018**, *116*, 3027–3041.

[140] S. G. Moustafa, A. Purohit, A. J. Schultz, D. A. Kofke, *Comput. Phys. Commun.* **2021**, *258*, 107554.

[141] B. A. De Moor, A. Ghysels, M.-F. Reyniers, V. Van Speybroeck, M. Waroquier, G. B. Marin, *J. Chem. Theory Comput.* **2011**, *7*, 1090–1101.

[142] E. L. Uzunova, H. Mikosch, *ACS Catal.* **2013**, *3*, 2759–2767.

[143] B. A. De Moor, M.-F. Reyniers, O. C. Gobin, J. A. Lercher, G. B. Marin, *J. Phys. Chem. C* **2011**, *115*, 1204–1219.

[144] G. Piccini, M. Alessio, J. Sauer, *Phys. Chem. Chem. Phys.* **2018**, *20*, 19964–19970.

[145] Y.-Q. Su, Y. Wang, J.-X. Liu, I. A. Filot, K. Alexopoulos, L. Zhang, V. Muravev, B. Zijlstra, D. G. Vlachos, E. J. Hensen, *ACS Catal.* **2019**, *9*, 3289–3297.

[146] C. Lo, B. Trout, *J. Catal.* **2004**, *227*, 77–89.

[147] M. E. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation*, Oxford University Press, Oxford, **2010**.

[148] G. Fogarasi, X. Zhou, P. W. Taylor, P. Pulay, *J. Am. Chem. Soc.* **1992**, *114*, 8191–8201.

[149] J. Baker, *J. Comput. Chem.* **1993**, *14*, 1085–1100.

[150] V. Bakken, T. Helgaker, *J. Chem. Phys.* **2002**, *117*, 9160–9174.

[151] T. Bučko, J. Hafner, J. G. Ángyán, *J. Chem. Phys.* **2005**, *122*, 124508.

[152] T. Bučko, *Theor. Chem. Acc.* **2018**, *137*, 164.

[153] E. B. Wilson, *J. Chem. Phys.* **1941**, *9*, 76–84.

[154] O. Farkas, H. B. Schlegel, *J. Chem. Phys.* **1998**, *109*, 7100–7104.

[155] C. Peng, P. Y. Ayala, H. B. Schlegel, M. J. Frisch, *J. Comput. Chem.* **1996**, *17*, 49–56.

[156] P. Pulay, G. Fogarasi, *J. Chem. Phys.* **1992**, *96*, 2856–2860.

[157] G. Ciccotti, M. Ferrario, *Mol. Simul.* **2004**, *30*, 787–793.

[158] T. Bučko, S. Chibani, J.-F. Paul, L. Cantrel, M. Badawi, *Phys. Chem. Chem. Phys.* **2017**, *19*, 27530–27543.

[159] J.-P. Ryckaert, G. Ciccotti, H. J. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327–341.

[160] M. Sprik, *Faraday Discuss.* **1998**, *110*, 437–445.

[161] P. Pulay, G. Fogarasi, F. Pang, J. E. Boggs, *J. Am. Chem. Soc.* **1979**, *101*, 2550–2560.

[162] J. Baker, A. Kessi, B. Delley, *J. Chem. Phys.* **1996**, *105*, 192–212.

[163] J. Andzelm, R. King-Smith, G. Fitzgerald, *Chem. Phys. Lett.* **2001**, *335*, 321–326.

[164] K. N. Kudin, G. E. Scuseria, H. B. Schlegel, *J. Chem. Phys.* **2001**, *114*, 2919–2923.

[165] C. Panosetti, K. Krautgasser, D. Palagin, K. Reuter, R. J. Maurer, *Nano Lett.* **2015**, *15*, 8044–8048.

[166] A. P. Bartók, R. Kondor, G. Csányi, *Phys. Rev. B* **2013**, *87*, 184115.

[167] M. Rupp, A. Tkatchenko, K.-R. Müller, O. A. von Lilienfeld, *Phys. Rev. Lett.* **2012**, *108*, 058301.

[168] K. M. Jablonka, D. Ongari, S. M. Moosavi, B. Smit, *Chem. Rev.* **2020**, *120*, 8066–8129.

[169] J. K. Nørskov, F. Studt, F. Abild-Pedersen, T. Bligaard, *Fundamental Concepts in Heterogeneous Catalysis*, Wiley, Hoboken, New Jersey, **2014**.

[170] T. Bučko, J. Hafner, *J. Phys.: Condens. Matter* **2010**, *22*, 384201.

[171] P. J. Dauenhauer, O. A. Abdelrahman, *ACS Cent. Sci.* **2018**, *4*, 1235–1243.

[172] R. Réocreux, C. Michel, P. Fleurat-Lessard, P. Sautet, S. N. Steinmann, *J. Phys. Chem. C* **2019**, *123*, 28828–28835.

[173] K. Alexopoulos, M.-S. Lee, Y. Liu, Y. Zhi, Y. Liu, M.-F. Reyniers, G. B. Marin, V.-A. Glezakou, R. Rousseau, J. A. Lercher, *J. Phys. Chem. C* **2016**, *120*, 7172–7182.

[174] K. S. Pitzer, W. D. Gwinn, *J. Chem. Phys.* **1942**, *10*, 428–440.

[175] R. B. McClurg, R. C. Flagan, W. A. Goddard III, *J. Chem. Phys.* **1997**, *106*, 6675–6680.

[176] J. Pfaendtner, X. Yu, L. J. Broadbelt, *Theor. Chem. Acc.* **2007**, *118*, 881–898.

[177] G. Piccini, J. Sauer, *J. Chem. Theory Comput.* **2013**, *9*, 5038–5045.

[178] G. Piccini, M. Alessio, J. Sauer, *Angew. Chem. Int. Ed.* **2016**, *55*, 5235–5237.

[179] V. Van Speybroeck, D. Van Neck, M. Waroquier, S. Wauters, M. Saeys, G. B. Marin, *J. Phys. Chem. A* **2000**, *104*, 10939–10950.

[180] V. Van Speybroeck, D. Van Neck, M. Waroquier, *J. Phys. Chem. A* **2002**, *106*, 8945–8950.

[181] P. Vansteenkiste, V. Van Speybroeck, G. B. Marin, M. Waroquier, *J. Phys. Chem. A* **2003**, *107*, 3139–3145.

[182] R. M. Barrer, J. A. Davies, *Proc. R. Soc. Lond. A. Mat.* **1970**, *320*, 289–308.

[183] H. J. F. Stroud, E. Richards, P. Limcharoen, N. G. Parsonage, *J. Chem. Soc. Faraday Trans. 1* **1976**, *72*, 942.

[184] J. Soto, A. Myers, *Mol. Phys.* **1981**, *42*, 971–983.

[185] G. B. Woods, J. S. Rowlinson, *J. Chem. Soc. Faraday Trans. 2* **1989**, *85*, 765.

[186] R. Q. Snurr, R. L. June, A. T. Bell, D. N. Theodorou, *Mol. Simul.* **1991**, *8*, 73–92.

[187] F. Karavias, A. L. Myers, *Langmuir* **1991**, *7*, 3118–3126.

[188] *Modelling of Structure and Reactivity in Zeolites*, (Ed.: C. R. A. Catlow), Academic Press, London, **1992**.

[189] P. R. Van Tassel, H. T. Davis, A. V. McCormick, *J. Chem. Phys.* **1993**, *98*, 8919–8928.

[190] M. W. Maddox, J. S. Rowlinson, *Faraday Trans.* **1993**, *89*, 3619.

[191] B. Smit, *J. Phys. Chem.* **1995**, *99*, 5597–5603.

[192] T. Bučko, J. Hafner, *J. Catal.* **2015**, *329*, 32–48.

[193] G. Kresse, J. Hafner, *Phys. Rev. B* **1993**, *48*, 13115–13118.

[194] G. Kresse, J. Hafner, *Phys. Rev. B* **1994**, *49*, 14251–14269.

[195] G. Kresse, J. Furthmüller, *Comput. Mater. Sci.* **1996**, *6*, 15–50.

[196] G. Kresse, J. Furthmüller, *Phys. Rev. B* **1996**, *54*, 11169–11186.

[197] P. E. Blöchl, *Phys. Rev. B* **1994**, *50*, 17953–17979.

[198] G. Kresse, D. Joubert, *Phys. Rev. B* **1999**, *59*, 1758–1775.

[199] T. Bučko, J. Hafner, S. Lebègue, J. G. Ángyán, *J. Phys. Chem. A* **2010**, *114*, 11814–11824.

[200] H. J. Monkhorst, J. D. Pack, *Phys. Rev. B* **1976**, *13*, 5188–5192.

[201] S. K. Schiferl, D. C. Wallace, *J. Chem. Phys.* **1985**, *83*, 5203–5209.

[202] N. Shklov, *Am. Math. Mon.* **1960**, *67*, 1022.

[203] *Numerical Recipes: The Art of Scientific Computing*, 3rd ed, (Ed.: W. H. Press), Cambridge University Press, Cambridge, UK, **2007**.

[204] R. Cameron, B., *Math. Enthus.* **2014**, *11*.

[205] K. K. Irikura, *J. Phys. Chem. Ref. Data* **2007**, *36*, 389–397.

[206] C. C. Marston, G. G. Balint-Kurti, *J. Chem. Phys.* **1989**, *91*, 3571–3576.

[207] T. L. Hill, *An Introduction to Statistical Thermodynamics*, Dover Publications, New York, **1986**.

[208] L. Smith, A. Davidson, A. Cheetham, *Catal. Lett.* **1997**, *49*, 143–146.

[209] A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, K. W. Jacobsen, *J. Phys.: Condens. Mat.* **2017**, *29*, 273002.

[210] P. H. Berens, D. H. J. Mackay, G. M. White, K. R. Wilson, *J. Chem. Phys.* **1983**, *79*, 2375–2389.

[211] W. G. Hoover, A. J. C. Ladd, B. Moran, *Phys. Rev. Lett.* **1982**, *48*, 1818–1820.

[212] D. J. Evans, *J. Chem. Phys.* **1983**, *78*, 3297–3302.

[213] S. Nosé, *J. Chem. Phys.* **1984**, *81*, 511–519.

[214] W. G. Hoover, *Phys. Rev. A* **1985**, *31*, 1695–1697.

[215] G. J. Martyna, M. L. Klein, M. Tuckerman, *J. Chem. Phys.* **1992**, *97*, 2635–2643.

[216] A. Bajpai, P. Mehta, K. Frey, A. M. Lehmer, W. F. Schneider, *ACS Catal.* **2018**, *8*, 1945–1954.

[217] G. Collinge, S. F. Yuk, M.-T. Nguyen, M.-S. Lee, V.-A. Glezakou, R. Rousseau, *ACS Catal.* **2020**, *10*, 9236–9260.

[218] F. Knoop, T. A. R. Purcell, M. Scheffler, C. Carbogno, *Phys. Rev. Mater.* **2020**, *4*, 083809.

[219] V. Van Speybroeck, K. De Wispelaere, J. Van der Mynsbrugge, M. Vandichel, K. Hemelsoet, M. Waroquier, *Chem. Soc. Rev.* **2014**, *43*, 7326–7357.

[220] W. Chen, L. Sun, B. Kozinsky, C. M. Friend, E. Kaxiras, P. Sautet, R. J. Madix, *J. Phys. Chem. C* **2020**, *124*, 1429–1437.

[221] G. Leibfried, W. Ludwig in *Solid State Physics*, *Vol. 12*, Elsevier, **1961**, pp. 275–444.

[222] O. Hellman, P. Steneteg, I. A. Abrikosov, S. I. Simak, *Phys. Rev. B* **2013**, *87*, 104111.

[223] R. Bianco, I. Errea, L. Paulatto, M. Calandra, F. Mauri, *Phys. Rev. B* **2017**, *96*, 014111.

[224] M. Zacharias, M. Scheffler, C. Carbogno, *Phys. Rev. B* **2020**, *102*, 045126.

[225] L. Monacelli, R. Bianco, M. Cherubini, M. Calandra, I. Errea, F. Mauri, *J. Phys.: Condens. Matter* **2021**, *33*, 363001.

[226] L. H. Sprowl, C. T. Campbell, L. Árnadóttir, *J. Phys. Chem. C* **2017**, *121*, 9655–9655.

[227] R. Baron, W. van Gunsteren, P. Hünenberger, *Trends Phys. Chem.* **2006**, *11*, 87–122.

[228] R. Baron, P. H. Hünenberger, J. A. McCammon, *J. Chem. Theory Comput.* **2009**, *5*, 3150–3160.

[229] O. Hellman, I. A. Abrikosov, *Phys. Rev. B* **2013**, *88*, 144301.

[230] C. Waitt, A. R. Miles, W. F. Schneider, *J. Phys. Chem. C* **2021**, acs.jpcc.1c05917.

[231] S. G. Moustafa, A. J. Schultz, E. Zurek, D. A. Kofke, *Phys. Rev. B* **2017**, *96*, 014117.

[232] R. Jinnouchi, J. Lahnsteiner, F. Karsai, G. Kresse, M. Bokdam, *Phys. Rev. Lett.* **2019**, *122*, 225701.

[233] C. H. Bennett, *J. Comp. Phys* **1976**, *22*, 245–268.

[234] G. König, S. Bruckner, S. Boresch, *J. Comput. Chem.* **2009**, *30*, 1712–1718.

[235] P. Domingos, *Data Min. Knowl. Discov.* **1999**, *3*, 409–425.

[236] R. Solomonoff, *Information and Control* **1964**, *7*, 1–22.

[237] J. Rissanen, *Automatica* **1978**, *14*, 465–471.

[238] P. D. Grünwald, *arXiv:math/0406077* **2004**.

[239] P. D. Grünwald, *The Minimum Description Length Principle*, MIT Press, Cambridge, Mass, **2007**.

[240] P. Domingos, *Commun. ACM* **2012**, *55*, 78–87.

[241] L. Himanen, M. O. Jäger, E. V. Morooka, F. Federici Canova, Y. S. Ranawat, D. Z. Gao, P. Rinke, A. S. Foster, *Comput. Phys. Commun.* **2020**, *247*, 106949.

[242] F. Faber, A. Lindmaa, O. A. von Lilienfeld, R. Armiento, *Int. J. Quantum Chem.* **2015**, *115*, 1094–1101.

[243] A. P. Bartók, M. C. Payne, R. Kondor, G. Csányi, *Phys. Rev. Lett.* **2010**, *104*, 136403.

[244] M. E. Potter, J. Amsler, L. Spiske, P. N. Plessow, T. Asare, M. Carravetta, R. Raja, P. A. Cox, F. Studt, L.-M. Armstrong, *ACS Catal.* **2023**, DOI 10.1021/acscatal.3c00352.

[245] J. Amsler, S. Bernart, P. N. Plessow, F. Studt, *Catal. Sci. Technol.* **2022**, *12*, 3311–3321.

[246] J. Amsler, P. N. Plessow, F. Studt, *Catal. Lett.* **2021**, *151*, 2595–2602.

[247] X. Zhang, R. Wang, X. Yang, F. Zhang, *Microporous Mesoporous Mater.* **2008**, *116*, 210–215.

[248] F. F. Madeira, N. Gnep, P. Magnoux, S. Maury, N. Cadran, *Appl. Catal. A: General* **2009**, *367*, 39–46.

[249] C. P. Nash, A. Ramanathan, D. A. Ruddy, M. Behl, E. Gjersing, M. Griffin, H. Zhu, B. Subramaniam, J. A. Schaidle, J. E. Hensley, *Appl. Catal. A: General* **2016**, *510*, 110–124.

[250] G. Garbarino, G. Pampararo, T. K. Phung, P. Riani, G. Busca, *Energies* **2020**, *13*, 3587.

[251] M. Zhang, Y. Yu, *Ind. Eng. Chem. Res.* **2013**, *52*, 9505–9514.

[252] K. Alexopoulos, M. John, K. Van der Borght, V. Galvita, M.-F. Reyniers, G. B. Marin, *J. Catal.* **2016**, *339*, 173–185.

[253] M. E. Potter, L.-M. Armstrong, R. Raja, *Catal. Sci. Technol.* **2018**, *8*, 6163–6172.

[254] H. Tong, *Petrochem. Technol.* **1985**, *14*, 92–93.

[255] R. Le Van Mao, T. M. Nguyen, G. P. McLaughlin, *Appl. Catal.* **1989**, *48*, 265–277.

[256] C. B. Phillips, R. Datta, *Ind. Eng. Chem. Res.* **1997**, *36*, 4466–4475.

[257] L. Pan, H. Li, *Petrochem. Technol.* **1985**, *14*, 154–157.

[258] L. Pan, H. Li, *Petrochem. Technol.* **1987**, *16*, 764–768.

[259] D. Wang, J. Li, M. Fan, *Chin. J. Catal.* **1992**, *13*, 234–236.

[260] T. Zhou, X.-x. Shi, L.-n. Wu, Q. Cui, H.-y. Wang, Y. Hu, H. Huang, *J. Chem. Eng. Chin. Univ.* **2011**, *3*, 453–458.

[261] H. Xin, X. Li, Y. Fang, X. Yi, W. Hu, Y. Chu, F. Zhang, A. Zheng, H. Zhang, X. Li, *J. Catal.* **2014**, *312*, 204–215.

[262] T. K. Phung, L. Proietti Hernández, A. Lagazzo, G. Busca, *Appl. Catal. A: General* **2015**, *493*, 77–89.

[263] D. Varisli, T. Dogu, G. Dogu, *Chem. Eng. Sci.* **2007**, *62*, 5349–5352.

[264] Z. Benliang, Z. Baozhong, *J. Northeast Normal Univ. (Nat. Sci. Ed.)* **1995**, *1*, 70–72.

[265] M. John, K. Alexopoulos, M.-F. Reyniers, G. B. Marin, *ACS Catal.* **2016**, *6*, 4081–4094.

[266] A. J. Hoffman, J. S. Bates, J. R. Di Iorio, S. V. Nystrom, C. T. Nimlos, R. Gounder, D. Hibbitts, *Angew. Chem.* **2020**, *132*, 18845–18853.

[267] M. Fečík, P. N. Plessow, F. Studt, *J. Phys. Chem. C* **2018**, *122*, 23062–23067.

[268] P. N. Plessow, F. Studt, *ACS Catal.* **2017**, *7*, 7987–7994.

[269] M. Fečík, P. N. Plessow, F. Studt, *ACS Catal.* **2020**, *10*, 8916–8925.

[270] M. Fečík, P. N. Plessow, F. Studt, *Catal. Sci. Technol.* **2021**, *11*, 3826–3833.

[271] C. Riplinger, F. Neese, *J. Chem. Phys.* **2013**, *138*, 034106.

[272] C. Riplinger, B. Sandhoefer, A. Hansen, F. Neese, *J. Chem. Phys.* **2013**, *139*, 134101.

[273] D. Feller, *J. Chem. Phys.* **1992**, *96*, 6104–6114.

[274] T. Helgaker, W. Klopper, H. Koch, J. Noga, *J. Chem. Phys.* **1997**, *106*, 9639–9646.

[275] F. Neese, *WIREs Comput. Mol. Sci.* **2012**, *2*, 73–78.

[276] *TURBOMOLE V.7.1 2016, a Development of Karlsruhe Institute of Technology, Karlsruhe, 1989-2019, TURBOMOLE GmbH, since 2007; Available from http://www.turbomole.com.* Turbomole GmbH, **2016**.

[277] F. Weigend, F. Furche, R. Ahlrichs, *J. Chem. Phys.* **2003**, *119*, 12753–12762.

[278] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.

[279] C. L. Janssen, I. M. Nielsen, *Chem. Phys. Lett.* **1998**, *290*, 423–430.

[280] W. Jiang, N. J. DeYonker, A. K. Wilson, *J. Chem. Theory Comput.* **2012**, *8*, 460–468.

[281] F. Weigend, M. Häser, H. Patzelt, R. Ahlrichs, *Chem. Phys. Lett.* **1998**, *294*, 143–152.

[282] T. H. Dunning, *J. Chem. Phys.* **1989**, *90*, 1007–1023.

[283] Y. Minenkov, E. Chermak, L. Cavallo, *J. Chem. Theory Comput.* **2015**, *11*, 4664–4676.

[284] Y. Minenkov, G. Bistoni, C. Riplinger, A. A. Auer, F. Neese, L. Cavallo, *Phys. Chem. Chem. Phys.* **2017**, *19*, 9374–9391.

[285] M. Saitow, U. Becker, C. Riplinger, E. F. Valeev, F. Neese, *J. Chem. Phys.* **2017**, *146*, 164105.

[286] F. Neese, F. Wennmohs, A. Hansen, U. Becker, *Chem. Phys.* **2009**, *356*, 98–109.

[287] D. Zhang, R. Wang, X. Yang, *Catal. Lett.* **2008**, *124*, 384–391.

[288] F. Bleken, M. Bjørgen, L. Palumbo, S. Bordiga, S. Svelle, K.-P. Lillerud, U. Olsbye, *Top. Catal.* **2009**, *52*, 218–228.

[289] W. Song, D. M. Marcus, H. Fu, J. O. Ehresmann, J. F. Haw, *J. Am. Chem. Soc.* **2002**, *124*, 3844–3845.

[290] D. Lesthaeghe, V. Van Speybroeck, G. B. Marin, M. Waroquier, *Angew. Chem. Int. Ed.* **2006**, *45*, 1714–1719.

[291] N. Tajima, T. Tsuneda, F. Toyama, K. Hirao, *J. Am. Chem. Soc.* **1998**, *120*, 8222–8229.

[292] A. D. Chowdhury, K. Houben, G. T. Whiting, M. Mokhtar, A. M. Asiri, S. A. Al-Thabaiti, S. N. Basahel, M. Baldus, B. M. Weckhuysen, *Angew. Chem. Int. Ed.* **2016**, *55*, 15840–15845.

[293] Y. Liu, S. Müller, D. Berger, J. Jelic, K. Reuter, M. Tonigold, M. Sanchez-Sanchez, J. A. Lercher, *Angew. Chem. Int. Ed.* **2016**, *55*, 5723–5726.

[294] X. Wu, S. Xu, W. Zhang, J. Huang, J. Li, B. Yu, Y. Wei, Z. Liu, *Angew. Chem. Int. Ed.* **2017**, *56*, 9039–9043.

[295] K. Van der Borght, R. Batchu, V. V. Galvita, K. Alexopoulos, M.-F. Reyniers, J. W. Thybaut, G. B. Marin, *Angew. Chem. Int. Ed.* **2016**, *55*, 12817–12821.

[296] R. H. Gil-Horán, J. C. Chavarría-Hernández, P. Quintana-Owen, A. Gutiérrez-Alejandre, *Top. Catal.* **2020**, *63*, 414–427.

[297] C. W. Ingram, R. J. Lancashire, *Catal. Lett.* **1995**, *31*, 395–403.

[298] K. De Wispelaere, K. Hemelsoet, M. Waroquier, V. Van Speybroeck, *J. Catal.* **2013**, *305*, 76–80.

[299] K. Hemelsoet, J. Van der Mynsbrugge, K. De Wispelaere, M. Waroquier, V. Van Speybroeck, *ChemPhysChem* **2013**, *14*, 1526–1545.

[300] S. L. C. Moors, K. De Wispelaere, J. Van der Mynsbrugge, M. Waroquier, V. Van Speybroeck, *ACS Catal.* **2013**, *3*, 2556–2567.

[301] C.-M. Wang, Y.-D. Wang, Z.-K. Xie, *J. Catal.* **2013**, *301*, 8–19.

[302] M. DeLuca, P. Kravchenko, A. Hoffman, D. Hibbitts, *ACS Catal.* **2019**, *9*, 6444–6460.

[303] M. DeLuca, C. Janes, D. Hibbitts, *ACS Catal.* **2020**, *10*, 4593–4607.

[304] A. D. Chowdhury, A. L. Paioni, K. Houben, G. T. Whiting, M. Baldus, B. M. Weckhuysen, *Angew. Chem. Int. Ed.* **2018**, *57*, 8095–8099.

[305] A. A. Arvidsson, P. N. Plessow, F. Studt, A. Hellman, *J. Phys. Chem. C* **2020**, *124*, 14658–14663.

[306] S. Namuangruk, P. Pantu, J. Limtrakul, *ChemPhysChem* **2005**, *6*, 1333–1339.

[307] J. Lu, Y. Liu, *J. Nat. Gas Chem.* **2011**, *20*, 162–166.

[308] M. Fečík, P. N. Plessow, F. Studt, *ChemCatChem* **2021**, *13*, 2451–2458.

[309] V. Van Speybroeck, K. Hemelsoet, L. Joos, M. Waroquier, R. G. Bell, C. R. A. Catlow, *Chem. Soc. Rev.* **2015**, *44*, 7044–7111.

[310] G. Li, E. A. Pidko, *ChemCatChem* **2019**, *11*, 134–156.

[311] C. Chizallet, *ACS Catal.* **2020**, *10*, 5579–5601.

[312] J. Rey, C. Bignaud, P. Raybaud, T. Bučko, C. Chizallet, *Angew. Chem.* **2020**, *132*, 19100–19104.

[313] T. V. Janssens, S. Svelle, U. Olsbye, *J. Catal.* **2013**, *308*, 122–130.

[314] P. Kumar, J. Thybaut, S. Teketel, S. Svelle, P. Beato, U. Olsbye, G. Marin, *Catal. Today* **2013**, *215*, 224–232.

[315] P. Kumar, J. W. Thybaut, S. Svelle, U. Olsbye, G. B. Marin, *Ind. Eng. Chem. Res.* **2013**, *52*, 1491–1507.

[316] S. Standl, F. M. Kirchberger, T. Kühlewind, M. Tonigold, M. Sanchez-Sanchez, J. A. Lercher, O. Hinrichsen, *Chem. Eng. J.* **2020**, *402*, 126023.

[317] G. J. Hutchings, F. Gottschalk, R. Hunter, *Ind. Eng. Chem. Res.* **1987**, *26*, 635–637.

[318] G. J. Hutchings, F. Gottschalk, M. V. M. Hall, R. Hunter, *J. Chem. Soc. Faraday Trans. 1* **1987**, *83*, 571.

[319] F. Salehirad, M. W. Anderson, *J. Catal.* **1996**, *164*, 301–314.

[320] S. R. Blaszkowski, R. A. van Santen, *J. Am. Chem. Soc.* **1996**, *118*, 5152–5153.

[321] P. E. Sinclair, C. R. A. Catlow, *J. Phys. Chem. B* **1997**, *101*, 295–298.

[322] S. R. Blaszkowski, R. A. van Santen, *J. Am. Chem. Soc.* **1997**, *119*, 5020–5027.

[323] D. Lesthaeghe, V. Van Speybroeck, G. Marin, M. Waroquier, *Chem. Phys. Lett.* **2006**, *417*, 309–315.

[324] J. Li, Z. Wei, Y. Chen, B. Jing, Y. He, M. Dong, H. Jiao, X. Li, Z. Qin, J. Wang, W. Fan, *J. Catal.* **2014**, *317*, 277–283.

[325] A. Comas-Vives, M. Valla, C. Copéret, P. Sautet, *ACS Cent. Sci.* **2015**, *1*, 313–319.

[326] Z. Wei, Y.-Y. Chen, J. Li, W. Guo, S. Wang, M. Dong, Z. Qin, J. Wang, H. Jiao, W. Fan, *J. Phys. Chem. C* **2016**, *120*, 6075–6087.

[327] Y. Liu, S. Müller, D. Berger, J. Jelic, K. Reuter, M. Tonigold, M. Sanchez-Sanchez, J. A. Lercher, *Angew. Chem.* **2017**, *129*, 7448–7448.

[328] D. B. Rasmussen, J. M. Christensen, B. Temel, F. Studt, P. G. Moses, J. Rossmeisl, A. Riisager, A. D. Jensen, *Catal. Sci. Technol.* **2017**, *7*, 1141–1152.

[329] O. Dewaele, V. L. Geers, G. F. Froment, G. B. Marin, *Chem. Eng. Sci.* **1999**, *54*, 4385–4395.

[330] M. Stöcker, *Microporous Mesoporous Mater.* **1999**, *29*, 3–48.

[331] W. Wang, A. Buchholz, M. Seiler, M. Hunger, *J. Am. Chem. Soc.* **2003**, *125*, 15260–15267.

[332] Y. Jiang, W. Wang, V. Reddymarthala, J. Huang, B. Sulikowski, M. Hunger, *J. Catal.* **2006**, *238*, 21–27.

[333] W. Wang, M. Hunger, *Acc. Chem. Res.* **2008**, *41*, 895–904.

[334] X. Chen, M. L. Neidig, R. Tuinstra, A. Malek, *J. Phys. Chem. Lett.* **2010**, *1*, 3012–3015.

[335] H. Yamazaki, H. Shima, H. Imai, T. Yokoi, T. Tatsumi, J. N. Kondo, *Angew. Chem. Int. Ed.* **2011**, *50*, 1853–1856.

[336] A. Hwang, D. Prieto-Centurion, A. Bhan, *J. Catal.* **2016**, *337*, 52–56.

[337] Z. Wei, Y.-Y. Chen, J. Li, P. Wang, B. Jing, Y. He, M. Dong, H. Jiao, Z. Qin, J. Wang, W. Fan, *Catal. Sci. Technol.* **2016**, *6*, 5526–5533.

[338] A. Schäfer, H. Horn, R. Ahlrichs, *J. Chem. Phys.* **1992**, *97*, 2571–2577.

[339] A. Schäfer, C. Huber, R. Ahlrichs, *J. Chem. Phys.* **1994**, *100*, 5829–5835.

[340] F. Weigend, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057.

[341] C. Chizallet, P. Raybaud, *ChemPhysChem* **2010**, *11*, 105–108.

[342] F. Leydier, C. Chizallet, A. Chaumonnot, M. Digne, E. Soyer, A.-A. Quoineaud, D. Costa, P. Raybaud, *J. Catal.* **2011**, *284*, 215–229.

[343] S. Schallmoser, T. Ikuno, M. Wagenhofer, R. Kolvenbach, G. Haller, M. Sanchez-Sanchez, J. Lercher, *J. Catal.* **2014**, *316*, 93–102.

[344] Y. Chu, X. Yi, C. Li, X. Sun, A. Zheng, *Chem. Sci.* **2018**, *9*, 6470–6479.

[345] R. van den Berg, G. Prieto, G. Korpershoek, L. I. van der Wal, A. J. van Bunningen, S. Lægsgaard-Jørgensen, P. E. de Jongh, K. P. de Jong, *Nat. Commun.* **2016**, *7*, 13057.

[346] Commercial ACS-grade Methanol by Merck KGaA, https://www.merckmillipore.com, **2020**.

[347] J. W. Moore, R. G. Pearson, A. A. Frost, *Kinetics and Mechanism*, 3rd ed, Wiley, New York, **1981**.

[348] S. Zones, R. Van Nordstrand, *Zeolites* **1988**, *8*, 166–174.

A. Appendix

Regression of Experimental Adsorption Free Energies

Adsorption free energies of Ar and N_2 in acidic chabazite zeolite were measured by Barrer and Davies[182] for a range of temperatures (138 K to 195 K for Ar and 78 K to 210 K for N_2). Their experimental data plotted in Figure A.1 exhibit a pronounced linear trend, which was used to determine $\Delta_{ads}G$ at 200 K.

Structural Model of Zeolite H-SSZ-13 used in Chapter 3

The studies presented in Chapter 3 reported simulations with the primitive unit cell of chabazite zeolite H-SSZ-13 with a Si to Al ratio of 11:1 (see Figure A.2). The acid proton was located in position O(1) according to the nomenclature of Jeanvoine et al.,[50] which belongs to one of the most populated proton sittings on chabazite.[208] The lattice parameters have been derived from the experimental geometry determined for a highly siliceous form (H-SSZ-13) of chabazite[348] (space group $R\bar{3}m$) with lattice constants of $a = 9.338$ Å and $\alpha = 95.0°$. The same unit cell dimensions were used in calculations of electronic, vibrational and rotational contributions to free energies of the isolated adsorbate particles Ar and N_2.

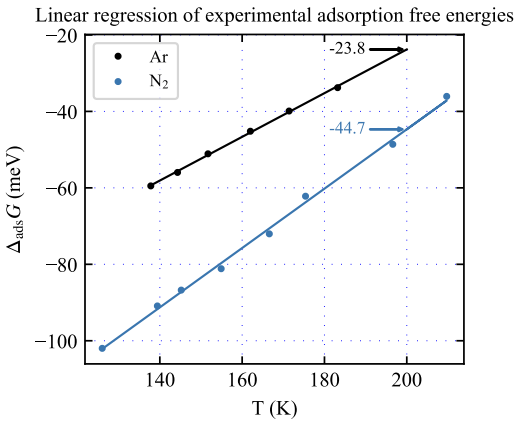


Figure A.1.: Linear regression of experimental adsorption free energies measured by Barrer and Davies[182] for Ar and N_2 in acidic chabazite zeolite. Adsorption free energies at 200 K are annotated. Reprinted with permission from [17]. Copyright 2021 American Chemical Society.

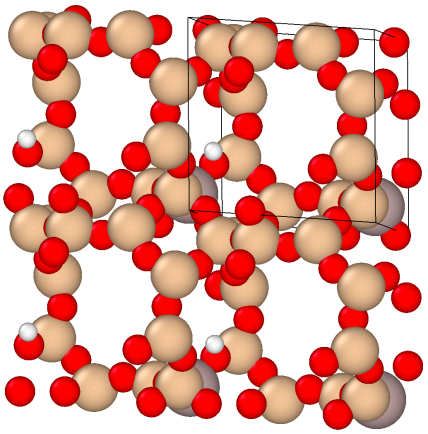


Figure A.2.: The frame in the upper right corner marks the boundaries of the primitive unit cell of chabazite zeolite H-SSZ-13. Color code: H (white), O (red), Si (beige), Al (gray). Reprinted with permission from [17]. Copyright 2021 American Chemical Society.

Different Convergence for Constrained and Free Reactant States

Using the rotamerization of ethane as an example, two λ-TI strategies for the calculation of anharmonic free energy barriers were compared in Section 3.3.2.1. The small deviation of the λ-TI$_b$ strategy from the λ-TI$_a$ and ξ-TI strategies was explained by a comparatively worse conversion over the same amount of simulation steps. Figure A.3 shows the convergence of the energy differences between the harmonic and anharmonic systems for the constrained reactant and the free reactant state. The much faster convergence of the constrained reactant state – also for higher values of λ – makes the λ-TI$_a$ strategy in Figure A.3a much more accurate and thus attractive for computational studies.

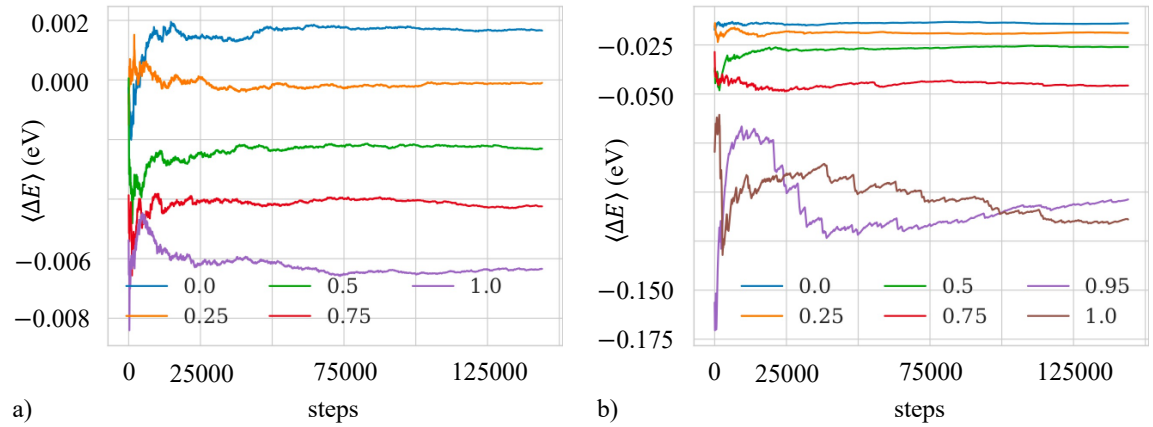


Figure A.3.: Convergence of the mean energy difference $\langle \Delta E \rangle = V_1 - V_{0,\boldsymbol{q}}$ for ethane computed for (a) the constrained reactant contributing to the λ-TI$_a$ strategy and (b) the free reactant state contributing to the λ-TI$_b$ strategy. The legend indicates different values of λ. Convergence improves when an increasing number of simulated steps is considered.

Ethylation of Butene

The free energy landscape for the ethylation of butene to hexene isomers becomes more clear when all states are referenced to EtOH in the gas phase. This has been done in Figure A.4, where all ethylation barriers from Figure 4.6 can be directly compared at 673.15 K and at a reference pressure of 1 bar.

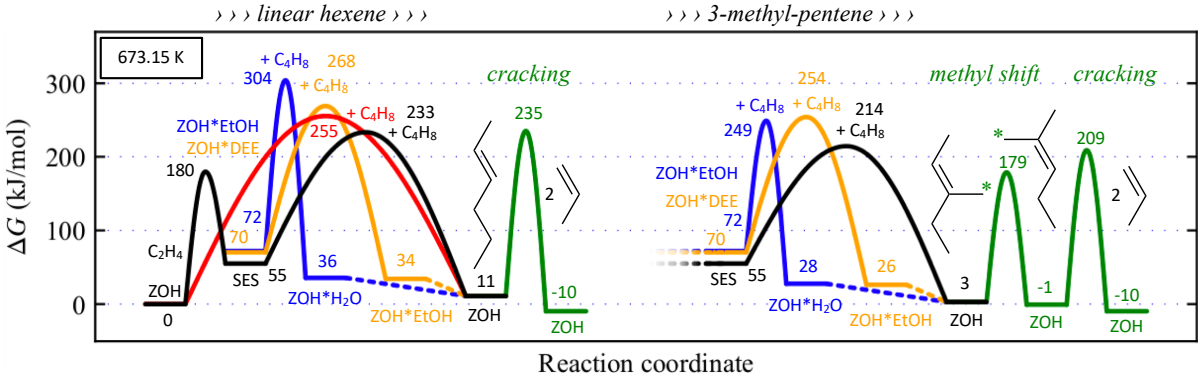


Figure A.4.: Gibbs free energy diagram for the ethylations of butene at 673.15 K and at a reference pressure of 1 bar. Direct ethylations can proceed via EtOH (blue), DEE (orange) or ethene dimerization (red). The stepwise ethylation (black) from ethene is mediated by the SES intermediate. The formation and cracking (green) of linear (left) and branched (right) hexene isomers can be compared. All states are referenced to ethene in the gas phase. Reprinted with permission from [245]. Copyright 2022 The Royal Society of Chemistry (CC-BY).

B. Acknowledgments

Several people have played great roles in shaping the author and thus this work. At this point I would like to thank everyone who accompanied and supported me all those years. Thank you very much![a]

In particular, Prof. Dr. Felix Studt, you fulfilled your role excellently in finding the right balance between experienced guidance and unperturbed freedom of research that suited me perfectly. Dr. Philipp N. Plessow, thank you for stimulating discussions, for your keen eyes and for radiating such a scientific aura. Dr. Tomáš Bučko, your vivid explanations are the best, what a great and fruitful collaboration! Thank you, Kevin M. Jablonka, Mila Krämer and Shenjun Zha, for your machine learning excursions.

A big thank you to all my colleagues! I very much enjoyed the pleasant and productive atmosphere. Thank you, Elisabeth, Michal and Tiago, for guiding my first steps at Campus North. Thank you, Kevin and Lukas, for offering a welcoming spare office with the most comfortable couch at Campus South.

From the research training group (RTG 2450), I want to highlight Julie, Mila, Rodrigo and Tarek for our great teamwork enabling me to have a pleasant term as spokesperson for the PhD student committee. Prof. Dr. Karin Fink and Prof. Dr. Volker Grimm, your regular reviews and helpful advice were very much appreciated. For our great collaboration in organizing the RTG symposium I thank Dr. Mariana Kozlowska, Prof. Dr. Marcus Elstner and Prof. Dr. Alexander Schug.

I very much appreciated helpful support by Sabine Holthoff and by the IKFT secretaries. Very special thanks go to Mr. Bergemann for his superior knowledge in office supply operations.

This line is dedicated to the legendary Mr. Riedel for his fantastic inspirations in math and physics. I also thank Dr. Villani for captivating science lessons and for arousing my fascination for chemistry.

For a unique life episode, Jule, Kevin, Kristina, Lea and Lukas, you are definitely worth a noble prize.
(not a typo)

Thank you, my friends and family! I am grateful for your reliable and compassionate support, for being there, for regular haystack lifting during summer, for collective gardening and bike tours.

[a] Besides, as usual, I gratefully acknowledge support by the GRK 2450, by the state of Baden-Württemberg through bwHPC (bwUniCluster and JUSTUS, RV bw17D01) and by the Helmholtz Association.

C. CV of the Author with List of Publications

Jonas Amsler ✉ jonas.amsler@kit.edu

…CV not available in digital format …

List of Publications

1. S. Tischer, M. Börnhorst, J. Amsler, G. Schoch, O. Deutschmann, *Phys. Chem. Chem. Phys.* **2019**, *21*, 16785–16797.
2. B. B. Sarma, J. Kim, J. Amsler, G. Agostini, C. Weidenthaler, N. Pfänder, R. Arenal, P. Concepción, P. Plessow, F. Studt, G. Prieto, *Angew. Chem. Int. Ed.* **2020**, *59*, 5806–5815.
3. J. Amsler, B. B. Sarma, G. Agostini, G. Prieto, P. N. Plessow, F. Studt, *J. Am. Chem. Soc.* **2020**, *142*, 5087–5096.
4. J. Amsler, P. N. Plessow, F. Studt, T. Bučko, *J. Chem. Theory Comput.* **2021**, *17*, 1155–1169.
5. J. Amsler, P. N. Plessow, F. Studt, *Catal. Lett.* **2021**, *151*, 2595–2602.
6. J. Amsler, S. Bernart, P. N. Plessow, F. Studt, *Catal. Sci. Technol.* **2022**, *12*, 3311–3321.
7. J. Amsler, P. N. Plessow, F. Studt, T. Bučko, *J. Chem. Theory Comput.* **2023**, DOI `10.1021/acs.jctc.3c00169`.
8. M. E. Potter, J. Amsler, L. Spiske, P. N. Plessow, T. Asare, M. Carravetta, R. Raja, P. A. Cox, F. Studt, L.-M. Armstrong, *ACS Catal.* **2023**, DOI `10.1021/acscatal.3c00352`.