# A computational investigation into the confinement induced thermo-kinetic effects of crystalline porous materials

# A computational investigation into the confinement induced thermo-kinetic effects of crystalline porous materials

by

Christopher Rzepa

Presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Doctor of Philosophy

in

Chemical Engineering

Lehigh University

Aug. 2022

Certificate of Approval

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

Date

_____

Dr. Srinivas Rangarajan, Ph.D., Committee Chair

_____

Dr. Jeetain Mittal, Ph.D., Committee Chair

_____

Dr. Daniel W Siderius, Ph.D., External Committee member

_____

Dr. Mark A. Snyder, Ph.D., Committee member

_____

Dr. Ganesh Balasubramanian, Ph.D., External Committee member

# ACKNOWLEDGMENTS

My time at Lehigh University can be best described as challenging, enlightening, and most importantly, delightful. Such an experience would not have been possible without people whom I would consider to be among my closest friends.

First, I would like to thank my PhD advisors Profs. Srinivas Rangarajan and Jeetain Mittal, who at the time took their chances by prematurely admitting a mediocre college graduate, but under whose guidance was cultivated into an academic researcher. I am forever grateful for their sharing of immense knowledge, providing honest, constructive feedback, and supporting me in the face of self doubt. Additionally, I would like to thank my committee members and collaborators: Profs. Mark Snyder, Ganesh Balasubramanian, Daniel W. Siderius, Vincent K. Shen, and Harold (Wickes) Hatch for offering their expertise throughout my research and making me infatuated with statistical thermodynamics.

Beneath the neatly organized data of a scientific publication lay the countless hours spent reading literature and performing experiments, to which are made tolerable by interacting with those around you. In this regard I would like to especially thank my lab partners: Dr. Huijie Tian, Dr. Bowen Li, Adhika Setiawan, Siddharth Prabhu, Chin-Fei Chang, and Aojie Li, Prof. Gregory Dignon, Roshan Regy, Nina Jovic, and Evan Pretti. Whose comradery and intellectually stimulating conversations have made my time within the Rangarajan and Mittal labs of Iacocca Hall and HST much more enjoyable. Additionally, I would like to thank my Lehigh colleagues: Dr. Daniyal Kiani, Dr. Manoj Silva, Dr. John McGlynn, Dr. John Sakizadeh and many others who have made my experience at Lehigh one of the most memorable. I would like to thank my friends from adolescence: Nicholas Marukos, Michael Bak,

Jerry Wang, Owen Walker, and Christopher Magenta for their support. In particular I would like to recognize Dr. Daniyal Kiani, for his enthusiasm during our failed fishing trips, but nevertheless sharing great laughs. Adhika Setiawan, in sharing among the most insightful conversations I ever had the pleasure of having. Nicholas Marukos, in demonstrating the work ethic and tenacity necessary in surviving undergraduate chemical engineering. Finally, Christopher Magenta, for being my lifelong friend.

The following people hold a special place in my heart, and without their love and support I would never have been able to earn my PhD. First, I would like to thank my fiancée Stephanie Creevey, who stood alongside me throughout the countless hours in the laboratory. My mother and father Maria and Adam Rzepa, who have devoted their lives for my success. Finally, my grandmother Wanda Śliwińska, for raising me into the person I am today. Kocham was wszystkich.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# ABSTRACT

The reactivity and selectivity of a chemical reaction may be significantly enhanced by the confining effects of hierarchically structured, heterogeneous, porous frameworks. While the composition of such frameworks span organic, metal-organic, and inorganic building units, their underlying feature is the presence of unique microporous topologies, which have a propensity for stabilizing adsorbates mainly through long-range dispersive interactions. Upon adsorption, the molecule is confined to the space of the framework's pore, where it sacrifices its degree of freedom (entropy) in favor of stabilizing interactions (enthalpy) with the framework structure. Under isothermal-isobaric conditions, the balance between entropic loss and enthalpic gain is dictated by the extent of confinement, which regulates the spontaneity, or thermodynamic favourability of the process. Excessive entropic loss may promulgate frameworks to act as molecular sieves, where prohibitively large species that cannot fit within their voids are excluded from adsorption and/or the reaction space. Likewise, insufficient confinement may result in the adsorbate merely equilibriating with the bulk gas phase, lacking the stability necessary to catalyze a particular process. Optimal confinement occurs when the geometry of the adsorbate matches the framework topology, acquiring enough stability to promote their respective reactions without limiting the product's rate of desorption, the reactants rate of adsorption, or limiting the formation of reactive intermediates. The continued discovery of new frameworks, and the hypothetical existence of hundreds of millions of others, make it conceivable and even likely that the optimal framework for many catalytic processes has yet to be discovered. Confinement driven heterogeneous catalysis is therefore a multifaceted problem; and despite our detailed understanding on an individual basis, our mechanistic understanding and ability to make empirical predictions remains limited.

In this thesis, we elucidate the phenomena of confinement by: 1) presenting its geometric relationship with the adsorption entropy, and the ability to make empirical predictions through interpretable linear correlations, 2) measuring confinement from a kinetic perspective on a zeolite promoted reaction system, and evaluating factors for framework optimization, 3) identifying readily quantifiable, geometric features of the adsorbate/framework system for rapid prediction of the adsorption entropy within zeolites.

# CHAPTER 1

# INTRODUCTION

Heterogeneous porous frameworks promote the reactivity and selectivity of reactions through their internal porous structures. Zeolites are a class of crystalline microporous frameworks which consist of channels with diameters on the order of 1-10Å. On the atomic level, their building units are composed of silicon atoms which are tetrahedrally coordinated to four oxygen atoms. In their purely siliceous forms, zeolites are charge neutral and have a propensity for gas adsorption with the ability to catalyze reactions through the confining nature of their channels.[1] Brønsted acid sites may be incorporated into the structure to further enhance reactivity; and they are formed by substituting the central silicon atom with an aluminum atom, which imparts a partial negative charge on the surrounding oxygen atoms that is compensated by a Bønsted proton. A Bønsted acid site is illustrated in Figure 1.1, where the transition state for a Diels-Alder reaction is stabilize by interacting with 1) the proton of the Brønsted acid, and 2) the framework channel of a zeolite named H-ZSM-5. Naturally, molecules with geometric diameters larger than the channel dimensions cannot adsorb or be generated within the framework. However, this effect may be much more subtle, such that the reaction becomes selective toward species that better "fit" within the channels.-a phenomena termed "Confinement".[2]

Despite there existing only a handful of unique zeolite frameworks (only 232 in 2017.[3]), determining the optimal zeolite for catalyzing a particular reaction continues to be a major challenge in the field. The extent of zeolite catalysis depends upon the degree of stability among rate enhancing intermediates and/or transition states. Although the reaction space of a system is multi-dimensional, it may be illustrated by a free energy diagram, where the peaks and troughs are dictated by the system's thermodynamic free energies as a function of the reaction coordinate. Figure

1.1 shows the free energy diagram for the Diels-Alder cycloaddition reaction between ethene and isoprene, with corresponding thermodynamic free energies plotted on the y-axis against each step of the reaction. Five products are possible between the reactants, which are color-coded according to their meta-/para- regiochemistries. The bold lines represent the unpromoted gas-phase reactions, which show two equilibrium states: the reactants and products, and a single non-equilibrium state: the transition state. In order to generate the product, the reactants must traverse through this transition state, which is also the largest thermodynamic free energy point along the path. The difference in free energy between the transition state and the reactants dictates the rate of product formation, which can be optimized by modifying the temperature (the probability that the reactants are in a high energy state) and pressure (probability that the reactants will collide). However, this transition state barrier is fundamentally determined by nature, and therefore the rates of product formation across all temperatures and pressures of the system are predefined.

Confinement imparted by the zeolite directly changes this free energy diagram, such that frameworks which better accommodate the shape of the transition state have smaller barriers. This is illustrated by the dashed lines in Figure 1.1, which show the zeolite promoted transition state free energies for each reaction are smaller relative to their unpromoted gas-phase reactions. Therefore, by manipulating the free energy surface, a catalyst can further accelerate the rate of a reaction. The promoted paths in Figure 1.1 also include two additional equilibrium states: the adsorbed reactants and products, which are typically assumed to be barrierless relative to the transition state. However, despite the smaller transition state barrier, there is the possibility of the zeolite over-stabilizing the reactants or products, such that the intrinsic barrier between the transition state and reactants is too large, and the reactants withold from reacting, or the products fail to desorb. Confinement is therefore multi-faceted;

where the optimal environment favors stabilizing rate controlling intermediates and transition states, while allowing reactants and products to adsorbed/desorb.

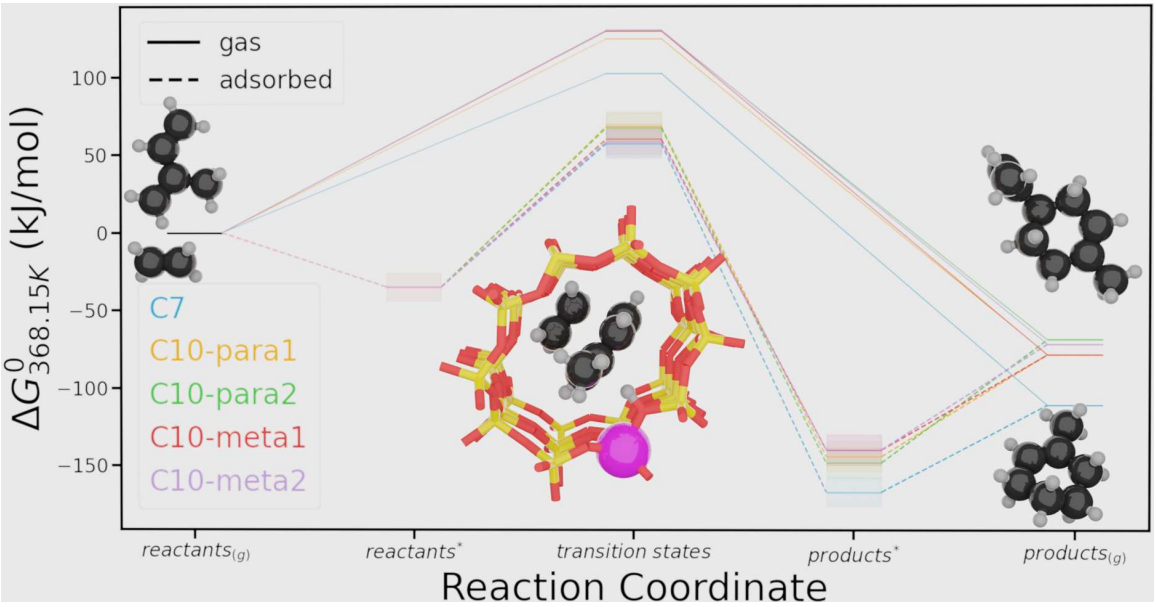


Figure 1.1: Free energy diagram of the Diels-Alder reaction between ethene and isoprene. the bold lines represent uncatalyzed gas-phase reactions, dashed lines represent zeolite catalyzed reactions. The reaction products, and corresponding coordinates are color-coded by their regio-chemistry.

An adsorbate's free energy can be deconstructed into contributions from energetic and entropic effects, where the former is a measure of the favorable molecule-framework dispersive interactions with the framework, and the latter is a measure of the unfavorable loss in the degrees of freedom of the molecule upon adsorption. Chemical reactions are typically ran under conditions of constant pressure and temperature, where the free energy of a closed system is defined by the Gibbs free energy:

$$\Delta G^0 = \Delta H^0 - T\Delta S^0 \tag{1.1}$$

where $\Delta G^0$, $\Delta H^0$ and $\Delta S^0$ are the Gibbs free energy, the enthalpy, and the entropy, respectively; where the superscript 0 indicates that the change is measured at some standard state (usually 298 K and 1 bar). Reactions can therefore be opti-

mized through the manipulation of these thermodynamic quantities by selecting the appropriate zeolite framework. Apart from the few documented zeolite frameworks, millions of hypothetical structures have been proposed based purely on geometric and energetic considerations[4]. It is therefore likely that the best frameworks for many reactions have yet to be discovered.

Within the uncatalyzed, ideal gas-phase, the entropy of a molecule can be rigorously calculated through fundamental statistical mechanical expressions. Ultimately, a molecule's entropy can be dissected into its respective vibrational, rotational, and translational degrees of freedom. Upon adsorption the molecule experiences a fractional loss in these modes commensurate to the size and shape of the framework structure, with the largest contributions to the entropy (translations and rotations) being impaired first; while the smallest contributions (vibrations) being largely preserved. It is therefore conceivable that geometric descriptors of the adsorbate/framework system should be capable in characterizing the adsorption entropy. In Chapter 2 we discuss the recent discovery of linear correlations between the entropy of the molecule in the adsorbed state and its standard gas-phase entropy within two dimensional catalytic surfaces[5, 6], and acidic/siliceous zeolites[7, 8]. We demonstrate how these entropic trends within zeolite frameworks can be computationally reproduced through Monte-Carlo simulations; and show that they are governed by physical, geometric, and readily quantifiable characteristics of the adsorbate/framework system. The work in this Chapter has been published within "The Journal of Physical Chemistry C" by Rzepa et al.[8]

The relationship between the adsorption entropy and the geometric aspects of the adsorbate/framework system suggest that machine-learned models, based on conveniently obtained physical descriptors, can be used to predict the thermodynamic quantities of rate determining intermediates and transition states. Consequently, the

extent of catalysis for a particular reaction may be readily predicted across many zeolite frameworks and thereby optimized. In Chapter 3 we determine the extent of confinement for a zeolite promoted reaction, and subsequently perform a kinetic analysis to determine the rate inhibiting intermediates and transition states. Ultimately, our analysis serves to establish a reference point for the kinetics and thermodynamics of the reaction, from which one may apply future models to optimize. In this chatper, we investigate the [4 + 2] Diels Alder (DA) reaction, henceforth called "DA reaction", which couples conjugated olefins to form six-membered cyclic products with various regio-chemistries. This particular reaction was chosen because it makes for an excellent archetype to study confinement-driven catalysis. Namely, this class of reactions has been well studied and extensively used within the pharmaceutical, agro-chemical, and oleochemical industries. DA reactions can occur uncatalyzed, being thermally driven to completion, but homogenous Lewis/Brønsted acid salts have been shown to promote them.[9, 10, 11, 12] Instinctively, acidic zeolites make for an excellent candidate to study their promotion; and recent experimental and computational studies have demonstrated this, but have been largely limited to furans.[13, 14, 15, 16, 17, 18, 19] The preservation of regio-chemistry throughout the reaction allows for precise control over the bulkiness of the reaction intermediates and transition states. In other words, a judicious choice of reactants can tailor transition states and products that "better fit" within the zeolite channels. The collection of computationally identified intermediates, transition states, and their thermodynamic values may then be fed into a reactor model, where macroscopic values such as the selectivity and rate controlling intermediates/steps can be determined. The work in this Chapter is currently under peer review within the "Catalysis Science and Technology Journal" published by the "Royal Society of Chemistry". In Chapter 5 section 5.1 we discuss a future outlook on how machine learned models may be used to optimize

this reaction.

Linear correlations are inevitably limited in their predictive capabilities; and given the vast realm of framework structures and candidate molecules, it becomes apparent that more predictive models are necessary in capturing complex structure-entropy relationships. Novel machine learning approaches based on topological features and atomistic graphs appear promising. In particular, machine learning models have shown to be adequate in predicting candidate structures for applications such as adsorption capacity[20, 21] and gas separations[22]. In Chapter 4 we show the limitations of the linear correlations from Chapter 2 through an expanded data set of almost four thousand zeolite-adsorbate systems. We transcend the predictive capabilities of linear correlations by building neural network models using readily-quantifiable geometric descriptors of the adsorbate/framework systems. Finally, we perform a local sensitivity analysis across each system in our data set to determine the most significant descriptors in predicting the adsorption entropy and infer their relationship to the principle of confinement.

# CHAPTER 2

# COMPUTATIONAL INVESTIGATION OF CORRELATIONS IN ADSORBATE ENTROPY FOR PURE-SILICA ZEOLITE ADSORBENTS

## 2.1 Motivation

Zeolites are crystalline, microporous frameworks that consist of unique channels and cages and have a propensity for gas adsorption. They have been historically used in light-gas separations and cracking of petrochemical feedstocks [23, 24]. Moreover, millions of hypothetical structures of zeolites have been proposed purely based on geometric and energetic considerations [4]. As a result, much work has been devoted to optimizing additional applications, such as methane storage [25, 26], carbon capture [27] and flue-gas treatment [28]. However, due to the large number of frameworks, identifying a zeolite whose geometry is well-suited to adsorb a particular molecule is difficult. This is further exacerbated by diffusion limitations [29], the realities of framework synthesis [30], thermal and mechanical stability, and postsynthesis treatment [31, 32]. Amidst these practical challenges, a reasonable first criterion for finding a suitable candidate is based on the thermodynamic spontaneity of the adsorption process, which is provided by the free energy of adsorption or, equivalently, the contributions from energetic and entropic effects.

Within siliceous zeolites and other neutral frameworks, the adsorption enthalpy is related to the strength of the van der Waals interactions between the molecule and pore [33]. On the other hand, the adsorption entropy can be interpreted as the measure of confinement associated with adsorption [7, 34, 35], where a molecule loses mobility and its degrees of freedom become hindered. The largest contributors to the entropy, translation and rotation, are the first to be impaired, while the

smallest contributor, vibration, is commonly assumed to be preserved. Techniques, such as gravimetry [36], chromatographic adsorption [34], and infrared operando spectroscopy [37], provide experimental routes in measuring the adsorption entropy. However, performing experiments for many molecule-framework combinations is impractical. Computational experiments of adsorption therefore make for an attractive alternative. Nevertheless, there remain caveats. For example, quantifying the entropy using ab initio-based methods (such as density functional theory) is computationally demanding [38, 39, 40, 41, 42] or relies on approximations [39]. Such approximations include (1) limiting translation to the surface area of the zeolite [43], which artificially restricts the entropy to a surface term; (2) invocation of the "harmonic approximation" based on the assumption of a very strong fluid–solid interaction that renders translations and rotations unimportant; or (3) employing the hindered translator/rotor model [44]. On the other hand, (classical) Monte Carlo (MC) and molecular dynamics simulations have proven successful in accurately calculating the adsorption entropy [45], with the primary limitation being the number of force fields available to accurately model the adsorbate–adsorbent interaction.

Although these simulations are numerically cleaner compared to experiment because they involve analyzing efficiently a greater number of molecule–zeolite combinations, it remains overwhelming to apply them to the millions of hypothetical frameworks. Novel topology-based, data-driven approaches have been shown to be adequate in predicting specific features such as adsorption capacity [27] and selectivity [22]. However, there are currently no similar models for entropy.

The use of empirical correlations offers an expedient route in predicting sensible thermodynamic quantities without resorting to experiments or simulations. De Moor et al. [33] used ab initio simulations to show that adsorption enthalpies and entropies for *n*-alkanes within acidic zeolites are linearly correlated with their carbon number.

Campbell and Sellers compiled a collection of experimental alkane entropies on two-dimensional catalytic surfaces [5, 6]. The key finding of their work was that the ratio of the adsorbed-phase entropy to the gas-phase entropy was approximately two-thirds. They proposed a general and elegant explanation, suggesting that the adsorption of a molecule from an unhindered gas phase onto a two-dimensional (2D) surface would eliminate a dimension of translational freedom, i.e., the adsorbate behaves as a 2D gas. This correlation was found to hold across many molecules, spanning $50\,R$ of entropy space with a standard deviation of only $2\,R$ (where $R$ is the universal gas constant). Campbell and Sellers' correlation was observed for other two-dimensional surfaces: Otyepková et al. calculated the adsorption entropy of a chemically diverse set of molecules adsorbed onto organic "van der Waals" materials using inverse gas chromatography and ab initio simulations [46, 47]. Their results showed an entropic loss of approximately $40\,\%$ relative to the gas-phase entropy. Likewise, Budi et al. calculated the adsorption entropy of a set of chemically diverse molecules adsorbed on mineral surfaces using density functional theory [48]. Although predicting a larger entropic loss relative to Campbell and Sellers' correlation, their data showed a strong linear dependence between the adsorbed-phase entropy and the gas-phase entropy. Dauenhauer and Abdelrhaman [7] expanded this idea to three-dimensional frameworks by compiling experimentally determined adsorption entropies for alkanes adsorbed in nine aluminosilicate zeolites. They showed that the entropic loss upon adsorption can be linearly correlated with the molecule's gas-phase translational and rotational entropies and that the occupiable volume of a zeolite is a useful descriptor in predicting such losses.

While the simple elegance of this nearly linear correlation of entropies is suggestive of an underlying physical origin, the extent of its applicability is unknown. This is of special concern considering that the aforementioned data is almost entirely for

industrially relevant alkanes. Therefore, within this Chapter, we avoid this practical constraint by using Monte Carlo simulations to quantify the adsorption entropies for a diverse assortment of molecules in zeolite-like materials. In particular, we use the Transferable potentials for Phase Equilibria (TraPPE) [49] force fields to calculate the adsorption entropies for 37 molecules across 10 functional categories within five pure-silicate zeolites. Our simulation results for this broad set of adsorbate molecules reveal that the correlation of the adsorbed-phase and gas-phase entropies persists, suggesting that these correlations may be much more broadly applicable than reported thus far. As a result, such correlative relationships may be exploited to predict sensible adsorption entropies for a wide range of industrially relevant fluids in zeolites and possibly other host materials.

## 2.2 Thermodynamics of Gas Adsorption

The thermodynamic quantity of primary interest here is the adsorption entropy, henceforth $\Delta s_{ads}$, which is the change in entropy of the adsorbate species upon transfer from the bulk (unconfined) state to the adsorbed (confined) state, on a molar basis. Quantification of $\Delta s_{ads}$ from laboratory or computational measurements depends on both the thermodynamic constraints on the starting and ending states (e.g., fixed pressure, fixed volume, etc.) and the thermodynamic conditions of the measurement (e.g., low pressure or infinite dilution, high pressure or near-saturation loading). The entropy results presented in ref [7] are obtained from various approaches, often based on either the Langmuir adsorption model or a modification thereof using the Henry's law constant [50, 51] or a high-level quantum-mechanical calculation [33], among others. Comparison between varying sources for the adsorption entropy is facilitated by the historical belief that $\Delta s_{ads}$ is largely unaffected by temperature [52, 53] and, for the purposes of those papers, by ignoring the effect of adsorbate loading. As a

generalization, however, the adsorption entropy can be defined through the principal thermodynamic relationship

$$\Delta G^0_{ads} = \Delta H^0_{ads} - T\Delta S^0_{ads} \tag{2.1}$$

where $\Delta G^0_{ads}$ and $\Delta H^0_{ads}$ are the Gibbs free energy of adsorption and the enthalpy of adsorption, respectively, where the superscript $^0$ indicates that the change is measured at some standard state (usually 298 K and 1 bar). From an experimental perspective, one can obtain $\Delta H^0$ and $\Delta G^0$ through measurements of the isosteric heat of adsorption ($Q_{st}$) and an adsorption equilibrium constant, respectively, which can then yield $\Delta s^0_{ads}$ by manipulation of eq 2.1.

For example, and as done for some of the data presented in ref [7], eq 2.1 may be manipulated by introduction of the Langmuir isotherm to express the adsorption entropy as [54, 55]

$$\frac{\Delta s^0_{ads}}{R} = \frac{\Delta h^0_{ads}}{RT} + \ln\left(\frac{k_H p^0}{\Gamma^\infty}\right) \tag{2.2}$$

in which $k_H$ is the Henry constant, $p^0 = 1\,\text{bar}$ is the standard-state pressure, and $\Gamma^\infty$ is the adsorption capacity; lowercase $s$ and $h$ identify those thermodynamic quantities as intensive molar quantities.

Inherent in the transformation of eq 2.1 to eq 2.2 are a number of assumptions, including (1) the aforementioned Langmuir isotherm describing the adsorbed phase, (2) ideal gas behavior for the bulk adsorptive phase, and (3) $p >> p^0$ such that the adsorption condition approaches the maximum loading $\Gamma^\infty$. (Additionally, we note that eq 2.2 is derived from an excess view of adsorption, with the usual caveat that the excess adsorption is an acceptable representation of the absolute adsorption [54, 56, 57].)

Equation 2.2 and similar expressions (cf. ref [55]) are, of course, not the only routes to quantification of the adsorption entropy. An alternative limiting case is that of infinite dilution or the Henry's law regime which facilitates the use of statistical mechanics for defining a version of $\Delta s_{ads}$ that, in turn, is easily implemented in numerical calculations. For a full discussion of the statistical mechanics and thermodynamics of adsorption in the Henry limit, we point the reader to the work of Sarkisov [58]. In particular, Sarkisov presented the key results for the Henry coefficient and differential enthalpy of adsorption for a system composed of a nondeformable adsorbent and a rigid (no internal degrees of freedom) adsorbate:

$$\begin{aligned} k_H &= \frac{\left\langle \exp\left[-\beta U_{fs}\left(\mathbf{r},\boldsymbol{\psi}\right)\right]\right\rangle}{RT\rho_S} \\ \Delta h_{ads}^{\infty} &= -R\left[\frac{\partial \ln k_H}{\partial(1/T)}\right]_{\Gamma} = \frac{\left\langle U_{fs}\left(\mathbf{r},\boldsymbol{\psi}\right)\exp\left[-\beta U_{fs}\left(\mathbf{r},\boldsymbol{\psi}\right)\right]\right\rangle}{\left\langle \exp\left[-\beta U_{fs}\left(\mathbf{r},\boldsymbol{\psi}\right)\right]\right\rangle} - RT \end{aligned} \tag{2.3}$$

(We note that $\Delta h_{ads}^{\infty} = -Q_{st}^{\infty}$, where $Q_{st}$ is the isosteric heat of adsorption.) In eq 2.3, $R$ is the universal gas constant, $\rho_S$ is the density of the adsorbent, $U_{fs}$ is the adsorbate–adsorbent (i.e., fluid–solid) potential energy, and $(\mathbf{r},\boldsymbol{\psi})$ identifies the adsorbate positions and orientations, respectively. Additionally, the $\langle\rangle$ brackets in this case indicate an average over all positions and orientations. Finally, the superscript $^{\infty}$ notation denotes the limit of infinite dilution.

The expressions in eq 2.3 show that the Henry coefficient and infinite dilution differential enthalpy of adsorption can be computed from spatial averages of terms related to $U_{fs}$. It is easy to compute these quantities in the context of (classical) molecular simulation after introducing some model of the fluid–solid potential energy (e.g., an all-atom Lennard-Jones (LJ) model, point charges, and appropriate combining rules [59]). In fact, no actual molecular simulation is necessary as both expressions

in eq 2.3 are spatial averages that may be computed via Monte Carlo integration.

Given the limit of infinite dilution ($N/V \to 0$), we may represent the adsorption process as the transfer of a single adsorbate molecule from a bulk gas state to a confined (adsorbed) state, where both states are represented by canonical ensembles at the same $T$ and volume $V$, i.e.,

$$\Delta s^{\infty}_{ads} = S_{ads}\left(N = 1, V, T\right) - S_{gas}\left(N = 1, V, T\right) \tag{2.4}$$

As we describe in the Appendix section, the entropy $S$ can be related to the canonical ensemble partition function, eventually resulting in the following expression for the adsorption entropy in the limit of infinite dilution:

$$\frac{\Delta s^{\infty}_{ads}}{R} = \frac{\Delta h^{\infty}_{ads}}{RT} + \ln\left[RT\rho_S k_H\right] + 1 \tag{2.5}$$

where $k_H$ and $\Delta h^{\infty}_{ads}$ are identical to the definitions in eq 2.3. We point out the similarity of the above equation to the adsorption entropy given in eq 2.2; the adsorption enthalpy and logarithm of the Henry constant appear in both. In contrast, however, the standard state chosen for our expression for $\Delta s^{\infty}_{ads}$ is the low-pressure limit, which may be interpreted as one adsorbate per simulation cell. In this low-pressure regime, the adsorbate entropy is highly sensitive to coverage. In particular, Campbell, Sprowl, and Árnadóttir have shown that the molar entropy of adsorbates on Pt(111) increases by 2.303 $R$ for every factor of 10 decrease in its coverage [60].

Our expression for the adsorption entropy offers an important computational advantage over the Langmuir-based approach in eq 2.2: all of the terms may be computed in a single Monte Carlo integration. As mentioned earlier, after selecting models of the adsorbent and the adsorbate–adsorbent potential energy, one can compute the adsorption entropy via straightforward integration. These integrations are rapid in

comparison to full Monte Carlo simulation (which would be necessary to compute both the Henry constant and adsorption capacity that are used in the Langmuir theory).

Finally, since the calculations we present in the following sections are actually for the adsorbed-phase entropy ($s^{\infty}_{ads}$) as a function of the gas-phase entropy ($s^{0}_{gas}$), it is necessary to clarify how those quantities are calculated. First, the molar bulk-phase entropy for our model molecules is calculated from a sum of rigorous statistical mechanical expressions for the single-molecule translational and rotational entropies for rigid molecules (denoted $s^{0}_{trans}$ and $s^{0}_{rot}$ in this work) [61, 62] and the molar vibrational entropies ($s^{0}_{vib}$) as obtained from the NIST CCCBDB [63]. We note that the translational entropy is computed using a volume corresponding to one molecule of an ideal gas at the selected standard state of 300 K and 1 bar. We admit that this is a mixing of reference states (a nonzero-pressure standard state for the gas phase, but infinite dilution for the adsorbed phase), but some choice along these lines is necessary to avoid a divergence in the gas-phase entropy that would appear at zero pressure. This convention is similar to that used in refs [5] and [7].

Finally, the adsorbed-phase entropy is estimated by

$$s^{\infty}_{ads} = s^{0}_{gas} + \Delta s^{\infty}_{ads} \tag{2.6}$$

Before proceeding, we briefly reiterate our reasoning for the use of our entropy approximation in eq 2.6. The adsorption entropy of the infinitely dilute state, e.g., one adsorbate per simulation cell, serves as a proxy for the actual, concentration-dependent adsorption entropy. Use of this adsorption entropy is a computationally advantageous choice, as it allows for the use of rapid Monte Carlo integration for the computation of the Henry constant and isosteric heat, as opposed to the comparatively lengthy simulation of the full adsorption isotherm that would be necessary

at a nondilute adsorbate concentration. Consequently, the Monte Carlo integration approach allows us to investigate many more adsorbates than would be possible via full simulation.

## 2.3 Computational Models and Methods

The following section describes the computational techniques that we used to calculate the adsorption entropy and, hence, the adsorbed-phase entropy from a relationship given in eq 2.5, as derived in the Appendix section. Our technique requires the selection of models for the zeolite adsorbents and adsorbate species and adaptation of a Monte Carlo integration technique (cf. Widom test particle method [64, 65]) to our adsorption system.

The adsorption system for our calculations is composed of a zeolitic adsorbent material and an adsorbate species. The adsorbents used here are the faujasite (FAU), Linde type A (LTA), mordenite (MOR), mordenite framework inverted (MFI), and ferrierite (FER) zeolite topologies, in their pure-silicate forms (i.e., no aluminum substitutions). The idealized structure of each silicate was obtained from the IZA database [66], and then periodically replicated to form an adsorbent cell of adequate size (cf. discussion of the cutoff radii below). The adsorbent framework was kept entirely rigid.

The adsorbate species used within this study were selected on the basis of (1) having an implicit hydrogen-based TraPPE force field, (2) the availability of vibrational entropy from the NIST Computational Chemistry Comparison and Benchmark Database (CCCBDB) [63], and (3) the ability to physically fit within our smallest cage framework (i.e., omitting molecules such as cyclo-octadecane and 2-ethylhexyl acrylate). The geometry of each adsorbate was obtained by using Open Babel [67] to convert the SMILES strings of each molecule into three-dimensional coordinates and

relaxing the structure by minimization of the GAFF potential [68]. These energy-minimized structures were used for the subsequent test particle insertions in the unaltered form, which allowed us to ignore energy terms related to bond bending, stretching, or rotation. The respective gas-phase entropies were calculated by summing the ideal gas equations for translational, rotational, and vibrational entropies at the same conditions (see Sec. 2.2). The gas-phase vibration entropy was obtained from the CCCBDB [63] with a CCSD(T) level of theory and a cc-pVDZ basis set and was assumed to be preserved upon adsorption.

As mentioned previously, the adsorbates were modeled using the TraPPE force field [49, 69] (i.e., Lennard-Jones parameters and point charges for constituent pseudo-atoms), while the adsorbents were modeled using the TraPPE-Zeo [70] force field. As specified in both TraPPE and TraPPE-Zeo, Lorentz–Berthelot combining rules were used to determine the cross-site Lennard-Jones parameters. The Ewald sum technique was used to compute the Coulomb energy of adsorbates whose TraPPE parameters include point charges, with consistent Ewald damping parameter $\alpha = 6.0/L_{min}$ ($L_{min}$ is the smallest side length of the simulation cell) and Fourier vectors with $k^2 < 27$.

The calculation of the adsorption entropy (as defined by eqs 2.3 and 2.5) was done via Monte Carlo integration (equivalent to the Widom test particle method [64, 65]), in the same mode described by Sarkisov [58]. First, a position and orientation of the adsorbate is generated randomly, and the adsorbate–adsorbent potential energy $U_{fs}(\mathbf{r}, \boldsymbol{\psi})$ for that position/orientation is computed for the TraPPE-derived model. Then, the appropriate terms in eq 2.3 are accumulated, and the procedure is repeated for $10^7$ random positions and orientations. This number of trials was satisfactory for numerical convergence for nearly all adsorbent/adsorbate combinations, where the $\Delta s^{\infty}_{\text{ads}}$ measurements were considered converged when the change in the measured value associated with an order-of-magnitude increase in the number of trials was less

than 5 %. The two exceptions to these convergence criteria were FAU/1,3,5-dioxane and MOR/ethyl-methyl-ether which were still varied by approximately 10 % when the number of trials was increased from $10^7$ to $10^8$. Subsequent calculations at $10^9$ trials for these two systems reduced the variation to less than 6 %. More importantly, the longer runs for these two systems did not affect any of the trends or conclusions that we discuss in the following sections. Finally, we also computed appropriate moments of the accumulated terms to allow for the estimation of uncertainty in the Monte Carlo integrals (which was not visible on the scale of plots presented here).

The actual Monte Carlo integrations were done using the FEASST [71] molecular simulation package developed at NIST. FEASST is primarily designed for Markov chain Monte Carlo molecular simulation, but its library functions may also be utilized for the position/orientation generation step and the subsequent energy calculations. Essentially, our program used conventional FEASST operations to build the system model and then run a loop that generated trial positions/orientations and called the energy calculator for the specified number of trial insertions. An example script implementing the Monte Carlo integration (Widom insertion) routine in FEASST v0.6.0 for the TraPPE Ethane/LTA silicate system is provided in the Appendix. Finally, we used conventional molecular simulation choices in our calculations: periodic boundaries were applied at all edges of the replicated zeolite cell, Lennard-Jones and Ewald real-space energies were cut at 15 Å, and the zeolite unit cells were periodically replicated to ensure that the length of each side of the simulation cell was at least twice this cutoff distance. By constructing the simulation cell to be at least twice the cutoff distance in each Cartesian direction, an adsorbate cannot interact with its periodic image and, thus, the spatial averages in eq 2.3 are truly the infinitely dilute limit, i.e., the partial occupancy of a unit cell does not affect the partition function [60].

## 2.4 Results and Discussion

### *2.4.1 Linear Correlation of Adsorbed- and Gas-Phase Entropies*

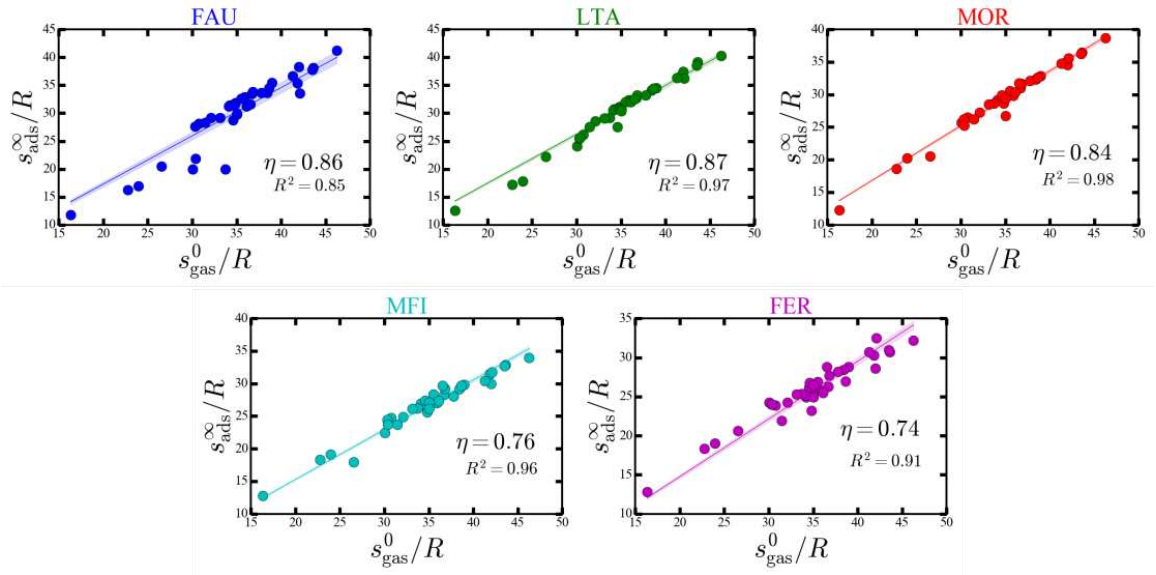


Figure 2.1: Plots of the adsorbed-phase entropy ($s^\infty_{ads} = s^0_{gas} + \Delta s^\infty_{ads}$) versus the respective gas-phase standard entropy ($s^0_{gas} = s^0_{trans} + s^0_{rot} + s^0_{vib}$) for the adsorbates listed below, adsorbed in FAU, LTA, MOR, MFI, and FER, at $T = 300\,\mathrm{K}$, as calculated using the infinitely dilute adsorption entropy described in Section 2.2. Points are the raw calculations from Monte Carlo (MC) integration, and lines are linear regressions of the MC data forced through the origin. The set of adsorbate molecules used to generate the plots consisted of 37 species in the following categories: aldehydes (2), alkanes (8), alkenes (7), ethers (5), sulfides (5), ketones (2), nitrile (1), cyclic alkanes (2), cyclic ethers (4), and aromatic molecules (1). The 95 % confidence interval of each regression is represented by the shaded region of the corresponding color. The 95 % confidence intervals for the fitted slopes ($\eta$) are FAU, 0.84–0.89; LTA, 0.86–0.88; MOR, 0.83–0.85; MFI, 0.76–0.77; FER, 0.73–0.75.

The adsorbents for our calculations consisted of the pure-silicate forms of FAU, LTA, MOR, MFI, and FER. The set of adsorbates consisted of 37 species among 10 functional categories: two aldehydes, eight alkanes, seven alkenes, five ethers, five sulfides, two ketones, one nitrile, two cyclic alkanes, four cyclic ethers, and one aromatic species; this yields 185 adsorbent-adsorbate combinations. Tabular results of $\Delta s^\infty_{ads}$ from our Monte Carlo integrations, the components of $s^0_{gas}$, and the resultant

$s^{\infty}_{ads}$ are given within the Appendix of Rzepa et al.[8] Figure 2.1 shows our calculated adsorbate entropies defined by eq 2.5, and plotted against their respective gas-phase entropies at the same temperature. The confidence intervals in $s^{\infty}_{ads}$ were below 0.02% of $\Delta s^{\infty}_{ads}$, and were therefore omitted. The primary and most important observation is that, for the most part, the adsorbed-phase entropy shows a remarkable degree of linear correlation with the gas-phase entropy.

Based on this visual linearity, we performed a least-squares regression of our data, with the requirement that the trend line passes through the origin, e.g., $s^{\infty}_{ads} = \eta_i s^{0}_{gas}$. Similar to ref [7], this added constraint is based on the conjecture that the adsorbed-phase entropy must be both positive and smaller than the gas-phase entropy, i.e., $0 < \eta_i < 1$. The $R^2$ correlation coefficients [1], except for that of FAU, are close to unity, indicating low deviation from the linear trend lines; the notable deviations for FAU will be discussed later.

Linearity in the correlation of $s^{\infty}_{ads}$ with $s^{0}_{gas}$ is the key feature of the results in Fig 2.1, as it corresponds to qualitatively similar observations in the entropy correlations disclosed in refs [7], [48], and [5]. More importantly, though, our results show that apparent linearity in the correlation of $s^{\infty}_{ads}$ with $s^{0}_{gas}$ persists for a larger set of chemically diverse adsorbates than was previously seen. As discussed in those references, such a simple correlation is an obvious opportunity for the development of an engineering correlation that allows for prediction of the adsorption entropy from a limited set of measurements (whether experimentally or computationally derived). Furthermore, we aim to build on the conclusions in refs [7], [48], and [5] by examining our own data in light of the physical arguments presented previously. As a brief review, the apparent linearity in adsorption entropy correlations derives from fundamental physical considerations. In ref [5], where their linear trend in the entropy

1. $R^2 = 0.85,\ 0.97,\ 0.98,\ 0.96$, and 0.91 for FAU, LTA, MOR, MFI, and FER, respectively.

correlations was approximately two-third, the authors offered a simplified explanation that the linear trend reflected a loss of approximately one-third of the translational and rotational entropies owing to adsorption. Dauenhauer and Abdelrahman [7] extended this argument to acidic zeolites by considering an additional loss of rotational entropy dependent on the adsorbate size and the pore volume. (Pure linear correlation of $s^{\infty}_{ads}$ with $s^{0}_{gas}$ is broken by the arguments in ref [7], but a visual linear correlation persists. We return to this point later in our work.) Before proceeding, we also note that a numerical comparison between our entropy values and those of ref [7] must be done with care; the gas-phase rotational and translational entropies (i.e., the $x$-axis of Fig 2.1) of our TraPPE-model molecules differ from experimental values. This is because the TraPPE molecules include pseudo-atoms, which represent hydrogens implicitly. Although the mass of the TraPPE-model molecules does not change, the symmetry number and principal moments of inertia within the rotational entropy equation are different.

The linear trend lines in Fig 2.1 yield the following slopes: FAU:0.86, LTA:0.87, MOR:0.84, MFI:0.76, FER:0.74. The slopes appear to be related to zeolite size as this sequence of decreasing slope corresponds to the sequence of decreasing pore size. Further discussion on this point is reserved for later in this paper (cf. Section 2.4.2). We immediately point out that all of these slopes are larger than the proposed two-third based on adsorption to a flat surface. Given that the $x$-axis is the total entropy, not just translational and rotational contributions, we should not expect a precise slope of two-third. These slopes are, additionally, larger than the slopes for linear trend lines computed from the data of Dauenhauer and Abdelrahman (not present in their paper, but easily obtained using data in the Appendix of ref [7]): FAU: 0.80, MOR: 0.75, MFI: 0.65, and FER: 0.59 (and other topologies not studied here).

As a specific example, Fig 2.2 reproduces the entropy correlation for MFI (same

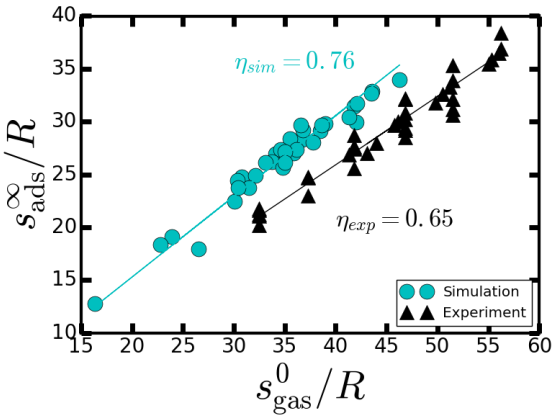


Figure 2.2: Comparison of the linear correlation of adsorbed- and gas-phase entropies for MFI adsorbent, from the present simulation/model-based results (at 300 K) and experimental calculations (at various temperatures) from Dauenhauer and Abdelrahman. As noted in the figure, the slopes of the linear trend lines are 0.76 and 0.65 for the simulation and experimental data sets, respectively. The simulation/model results are for the diverse set of adsorbates studied here, whereas the experimental results are for alkanes of up to eight carbons, with propane as the smallest adsorbate (and the corresponding smallest $s^0_{gas}$). We note that the model-based entropies are offset compared to the experiment due to the pseudo-atom structure of TraPPE molecules mentioned in the text.

as Fig 2.1) but includes the data from ref [7] for the same material; the slope based on experimental measurements is 0.65 versus 0.76 for our model-based measurements. Qualitatively identical differences were found for FAU and MOR, the other two zeolite structures common to the two sets of results. We highlight two points in examining this difference. First, our results are for stock (i.e., nontuned), model-based representations of the adsorbates (e.g., rigid molecules composed of TraPPE pseudo-atoms) and adsorbents (e.g., LJ and point charge representations of silicon and oxygen). Second, the adsorbents are nonacidic, pure silicates in our models, versus aluminum-substituted (with varying Si–Al ratios) and cation-balanced zeolites in the experiments. From our perspective, the most critical result is that the linear scaling appears for a wide range of adsorbate species even without specifically tuning the potential energy models for our adsorption systems or simulating the actual zeolites. This is

again suggestive of the arguments for systematic entropy loss given in the previous works. However, the difference in slopes highlighted in Fig 2.2 is an important distinction between our work and that referenced above. Our results suggest that while the appearance of a linear trend in the correlation of adsorbed-phase and gas-phase entropies may derive from underlying, common physical effects, the strict details of the total entropy loss function for our model systems differ from the experiment.

### *2.4.2 Correlation of Adsorbed-Phase Entropy with Adsorbent Characteristics*

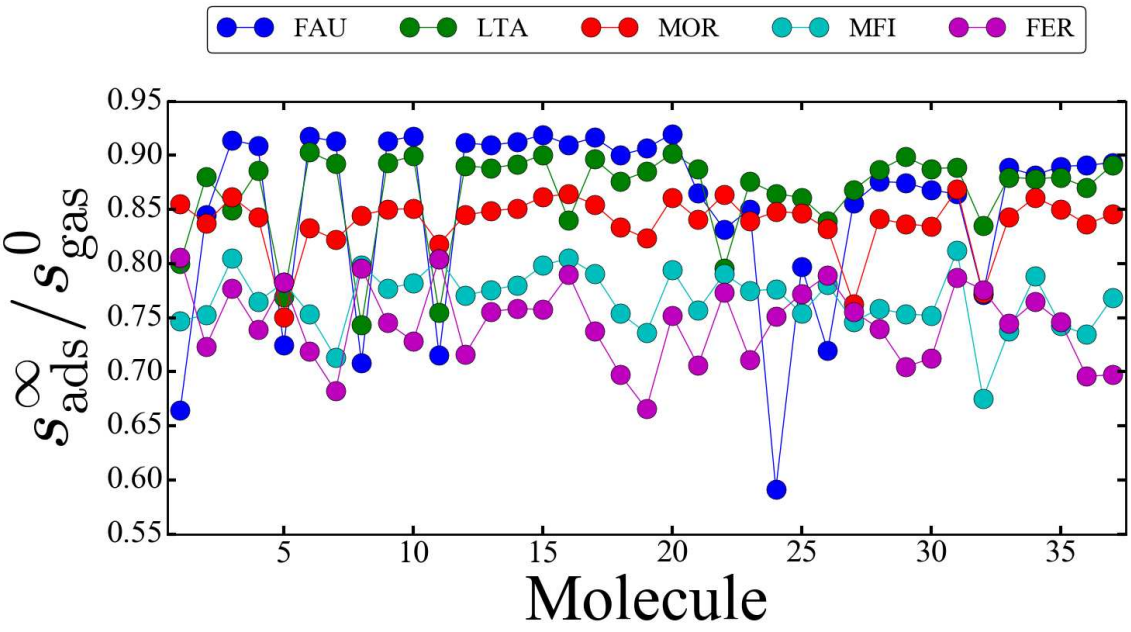


Figure 2.3: Ratio of the adsorbed-phase to gas-phase entropy, $s^{\infty}_{ads}/s^{0}_{gas}$, for each adsorbate molecule identified by an integer label. Connecting lines are simply an aid to the eye, with no physical interpretation implied. The symbol and connecting line colors identify the adsorbent framework type and correspond to the color scheme in Fig 2.1. As discussed in the text, the outliers are small adsorbates and chemically unique cyclic ethers.

In this subsection, we discuss the variance of the linear correlations in Fig 2.1

across the set of adsorbates and adsorbents studied here. The first metric we examine more closely is the ratio $s^{\infty}_{ads}/s^{0}_{gas}$ for each adsorbate–adsorbent combination, as forcing the trend lines in Fig 2.1 through the origin is analogous to taking this ratio to be constant for each adsorbent. In doing so, we look more closely at how constant that ratio actually is. Figure 2.3 plots $s^{\infty}_{ads}/s^{0}_{gas}$ for all 185 adsorbate–adsorbent combinations, with the adsorbate species identified by an integer on the $x$-axis according to the Appendix of Rzepa et al.[8] and the adsorbents identified by the same color coding as in Fig 2.1. Adsorbates are also grouped by chemical type (See Appendix of Rzepa et al.[8]) though we stress that the $x$-axis integers should not be used to draw any correlative conclusions; the purpose of this figure is to put every entropy ratio from our calculations on a single plot. As expected from the results in Fig 2.1, for a given adsorbent, the entropy ratio is remarkably uniform across the adsorbates, with only a few noticeable outliers. Furthermore, the data in the figure appear to be clustered into three adsorbent groupings. Based on the "largest cavity diameter" (LCD) descriptor of First et al. [72] (the LCD being the diameter of the largest sphere that the adsorbent can accommodate), our zeolites are grouped into the following adsorbent categories: (1) spherical pores with an LCD of $\approx$ 7 Å (FER and MFI), (2) cylindrical channels with an LCD of $\approx$ 7 Å (MOR), and (3) spherical pores with an LCD of $\approx$ 12 Å (FAU and LTA) [2]. The implication is that the entropy ratio is effectively a constant for materials with similar characteristic pore size and shape and that (again, except for a few prominent outliers) the adsorbate characteristics are relatively unimportant. This raises the possibility that models based primarily on adsorbent characteristics may be an effective tool for modeling the adsorbed-phase entropy.

2. The pore size descriptors for FAU, LTA, MOR, MFI, and FER are MSD = 11.24, 11.05, 6.7, 6.36, and 6.31 Å [73]; LCD = 11.9, 11.7, 6.5, 7.0, and 7.0 Å [72], $V_{occ}$ = 370.0, 311.4, 223.0, 177.4, and 198.8 Å$^3$ [73]

While the predominant trend in Fig 2.3 is a roughly uniform entropy ratio for each adsorbent across our set of adsorbates, there are some noticeable outlier adsorbates, particularly for the FAU and LTA topologies. For example, adsorbates 1, 5, 8, 11, and 26 are acetaldehyde, methane, ethane, ethylene, and dimethyl ether, respectively; these are all relatively small molecules. Furthermore, the adsorbents (FAU and LTA) have small secondary pores that are connected to the primary pores (LCD $\approx$ 12 Å) via inaccessible apertures. It is known that these small pores can be accessed by small adsorbates in Monte Carlo-type simulations since the adsorbate positions are generated randomly rather than by physical trajectories [74], resulting in unphysically large adsorption, which would be reflected in a larger entropy of adsorption and, hence, smaller entropy ratio. Our integration technique also relies on the random generation of adsorbate positions; we have confirmed by direct visualization that these outliers result from adsorbates accessing pores that are actually inaccessible. The other group of prominent outliers are adsorbates 22, 24, and 32, corresponding to 1,3-dioxolane, 1,3,5-trioxane, and acetonitrile, respectively. These adsorbates are two cyclic ethers and one nitrile. In this case, we suspect that the unique chemical functionality of these adsorbates, in comparison to the other TraPPE species, results in a larger-than-expected entropy loss.

Finally, the $\eta$ slopes in Fig 2.1 and the examination of adsorbate-specific entropy ratios in Fig 2.3 are suggestive of an entropy loss model based primarily on certain adsorbent characteristics. As pointed out above, $\eta$ is roughly the same for (1) FAU and LTA and (2) FER and MFI; the adsorbents in each of these two groups have roughly the same LCD and predominantly cage-like, spherical pores. The MOR topology has, as mentioned previously, an LCD similar to FER and MFI but with channel pores. Other pore descriptors are, of course, available as well, and we examine two others here. First, as a compliment to the LCD descriptor, we include the "maximum

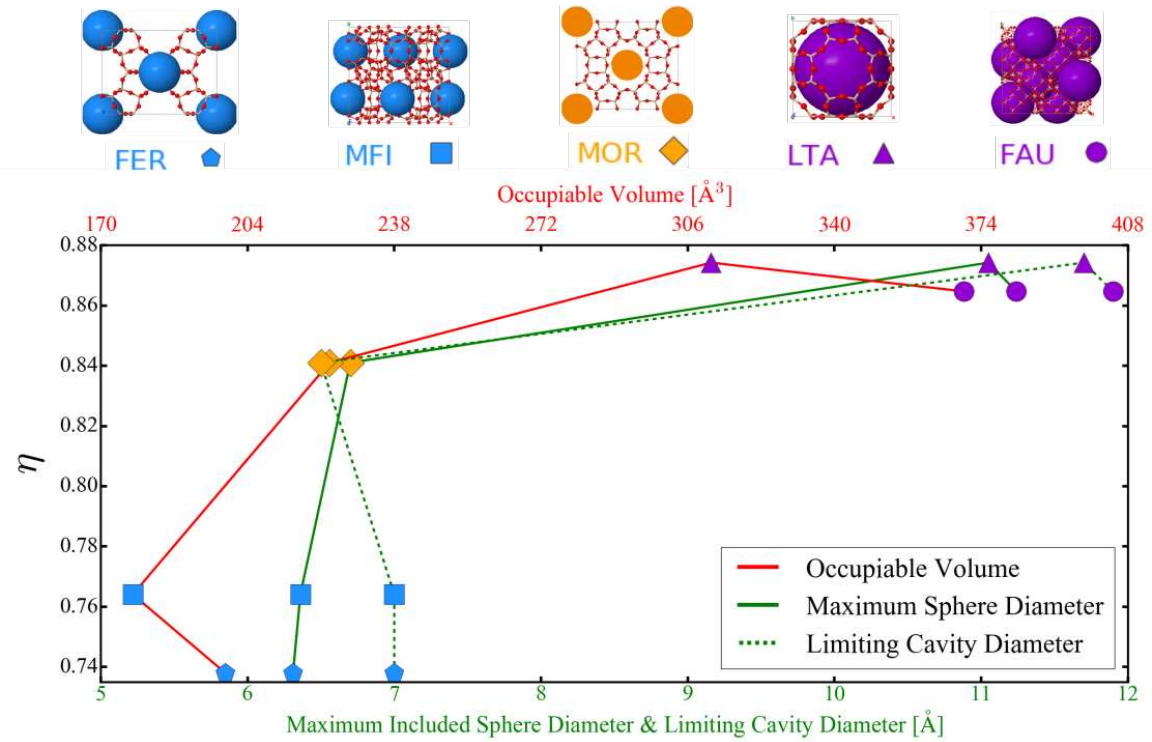


Figure 2.4: Plot of $\eta$ from Fig 2.1 versus the LCD, MSD, and occupiable volume (see the main text for definitions and values[2]). Symbols are color-coded as follows. Blue: materials with an LCD of $\approx 7$ Å, with spherical pores; orange: material with an LCD of $\approx 7$ Å, with cylindrical channels; purple: materials with an LCD of $\approx 12$ Å, with spherical pores. Connecting lines are an aid to the eye to identify the MSD (solid green), LCD (dotted green), and occupiable volume (solid red) for each material. Zeolite graphics are from the Zeomics online database.

included sphere diameter" (MSD), which is the largest sphere pore identified in a calculated pore size distribution [73]. Second, we also include a different type of geometric descriptor, the "occupiable volume" ($V_{occ}$), which is defined as the volume per 1000 Å of the crystal cell that can be accessed by the center of probe molecules with diameter 2.8 Å [73]. These three pore size descriptors capture both the size of the largest pore features and the overall pore volume.

In Fig 2.4, we plot the $\eta$ slope for each adsorbent as a function of these three pore metrics; diameter-based metrics are on the lower $x$-axis and occupiable volume is on the upper $x$-axis. The important result shown in Fig 2.4 is that, regardless of the

metric used to characterize the zeolite adsorbent, the slope of the entropy relationship in Fig 2.1 follows the same qualitative trend. Starting at the largest pore adsorbents in terms of any of the three metrics, $\eta$ decreases slowly with decreasing pore size, before decreasing more rapidly to values in the vicinity of $\eta = 0.75$. Despite plotting all three metrics on the x-axes of Fig 2.4, we imply no quantitative relationship between MSD/LCD and $V_{occ}$; the purpose is to show the common trend in the correlation of $\eta$ with different size metrics. (As an aside, we note that the upper and lower x-axes have common scaling, i.e., they share a common $x = 0$ origin [though not visible] given the upper bounds of LCD/MSD = 12 Å and $V_{occ} = 408\,\text{Å}^3$; we acknowledge that the trend in the figure can be distorted by adjusting the scale and/or origin. Our plot attempts to avoid biasing the interpretation of the data by using a common origin.) Based on Fig 2.4, one may be motivated to construct an empirical fit of $\eta$ as a function of one or more pore size metrics, from which one could then approximate the adsorbed-phase entropy from the gas-phase entropy. However, we suggest that a larger set of zeolites with more variation in pore size be examined computationally to provide a better basis for generating an empirical predictor of $\eta$.

Overall, we wish to stress that the entropy ratio is primarily a function of the adsorbent characteristics and that it may be quite adequate for engineering purposes to approximate the adsorbed-phase entropy as a fraction of the gas-phase entropy, with that fraction being largely independent of the adsorbate identity. This simple, fractional scaling of $s^{\infty}_{ads}$ with $s^{0}_{gas}$ is, of course, a highly simplified model of the adsorbed-phase entropy and includes minimal physical insight or intuition. As shown in the next section, one can introduce physically intuitive arguments into more complicated models of entropy than that presented above.

### *2.4.3 Empirical Model of Entropy Based on Adsorbent Identity*

While the linear entropy scaling in Figures 2.1 and 2.3 is quite successful for the TraPPE-based adsorption systems, it is ultimately an oversimplified correlation. We seek a physics-based structure–topology–entropy relationship that is generic and applicable across molecule classes and zeolites. As a first step to this end, motivated by the arguments in refs [5] and [7], we considered a simple linear model wherein each contributor to the entropy is affected differently. However, it has been argued that the translational contribution to the adsorbate entropy (for alkanes) is relatively uniform across frameworks, with the emphasis being on the disparate rotational contributions [7, 75]. For example, Denayer et al. have demonstrated that the driving force behind the zeolitic separation of alkane/isoalkane mixtures is exclusively the difference in rotational loss, represented by a ratio between the molecule's radius of gyration with the cavity's largest inscribed van der Waals radius [75]. Therefore, we consider a model where $s^{\infty}_{ads}$ is composed of (1) two-third translational entropy (equivalent to the loss of one-third of the gas-phase translational entropy) and (2) a fraction of the gas-phase rotational entropy. As mentioned previously, the vibrational entropy is assumed to be unaffected by adsorption or confinement.

We propose the simplest model consistent with the proposition that the adsorbed-phase entropy is composed of the translational and rotational entropies and all of the vibrational entropy of the gas state. Mathematically, this may be stated as

$$s^{\infty}_{ads} = \alpha s^{0}_{trans} + \beta s^{0}_{rot} + s^{0}_{vib} \quad \text{where} \quad \alpha = 2/3 \quad \text{and} \quad 0 < \beta < 1 \tag{2.7}$$

Then, motivated by our previous results, which showed that the linear correlation of adsorbed- and gas-phase entropies could be based primarily on adsorbent identity, we make the further approximation that $\beta$ is constant for each adsorbent species

irrespective of the adsorbates. In other words, the loss in rotational entropy is dictated exclusively by the adsorbent characteristics and the loss in translational entropy is the aforementioned one-third of the gas-phase term. We make this approximation owing to the results in the previous sections showing the predominant effect of the adsorbent in correlations for the adsorption entropy.

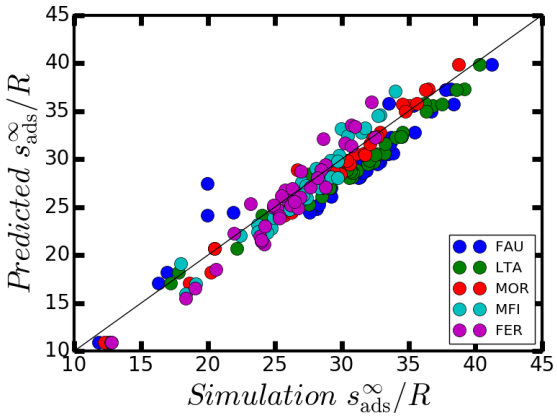


Figure 2.5: Plot of the adsorbed-phase entropy $s^{\infty}_{ads}/R$ as computed from the multilinear regression in eq 2.7 versus that calculated in the Monte Carlo integration ("simulation") described in Section 2.3. Particular zeolite topologies are identified by the color codes in the legend. The $y = x$ line is plotted for evaluation of the multilinear fit.

| **Topology** | MSD(LCD) | $\beta$ |
|---|---|---|
| FAU | 11.2(11.9) | 1.00 |
| LTA | 11.1(11.7) | 1.00 |
| MOR | 6.7(6.5) | 1.00 |
| MFI | 6.4(7.0) | 0.80 |
| FER | 6.3(7.0) | 0.71 |

Table 2.1: Fitting Parameters for $\beta$ ($\alpha = 2/3$) for the Multilinear Regression of $s^{\infty}_{ads}$ to eq 2.7 for Each Zeolite Popology Studied Here with It's Respective MSD

Based on these arguments, we performed a simple multilinear least-squares regression of our entropy data with eq 2.7 and the constraints shown. The full set of results are shown in Fig 2.5, where the calculated ("simulated") $s^{\infty}_{ads}$ is on the $x$-axis and the predicted value (eq 2.7 is on the y-axis. The $\beta$ values from the regressions

are given in Table 2.1, with the $\alpha$ parameter held constant at two-third. As is hoped for this type of model, the data are clustered around the $y = x$ line, with only a few prominent outliers. In fact, the three most noticeable outliers are for the FAU material, and these correspond to 1,3,5-trioxane, acetaldehyde, acetonitrile outliers already pointed out in our discussion of Fig 2.3. The RMSE of the multilinear fits are 2.6, 1.8, 1.1, 1.4, and 1.8 for FAU, LTA, MOR, MFI, and FER, respectively. The $\alpha$ and $\beta$ coefficients may be thought of as survival coefficients for the translational and rotational entropy, so trends in them may be helpful for physical interpretation of the predicted results. First, we note that the zeolite materials are in order of decreasing MSD in Table 2.1. The table then reveals that $\beta$ decreases with decreasing pore size. The simplified model thus offers the following physical interpretation: decreasing pore size results in a proportionately larger loss of rotational entropy, entirely consistent with the previous results [7]. In other words, smaller pores are a hindrance to free rotation of an adsorbate, resulting in the loss of rotational entropy.

However, we stress that further work is needed to clarify whether the trend in $\beta$ is a real effect or an artifact of an oversimplified model. By the same token, we emphasize that our pool of molecules excludes especially large species; we anticipate that data sets with extraordinarily hindered molecules will have much lower $\beta$ values, but will ultimately follow a similar trend.

Ultimately, we find that the adsorbed-phase entropy can be fit to a simple model based on the survival of components of the gas-phase entropy and that this model is quite accurate except for a few outliers. Thus, in addition to the extremely simple linear correlation of the adsorbed- and gas-phase entropies presented in Fig 2.1 and associated discussion, we offer a more complex, but more physically intuitive, model based on eq 2.7 that may be used to predict the adsorbed-phase entropy solely from the gas-phase entropy with a satisfactory degree of accuracy.

## 2.5 Summary and Conclusions

The entropy of an adsorbed molecule is an important thermophysical property in evaluating the use of adsorbent for applications including, but not exclusive to, gas separation, catalysis, and energy storage. The number of actual and hypothetical zeolite materials, as only one example of a class of industrially relevant and resilient adsorbent materials, necessitates the provision of predictive models for the adsorbed entropy, whether those models are based on robust physical principles or engineering correlations. Previous works [5, 7, 33] have demonstrated that the adsorbed-phase entropy may be predicted with good accuracy by applying straightforward loss functions to the gas-phase entropy, though the extent of those correlations was not clear as the correlations were generated for a limited set of adsorbates, primarily alkanes.

Here, we examined a larger and more chemically diverse set of adsorbates by computing the entropy in a model-based approach using 37 adsorbates available in the TraPPE force field library and five pure-silicate zeolites, for a total of 185 unique calculations. Our results show that linear correlation of the adsorbed- and gas-phase entropies persists for adsorbates and adsorbents not previously considered, though the form of the linear correlation differs from previous work [5, 7]. Furthermore, the correlation may be cast as a highly simplified linear scaling function applied to the gas-phase entropy, where the scaling coefficient is a constant for a particular adsorbent material, independent of the identity of the adsorbate species (Fig 2.3). Additionally, our data suggest that the scaling coefficient correlates with pore size characteristics (Fig 2.4), though more adsorbent materials need to be evaluated to confirm the observed correlation.

Beyond the simple linear scaling of $s^{\infty}_{\text{ads}}$ with $s^{0}_{\text{gas}}$, we also fit the adsorbed-phase entropy to a multilinear expression based on physical arguments about the loss of translational and rotational entropies (eq 2.7). The multilinear fitting model success-

fully reproduced the TraPPE-model-based entropy calculations, which offers an alternative approach to modeling the adsorbed-phase entropy. In addition, this multilinear model is suggestive of opposing trends in the translational and rotational entropies (decreasing rotational entropy and increasing translational entropy with decreasing pore size), but more work is needed to determine whether this is a mathematical artifact or a real trend. Regardless, our model is consistent with prior approaches that model the adsorbed-phase entropy according to survival functions applied to the rotational and translational entropies.

Finally, as an overall summary, our results suggest that simple linear models can be used to approximate the entropy of an adsorbed phase, whether through a single scaling factor ($\eta$ in eq 2.5) or through a more complicated multilinear expression (eq 2.7). In either case, our results point to engineering correlations based on a limited set of actual measurements; for example, evaluation of a new material could be done by measuring the Henry constant and isosteric heat of a few adsorbates and then using that limited set of data to generate $\eta$ or $\alpha$ and $\beta$, based on the trends identified here. While these correlations have been identified specifically for pure-silicate zeolites, it is anticipated that the resultant trends may also appear for materials such as metal-organic frameworks and microporous polymers, which should be the objective of future work.

# CHAPTER 3

# DFT BASED MICROKINETIC MODELING OF CONFINEMENT DRIVEN [4+2] DIELS-ALDER REACTIONS BETWEEN ETHENE AND ISOPRENE IN H-ZSM5

## 3.1 Introduction

One of the most successful advancements within the petro-chemical industry has been the application of zeolites, which are crystalline, alumino-silicate frameworks that consist of ordered, cage-like structures. Their utility arose from their distinct microporous topologies, which typically consist of subnanometer channels with various dimensionalities. Historically, they have been applied in the catalytic cracking of hydrocarbons[76] and light gas separations[77]; but have shown function in the conversion of methanol to olefins[78, 79], Friedel-Crafts alkylation[80], Beckmann rearrangement[81], Fischer-Tropsch synthesis[82], among others.[83] To an extent, all of these processes are promoted by the principle of confinement; or how well the molecules fit within the zeolite's pores and active site. This concept suggests that the adsorbates conform to the curvature of the channel, becoming stabilized by maximizing their van der Waals interactions with the framework while limiting their entropic loss.[2] Consequently, confinement promulgates zeolites as molecular sieves, where prohibitively large species that cannot fit within their voids are excluded from adsorption and/or the reaction space. Likewise, smaller species paired with large framework features may not incur any significant stability; and merely equilibriate with their bulk gas phase.[84] Optimal confinement occurs when the geometry of the reaction intermediates and/or transition states "match" their channel dimensions, ac-

quiring enough stability to catalyze their respective reactions. For example, the rates of dimethyl ether carbonylation have been shown to be significant within zeolites containing eight-membered ring (8MR) channels, but were undetectable within zeolites containing 10+ MR channels exclusively.[85] Such specificity was attributed to the enthalpic stabilization of carbo-cationic transition states by the closer proximity of the surrounding framework's oxygen atoms within the 8 MR than 10+ MR channels. Yet configurations which maximize the adsorbate's van der Waals interactions with the framework (enthalpic stability) may not be preferred due to the entropic penalty incurred from the adsorbate's loss in mobility; and as a result, adsorbates have been shown to prefer non-intuitive arrangements within the zeolite channel. For example, high-temperature alkane activation reactions[85] have shown a preference toward 8-MR channels, but contain transition states too large to be fully encompassed within them. It was discovered that these transition states were only partially confined, extending the majority of their geometries into the adjacent 12 MR channel. Such "loose" configurations were preferred because they provided enough entropic freedom to exceed the enthalpic stability otherwise gained through complete adsorption within the 12 MR channels. Overstabilizing adsorbates is also possible, and has been shown to impede reaction rates through increased diffusion limitations[86, 87], restrained product desorption, and increased intrinsic reaction barriers[14]. However confinement is not exclusive to zeolites; and has been shown to be a significant factor in catalyzing the oxidative dehydrogenation of hydrocarbons within mixed metal-oxides[88, 89] and manganese oxide based octahedral molecular sieves[90], as well as catalyzing Diels-Alder reactions within macromolecules.[91] Understanding confinement is therefore a multifaceted problem; and despite our detailed understanding on an individual basis, our mechanistic understanding and ability to make empirical predictions remains limited.

In this work, we explore the effects of zeolite confinement on a set of Diels-Alder cycloaddition reactions. These reactions are an excellent archetype to study confinement because, 1) they can occur uncatalyzed, allowing for the effects of the zeolite framework to be separated from the reaction chemistry, and 2) their capability in producing cyclic regio-isomers provides a span of variously sized, but chemically identical, species. These reactions have been well studied and applied in the pharmaceutical, agro-chemical, and oleo-chemical industries for many decades.[92] Recently, tandem Diels-Alder-dehydration chemistries have been suggested for the synthesis of aromatic monomers from bio-mass based alternatives.[93] The [4 + 2] Diels Alder reaction, henceforth called "DA reaction", is particularly interesting because it forms a six-membered cyclic product (the cycloadduct) by coupling a 1,3-conjugated olefin (diene) with a double bonded moiety (dienophile). The cycloaddition reactions of such a diene + dienophile, followed by tandem dehydrogentation on a multifunctional catalyst, has been shown to result in valuable aromatic products.[16, 18] At the core of this reaction, four $\pi$-electrons from the diene and two $\pi$-electrons from the dienophile are broken, forming two new $\sigma$-bonds that enclose the product's six-membered ring. This reaction typically occurs through a concerted mechanism, with the formation of the two $\sigma$-bonds occuring simultaneously through a single, pericyclic transition state. However, a stepwise mechanism is also possible, involving the formation of a diradical or zwitterionic intermediate.[94] Although they may be thermally driven to completion, the incorporation of homogeneous Lewis and Brønsted acids have been shown to not only catalyze these reactions, but also promote the selectivity of certain stereo and regio-isomers.[9, 10, 11, 12] Zeolites are particular for this endeavor, because they offer Lewis/Brønsted acidity alongside confinement; which can limit undesirable byproducts or isomers and promote catalysis with particular selectivity.[84] Moreover, the preservation of regio-chemistry throughout the DA reaction allows one

to control the bulkiness of the products and/or transition states, meaning a judicious choice between zeolite and reactants, whose products and/or transition state geometries appear to better fit within the channels, can be posited a-priori.

Despite being relatively unexplored, zeolites have been shown to catalyze DA reactions. However, the research has been largely focused on the DA cycloaddition between 2,5-dimethylfuran and ethene with subsequent dehydration to form xylenes; particularly due to the readily available synthesis of furans from bio-mass.[95] Williams et al.[17] have studied this reaction within zeolite HY, achieving a 75% selectivity toward para-xylene. Their findings suggest that the DA reactions occurred without a catalytic active site; rather, the reaction was promoted by confinement and the role of the Brønsted acid site was instead attributed to the catalysis of the dehydration step. Nikbin et al.[96, 16] have studied the application of HY and various alkali-exchanged Y zeolites for the same reaction using a combination of density functional theory and hybrid quantum mechanics/molecular mechanics calculations. Interestingly, their calculated DA activation barriers in HY were larger than their respective gas phase reactions, suggesting that they occurred uncatalyzed at the Brønsted active sites. Moreover, their results showed that alkali-exchanged Y zeolites (LiY, NaY, KY, RbY, CsY), in the form of Lewis acids, exhibited marginal catalytic activity toward DA reactions. On the contrary, the dehydration steps were shown to be significantly catalyzed at the Brønsted sites; and therefore they advocated that a bifunctional catalyst, one comprised of both Lewis and Brønsted acids, might better catalyze the overall process. Rohling et al.[13, 14] have investigated the role of low-silica, alkali-exchanged Y zeolites (LiY, NaY, KY, RbY, CsY) on the same reaction using Density Functional Theory and microkinetic modeling. The scope of their work was to determine the collaborative effect of multiple cation sites on the DA cycloaddition/dehydration reactions against the standard, single site model. Their

results suggest that low-silica alkali-exchanged Y zeolites are highly active catalysts for these reactions, owing to a combination of confinement-induced initial-state destabilization and transition state stabilization via ionic interactions with the collective alkali cations. In a subsequent work, Rohling et al.[15] have computationally investigated the catalytic effect of d-block cation exchanged high-silica Y zeolites on the DA cycloaddition step; and concluded that cations with less filled d-shell orbitals exhibited lower activation barriers. Apart from the familiar concerted single-step DA cycloaddition path; they also found a two-step path, which ultimately depended on the relative size of the cation. Other DA cycloaddition reactions have also been catalyzed by zeolites. For example, Bernardon et al.[19] have experimentally investigated DA reactions between isoprene and methyl-acrylate within different zeolite frameworks of various acid site densities. Among those frameworks, H-ZSM5 had the highest productivity and upwards of 91% selectivity toward the para- over the meta-cycloadduct. Ultimately, this catalytic effect was attributed to confinement. Apart from zeolites, macromolecules have also been shown to act as potential DA catalysts. Chakraborty et al.[91] have investigated the application of the macrocyclic molecule "cucurbit[7]uril" (CB[7]) as a catalyst for the DA reactions between benzene, furan, cyclopentadiene, and thiophene with ethene using Density Functional Theory. All reactions exhibited enhanced rate constants within the CB[7] molecule relative to their free state. However, the activation energy for each reaction was less favorable within CB[7] than their respective free state, suggesting that the activation entropy facilitated these reactions.

The objective of this work is to use periodic Density Functional Theory (DFT) based microkinetic modeling (MKM) to study the effects of confinement on the kinetics of the DA cycloaddition reactions between ethene (C2) and isoprene (C5) within H-ZSM5. The resulting DA cycloadducts are outlined within Figure 3.1, and there is

a single C7 product formed between isoprene and ethene (IUPAC: "1-methylcyclohex-1-ene"); there are four C10 products that are formed from two isoprene molecules, and we have organized them by their "para-" and "meta-" regiochemistry. For brevity, we will be referring to the C10 products (and corresponding reactions) as "C10-para-1,2" and "C10-meta-1,2" with the exception of 1-methylcyclohex-1-ene which will be referred to as "C7". Our choice of reactants provides an array of variously shaped (meta- vs. para-) and sized (C7 vs. C10) transition states/products whose corresponding reactions can occur uncatalyzed. This allows us to gauge the extent of confinement relative to their uncatalyzed gas phase reactions; and infer discriminating factors based on the specie's size/shape. We have chosen to work with H-ZSM5 in view of its well studied applications as a catalyst for light olefin reactions[97, 98, 99], pervasive use in industry[76], and superiority in catalyzing similar DA reactions.[19]

## 3.2 Methods

The adsorption calculations have been simulated using one full periodic unit of the ZSM-5 orthorhombic structure taken from the International Zeolite Database (IZA).[3] The unit cell consisted of 192 Oxygen and 96 Silicon atoms with 12 distinguishable tetrahedral sites (T-sites); we follow the nomenclature of IZA to identify the T-sites. The Brønsted site was formed by substituting a single Si atom with Al and adding a hydrogen on the most stable adjacent oxygen. Based on the work of Ghorbanpour et al.[100], the T7 site was chosen as the Brønsted site; it was considered to be energetically favorable, providing good access due to its location at the intersection of the straight and sine channels. The effects of topologically identical locations on adsorption within the zeolite supercell were deemed negligible based on a prior work.[101]

All electronic structure calculations were performed using the Vienna ab initio

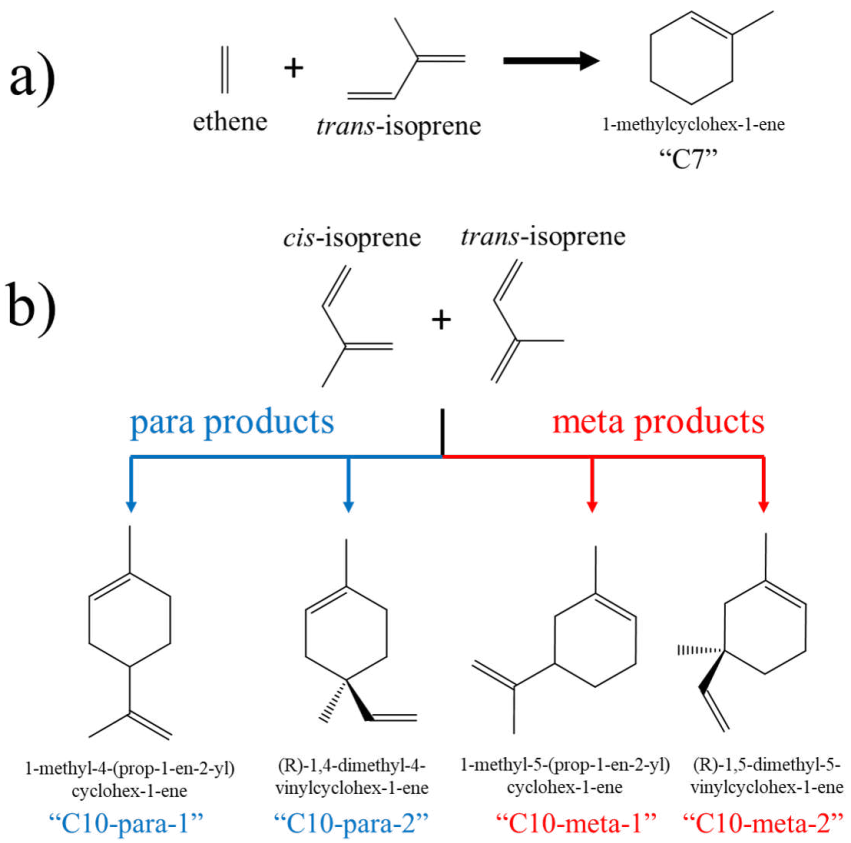


Figure 3.1: The DA cycloaddition reactions between ethene and isoprene considered within this work. a) The DA cycloaddition reaction between ethene (diene) and *trans*-isoprene (dienophile) to form the C7 cycloadduct, IUPAC: "1-methylcyclohex-1-ene". b) The possible DA cycloaddition reactions between *cis*-isoprene (diene) and *trans*-isoprene (dienophile) to form four possible C10 cycloadducts. The cycloadducts were segregated based on their respective para- and meta- regiochemistries.

simulation package (VASP)[102, 103] through plane wave density functional theory (DFT). PAW potentials[104] with generalized gradient approximation (GGA) using the Perdew-Wang-Ernzerhof (PBE)[105] exchange correlation functional with Grimme-D2(DFT-D2)[106, 107] dispersion corrections were used in view of its accuracy for small adsorbates in other zeolites.[101, 108] The convergence criterion for electronic relaxation was $10^{-4}$ eV with a plane-wave energy cutoff of 400 eV; ionic convergence was set with the criteria that the force on each ion be smaller than 0.02 eV/Å. Gaussian smearing with width 0.1 eV was used; all energies were extrapolated

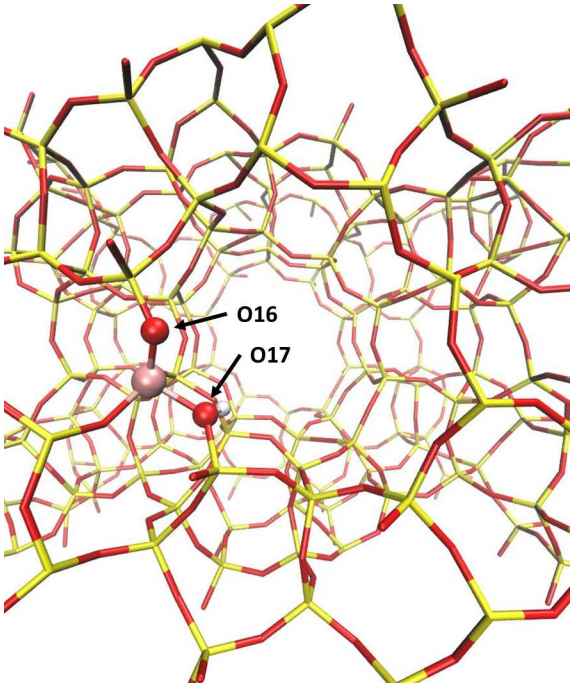


Figure 3.2: The lone structure of H-[Al]ZSM-5, where the Brønsted proton is bond to O17 pointing along the straight channel toward an adjacent O17. The O16 oxygen is also emphasized. Key: silicon (yellow), oxygen (red), hydrogen (white), aluminum (pink).

to 0 K. Gas-phase calculations were performed with a minimum of 10 Å of vacuum between periodic images. Transition states were identified by using the climbing image nudge elastic band method (NEB)[109] with seven images between the reactants and product. The ionic convergence for the NEB calculations was set with the criteria that the force on each ion within each image be smaller than 0.1 eV/Å. As a consequence of the large unit cell size, the Brillouin zone was sampled using only the gamma point. Keeping atoms constrained, optimization of the zeolite cell parameters resulted in: ($20.2 \times 19.9 \times 13.3$ Å,$\alpha$,$\beta$,$\gamma$=90°).

The enthalpies, entropies and Gibbs free energies for reactants, products, and transition states were derived using statistical thermodynamics. The harmonic normal mode approximation was used in determining the vibrational frequencies; for which the Hessian was calculated using finite difference implemented in VASP, with

displacement steps of ± 0.015 Å and a convergence criterion of $10^{-6}$eV. The translational and rotational modes for the bulk-phase species were decoupled from the vibrational modes and their thermodynamic contributions were calculated according to the formal statistical thermodynamic treatment of an ideal gas. For adsorbed states, all imaginary frequencies were replaced with a $100cm^{-1}$ cutoff, based on a sensitivity analysis from a previous work.[108] For transition states with more than a single imaginary frequency, the respective modes were visualized and replaced. The enthalpies were calculated by taking the sum of the DFT-calculated ground state electronic energy, the zero point vibrational energy, and temperature contributions from the constant pressure heat capacity, which was derive by fitting the entropy to a set of Shomate Parameters. Finally, the standard Gibbs Free Energy was derived from it's classical definition of the enthalpy and entropy. (See equations 3 and 4 in A.2 of the Appendix.)

The thermodynamic quantities for adsorbed species were also calculated using this method, but were given in terms of an upper and lower bound based on approximations of the adsorbate entropy outlined in A.2 of the Appendix. The reality is that our thermodynamic adsorbate quantities fall somewhere between these two approximations[110, 111]; we therefore report both values as bounds for our thermodynamic and kinetic parameters. Within this paper, these approximations are defined as "Free translator" and “Harmonic Oscillator” for the upper and lower entropic bounds respectively. Adsorption/desorption steps were constructed by forming pseudo-transition states which have the entropy of their corresponding adsorbate under the Free Translator approximation, and the enthalpy of their ideal-gas state, this derivation and rationale is outlined within A.3 of the Appendix. The derivation of our MKM and relative analyses (conversion, selectivity, yield, apparent activation energy, apparent reaction order) can also be found in A.4 of the Appendix.

## 3.3 Results and Discussion

### *3.3.1 Reactant adsorption*

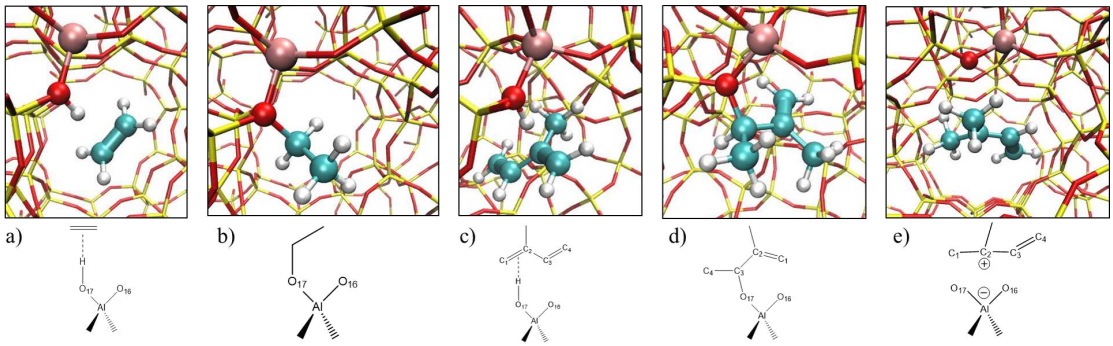


Figure 3.3: The most energetically stable adsorption configurations of our reactants. The Brønsted proton is bond to O17 for each image; and the aluminum, proton, and oxygen (O17) were emphasized using spheres. The reactant atoms were also represented as spheres, with double bonds emphasized using thicker diameters. Key: silicon (yellow), oxygen (red), hydrogen (white), aluminum (pink), carbon (turquoise) a) Ethene physisorption, where there is a distinct interaction between the Brønsted proton and ethene's double bond. b) Ethene chemisorption, where the primary carbon has been protonated; and an alkoxide bond has formed between O17 and the remaining carbon. c) *Trans*-isoprene physisorption, where the Brønsted proton at the acid site is interacting with the double bond between the primary and tertiary carbons. d) Isoprene chemisorption, where the primary carbon "$C_4$" has been protonated and an alkoxide bond has formed between the tertiary carbon "$C_3$"" and O17. e) Carbenium isoprene chemisorption, where the tertiary carbon "$C_3$" has been protonated.

We begin our analysis by validating the adsorption of our stable intermediates at standard conditions against available results. Table 3.1 shows our reactant DFT energy (non ZPE corrected), enthalpy, entropy, and Gibbs free energy change upon adsorption at standard conditions (298.15 K and 1 atm). Adsorption was defined by equation 3.1, where "$X$" represents the thermodynamic quantity of interest and the subscripts $molecule^*$, $molecule_{(g)}$, $Zeolite$, represent the adsorbed, gas and lone zeolite states respectively. Due to the propensity of olefins to undergo oligomerization reactions within acidic zeolites (see section 3.3.3), we have investigated three possible

| | Adsorbate | $\Delta E_{DFT}$ [kJ/mol] | $\Delta H^0$ [kJ/mol] | $\Delta S^0$ [J/mol/K] | $\Delta G^0$ [kJ/mol] |
|---|---|---|---|---|---|
| **physisorbed** | ethylene | -63 | -58(-59) | -143(-107) | -15(-27) |
| | *trans*-isoprene | -100 | -96(-98) | -191(-147) | -39(-54) |
| **alkoxide** | ethene | -108 | -96 | -171 | -45 |
| | isoprene | -90 | -80 | -208 | -18 |
| **carbenium** | isoprene | -70 | -67(-68) | -179(-136) | -13(-27) |

Table 3.1: Standard enthalpy ($kJ/mol$), entropy ($J/mol/K$), and Gibbs free energy ($kJ/mol$) of adsorption at 298.15 ($K$) and 1 ($atm$) with the Free Translator approximation quantities housed within parenthesis

scenarios for the adsorption of each reactant: 1) A physisorbed $\pi$-complex, where the Brønsted proton is interacting with the olefin but no distinct bond is formed; 2) Protonation of the olefin, resulting in the formation of a covalent alkoxide bond with an oxygen atom at the acid site; 3) Protonation of the olefin, resulting in a stable carbenium ion. Theoretically, protonation may occur on any of the carbon atoms involved in its $\pi$-bonds. Moreover, the Brønsted proton at the acid site has been shown to be mobile, capable of migrating among adjacent oxygen atoms.[112] As a result, the alkoxide bond may theoretically be formed with any of the adjacent framework oxygen atoms. We have examined each of these possibilities, but only the most favorable states are discussed here.

$$\Delta X_{adsorption} = X_{molecule^*} - X_{molecule_{(g)}} - X_{Zeolite} \tag{3.1}$$

## Ethene

Our relaxed complex for the physisorbed state of ethene is shown in Figure 3.3 a). Consistent with other works[108, 113], the Brønsted proton is attracted to the center of the $\pi$-bond of ethene. The formation of this physisorbed structure is exergonic, with a Gibbs free energy of -15(-27)kJ/mol. The most favorable chemisorbed structure of ethene is shown in Figure 3.3 b), with the alkoxide bond formed with oxygen O17

(also the most stable oxygen for the Brønsted site). The chemisorption of ethene was found to be more favorable by -18(-30) kJ/mol than physisorption, with a Gibbs Free Energy of -45 kJ/mol. No stable carbenium ion was found for ethene, which is consistent with other works which suggest that small olefins exist as carbeniums only through transition states during chemisorption.[114, 115]

Nguyen et al.[113] have computationally investigated the adsorption of C2-C8 linear alkenes in various acidic zeolites at the QM-Pot(MP2//B3LYP:GULP) level of theory. Specifically, they have calculated the standard enthalpy, entropy, and Gibbs free energy change for the physisorption, chemisorption and protonation of ethene within H-ZSM5. Their calculated Gibbs free energy of ethene physisorption was -20 kJ/mol, which falls within our range of -15(-27) kJ/mol. Their entropy and enthalpy are also consistent with our values. On the other hand, their Gibbs free energy of chemisorption was -69 kJ/mol, approximately 24 kJ/mol less than our value. A direct comparison between our chemisorption energies (non ZPE corrected) shows that DFT-D2 dispersion corrections predict a higher binding energy by 22 kJ/mol.

In a separate study[108], we assessed the interactions between the $\pi$-bond of ethene with the active sites of Brønsted and cation exchanged ZSM-5 zeolites by measuring its heat of adsorption using microcalorimetry and periodic DFT calculations at low temperature. As a comparison, the adsorption of ethane was also quantified, given its lack of a $\pi$-bond but otherwise similar physical properties to ethene. While both adsorbates showed identical experimental heats of adsorption within siliceous ZSM-5, ethene showed a significantly stronger adsorption within H-ZSM5 (8 kJ/mol) than the adsorption of ethane (2 kJ/mol). These results suggest that while ethene and ethane share similar confinement effects, ethene exhibits additional Brønsted-$\pi$ interactions at the active sites within H-ZSM5. The PBE(DFT-D2) functional correctly captured this enhanced stability of ethene, but it overestimated its experimental

heat of adsorption by 20 kJ/mol. A comparison among different DFT functionals (Grimme-D2[106, 107], Grimme-D3[107], Tkatchenko-Scheffler[116], DDsC[117, 118], vdw-Df(2)[119, 120, 121]) for this system showed a wide variation in the binding energies of ethene/ethane (almost 32 kJ/mol), but all correctly captured that ethene interacts more strongly than ethane at the Brønsted acid site. Accordingly, while our DFT calculations may overpredict experimental values, we expect our model to correctly capture qualitative differences.

### Isoprene

Our relaxed physisorbed complex of isoprene is shown in Figure 3.3 c). The Brønsted proton resides on O17; and is directed toward the center of the $\pi$-bond between the primary and tertiary carbons (labelled as "$C_1$" and "$C_2$" within Figure 3.3 c) respectively). This configuration is consistent with other physisorbed olefins of similar size within acidic zeolites.[122, 113, 123, 124] The adsorption was exergonic, with a Gibbs free energy of -39(-54)kJ/mol.

The most stable chemisorbed state of isoprene is shown in Figure 3.3 d), where the primary carbon "$C_4$" has been protonated; and an alkoxide bond formed between the secondary "$C_3$" carbon with the host oxygen O17. The Gibbs free energy of this chemisorbed state was -18 kJ/mol. Unlike ethene, we have found a stable tertiary-carbenium ion; and its relaxed structure is shown in Figure 3.3 d). Adsorption of this tertiary-carbenium was also found to be exergonic, with a Gibbs free Energy of -13(-27) kJ/mol. The resulting alkoxides from this tertiary-carbenium were found to be very energetically unfavorable. (See Figure 3.9 of Section 3.3.3) Ultimately, the physisorbed state of isoprene was found to be more thermodynamically favorable than the chemisorbed and carbenium states by 12(40) kJ/mol and 26(25) kJ/mol respectively.

Nguyen et al.[123] have looked into the physisorption, chemisorption, and protonation of isobutene in various acidic zeolites using periodic density functional theory at the PW91-D//PW91 level of theory. On the basis of their calculations, the standard Gibbs free energy of formation for physisorbed isobutene within H-ZSM-5 is slightly more stable than it's tertiary-butyl carbenium ion, but significantly more stable than its chemisorbed counterparts.

Moreover, Cnudde et al.[125] have computationally studied the temperature dependent stability of C4-C5 alkene cracking intermediates within H-ZSM-5 using ab-initio and molecular dynamics simulations. The results of their static simulations (non MD) have shown that physisorption is more stable than chemisorption, regardless of the alkene (branched vs. linear) or temperature (323-773 K).

Their dynamic simulations (MD) at 323 K showed systemic deprotonation for linear carbenium species, whilst the physisorbed and alkoxide species remained stable. However the MD simulations of branched intermediates have shown that the physisorbed and chemisorbed species experienced prompt protonation into stable carbenium ions.

These studies are somewhat consistent with our results, because we have found that the physisorbed complex of isoprene is more stable than its chemisorbed and carbenium counterparts. However, our tertiary-carbenium state is only "significantly" more stable than chemisorbed isoprene if we assume our entropy approximation is more accurately described by the Free Translator. While MD simulations to analyze the stability of isoprene adsorption are outside the scope of this paper, we postulate that the range of our thermodynamic approximations encompasses the stability of these higher free energy states.

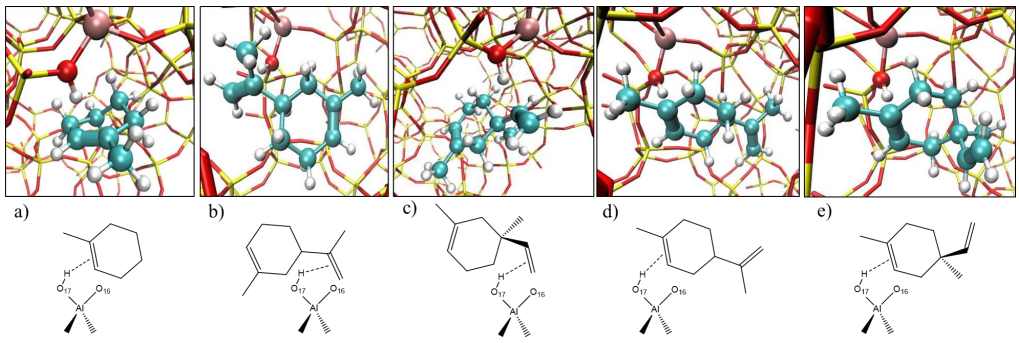


Figure 3.4: The most energetically stable adsorption configurations of our Diels-Alder products. The Brønsted proton is bond to O17 for each image; and the aluminum, proton, and oxygen (O17) were emphasized using spheres. The reactant atoms were also represented as spheres, with double bonds emphasized using thicker diameters. Key: silicon (yellow), oxygen (red), hydrogen (white), aluminum (pink), carbon (turquoise) a) Product C7 physisorption, where the Brønsted proton interacts with C7's double bond. b) Product C10-meta1 physisorption, where the Brønsted proton interacts with C10-meta1's external double bond. c) Product C10-meta2 physisorption, where the Brønsted proton interacts with C10-meta2's external double bond. d) Product C10-para1 physisorption, where the Brønsted proton interacts with C10-para1's internal double bond. e) Product C10-para2 physisorption, where the Brønsted proton interacts with C10-para2's internal double bond.

### *3.3.2 Diels-Alder product adsorption*

The most stable configurations of our adsorbed DA products are shown in Figure 3.4; and their respective DFT energy (non ZPE corrected), enthalpy, entropy, and Gibbs Free energy change upon adsorption (defined by equation 3.1) at 298.15 K and 1 atm are given in table 3.2. Synonymous with the configurations of the physisorbed reactants, each product exhibited an interaction between the Brønsted proton on O17 with one of its $\pi$-bonds. For the C7 product this interaction occurred with its lone $\pi$-bond; but for the C10 products, it was observed with either the $\pi$-bond of the dienophile (external to the six-membered ring) or with the $\pi$-bond of the diene (internal to the six-membered ring). The C10-meta products preferred this interaction with their external $\pi$-bond, while the C10-para products preferred to interact with their internal $\pi$-bonds. The Gibbs free energy of adsorption for the C10-para products

| **adsorbate** | $\Delta E_{DFT}$ [kJ/mol] | $\Delta H^0$ [kJ/mol] | $\Delta S^0$ [J/mol/K] | $\Delta G^0$ [kJ/mol] |
|---|---|---|---|---|
| $C7$ | -127 | -121(-123) | -203(-157) | -61(-76) |
| $C10-para1$ | -146 | -138(-140) | -224(-175) | -71(-88) |
| $C10-para2$ | -155 | -150(-151) | -218(-169) | -85(-101) |
| $C10-meta1$ | -141 | -135(-136) | -226(-177) | -67(-83) |
| $C10-meta2$ | -145 | -138(-140) | -218(-169) | -73(-89) |

Table 3.2: Standard enthalpy ($kJ/mol$), entropy ($J/mol/K$), and Gibbs free energy ($kJ/mol$) of adsortpion at 298.15 ($K$). The lower and upper bounds are given by the harmonic and Free Translator approximations defined in section A.2 of the Appendix, with the Free Translator quantities housed within parenthesis

were on average 8 ($kJ/mol$) more favorable than the C10-meta products, with the C10-para2 product being most strongly adsorbed by an average of 14 ($kJ/mol$).

### *3.3.3 Reaction pathways*

The adsorption of small olefins at the Brønsted acid sites of zeolites have a propensity to oligomerize at temperatures as low as 300 K.[126, 127, 128, 129, 130, 131] At higher temperatures, these oligomers have been shown to further undergo cyclization reactions, which can subsequently dehydrogenate to form a myriad of cyclic products.[132, 133] In particular, our C7 product may be formed via the 1,6 ring closure/dehydrogenation of a primary alkoxide, formed by the oligomerization between isoprene and ethene (see Figure 3.5). The extent that these reactions may compete with DA cycloaddition is otherwise unknown; but Bernardon et al.[19] have run similar DA reactions (isoprene with methyl-acrylate) within H-ZSM5 at temperatures between 293.15 to 363.15 K without appreciable formation of such byproducts. Therefore to limit these unwanted reactions while maintaining sufficient catalytic activity (see section 3.3.4), the conditions for our analysis were chosen to be 368.15 K and 1 atm.

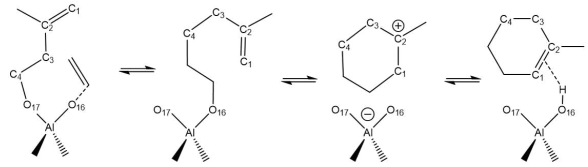

Figure 3.5: Proposed mechanism for the oligomerization of primary isoprene alkoxide with ethene, with subsequent cyclization and dehydrogenation to form product C7.

## Diels-Alder cycloaddition

Figure 3.6 displays the Gibbs free energy surface (Data given in Table A1 of the Appendix) of our adsorbed and gas-phase DA reactions at 368.15 K and 1 atm. The primary and most important observation is that the apparent barrier for each adsorbed DA cycloaddition reaction is substantially smaller than their gas-phase analogs, suggesting that this reaction is catalyzed within H-ZSM5. All DA reactions showed an average 37(56) kJ/mol lower apparent activation barrier within H-ZSM5 relative to their bulk gas phase. The apparent activation barrier for the C7 reaction was on average 6(26) kJ/mol smaller relative to the other DA reactions, but its bulk gas phase barrier was also 26 kJ/mol smaller among the bulk gas phase reactions. However, the C10 reactions showed a more significant change in their apparent activation barriers relative to their bulk gas phase reactions, being on average 46(66) kJ/mol and 22(35) kJ/mol lower, with C7 being only 8(8) kJ/mol lower. Among the C10 reactions, the meta regio-isomers both exhibited lower apparent activation barriers by 9 kJ/mol relative to the para regio-isomers. Because the entropic approximations are largely cancelled when comparing among these reactions, we attribute such stability differences to the DFT energies. Therefore, our data show: 1) The larger C10 TSTs are more stabilized by the framework than the C7 TST, and 2) The meta regio-isomers are more stabilized by the framework than the para regio-isomers, relative to their unctalyzed gas-phase reactions.

The enhanced stabilization of these TSTs relative to their gas phase phase is

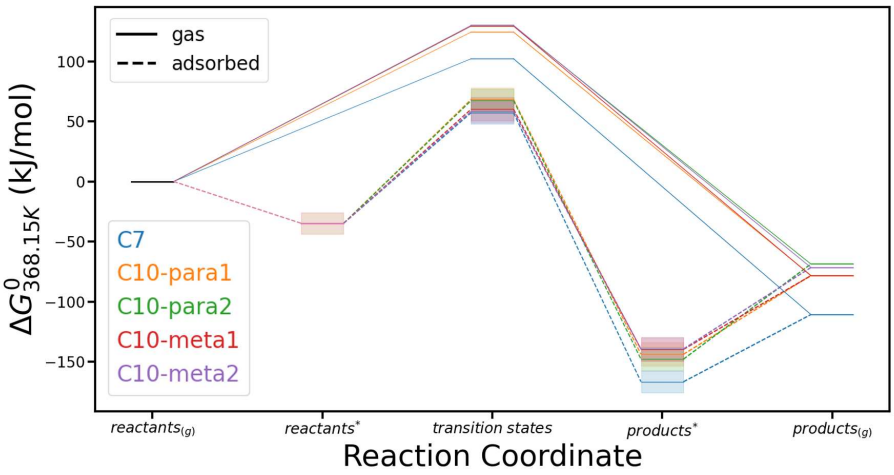


Figure 3.6: The Gibbs free energy surface of our adsorbed (dashed line) and gas (solid line) DA reactions at 368.15 (K). The x-axis represents the reaction coordinate, which is segregated between the adsorbed and gas states; represented by "*" and "(g)" super/subscripts respectively. All path energies have been normalized relative to their respective gas-phase reactants. The adsorbed reactants are non-interacting and their energies were taken as the sum of their most stable physisorbed configurations from Table 3.1 with their corresponding secondary reactant from the bulk gas phase. The product energies were taken from their most stable configurations outlined in Table 3.2. The shaded regions represent the thermodynamic limits defined by the harmonic and Free Translator approximations, with each state representing an average between those two limits.)

largely attributed to the favorable confinement induced by the H-ZSM5 framework rather than the Brønsted-$\pi$ interaction with the acid site. In Figure 3.8, we show the apparent (non ZPE corrected) DFT energy barriers for the catalyzed and gas phase DA cycloaddition reactions, previously outlined within Figure 3.6 and following the same color scheme. These energies were separated into their Kohn Sham (PBE) and Grimme-D2 contributions (defined in section 3.2) and plotted within Figure 3.8 using cross hatches and hatches respectively, while the uncatalyzed gas-phase and catalyzed adsorbed phase reaction energies were plotted with solid line and dash-dot borders respectively. The most significant observation from Figure 3.8 is that the differences in PBE energy barriers for the catalyzed and gas phase are much smaller than for the

Grimme-D2 dispersion correction energies. This suggests that the TSTs are mainly stabilized by the long-range van der Waals forces provided by the framework (i.e confinement), rather than the covalent interaction between the Brønsted proton and the $\pi$-bond of the TST.

The TST configurations corresponding to our DA reactions are shown in Figure 3.7 and in all cases it was found that the most energetically stable configurations preferred *cis*-isoprene (the diene) to reside at the acid site. These structures largely resemble their co-adsorbed states rather than their products, which was also observed in the gas-phase (see Figure A1 of the Appendix). Unlike the configurations of the adsorbed products shown in Figure 3.4, the Brønsted proton preferred to interact more closely with one of the carbon atoms involved in a $\pi$-bond rather than the $\pi$-bond itself. In particular, the Brønsted proton preferred to interact with the secondary carbon of the *cis*-isoprene (diene) for the C7 TST, represented within Figure 3.7 a). However, for the C10 TSTs, this interaction ranged between the primary and tertiary carbon atoms of the *cis*-isoprene (diene), ultimately depending on whether the final product was para or meta. The C10-para TSTs are represented within Figure 3.7 b) and c); and it can been seen that the Brønsted proton preferred to interact with the tertiary carbon. However, for the C10-meta TSTs (represented within Figure 3.7 d) and e)), the Brønsted proton preferred to interact with the primary carbon.

## Competitive chemisorption and cyclization pathways

Historically, two pathways for olefin oligomerization have been considered: 1) a concerted mechanism in which protonation and C-C coupling occur simultaneously, or 2) a step-wise mechanism, where protonation results in the formation of an intermediate alkoxide before coupling with an additional olefin. Svelle et al.[134] have investigated both pathways using DFT in the dimerization of linear alkenes; but the

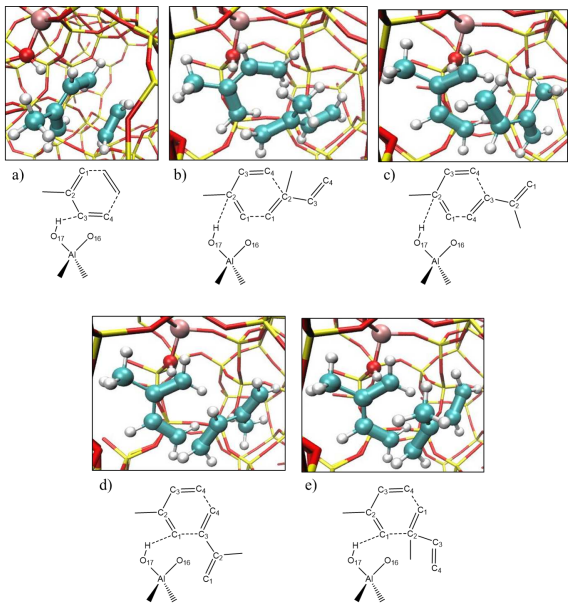


Figure 3.7: The most energetically stable adsorption configurations of our Diels-Alder transition states (TST). The Brønsted proton is bond to O17 for each image; and the aluminum, proton, and oxygen (O17) were emphasized using spheres. The reactant atoms were also represented as spheres, with double bonds emphasized using thicker diameters. Key: silicon (yellow), oxygen (red), hydrogen (white), aluminum (pink), carbon (turquoise) a) Product C7 TST, where the Brønsted proton interacts with the diene's secondary carbon "$C_2$". b) Product C10-para2 TST, where the Brønsted proton interacts with the tertiary carbon "$C_2$" on the diene. c) Product C10-para1 TST, where the Brønsted proton interacts with the tertiary carbon on the diene. d) Product C10-meta1 TST, where the Brønsted proton interacts with the diene's primary carbon "$C_1$". e) Product C10-meta2 TST, where the Brønsted proton interacts with the diene's primary carbon "$C_1$".

prevailing pathway was inconclusive. On the other hand, Shen W.[135] have concluded that while the concerted mechanism for ethene dimerization is preferred in large pore zeolites, both pathways are competitive for smaller pore zeolites (i.e H-ZSM5). Without claiming exclusivity, we have investigated the step-wise mechanism; which has been predominantly chosen in works concerning olefin cyclization and $\beta$-scission.[130, 132, 136]

Unlike oligomerization, the mechanism for olefin chemisorption is well established, and is illustrated within Figure 3.9.[122, 131, 113, 123, 130] The path begins with a

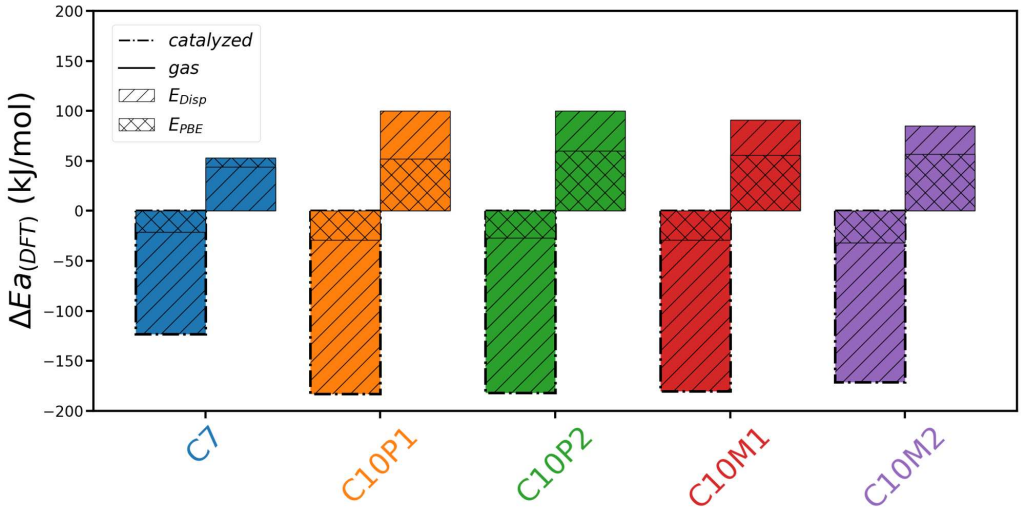


Figure 3.8: The apparent DFT energy barrier (non ZPE corrected) for the catalyzed (alternating line-dot border) and gas phase (straight-line border) DA cycloaddition reactions outlined within Figure 3.6. The apparent barrier was defined as the difference between the TST and the sum of zeolite structure with the gas phase reactants. The DFT energy has been separated between the Perdew-Wang-Ernzerhof (PBE) energy (cross-hatch filled) of the exchange correlation functional and the Grimme-D2 dispersion correction energy (line-hatch filled). The differences between the catalyzed and gas phase PBE energies are small for each reaction. However, the differences between the catalyzed and gas phase dispersion energies are large for each reaction.

physisorbed complex, where the Brønsted proton interacts with the $\pi$-bond of the olefin. Chemisorption is initiated by protonation of either $\pi$-bonded carbon in the olefin, ultimately traversing through a cationic TST before forming a covalent alkoxide bond with the adjacent oxygen atom in the framework. The rate of this step has been shown to depend upon the stabilization of this cationic TST, with tertiary carbons preferring to house the positive charge over secondary or primary carbons.[113, 123] However, protonation of olefins containing tertiary carbons (i.e isoprene) have been found to form stable carbenium intermediates in the form of physisorbed carbocationic states. For example, Ngueyn et al.[123] have shown tert-butyl carbeniums to be more stable than their alkoxide analogs at temperatures as low as 300 K.[123]

In Figure 3.10 we show the Gibbs free energy surface for the physisorption, pro-

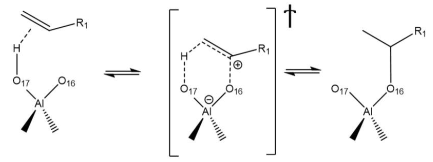


Figure 3.9: General mechanism for olefin chemisorption at a zeolite's Brønsted site. The formation of the alkoxide bond is formed with an adjacent framework oxygen.

tonation, and subsequent chemisorption of ethene and isoprene at 368.15 K and 1 atm. Although we have considered many possible adsorption configurations at these conditions, the most thermodynamically favourable states remained those discussed within section 3.3.1. In particular, the Brønsted proton preferred to reside at O17, being 10 ($kJ/mol$) more stable than at O16. Physisorbed ethene and *trans*-isoprene continued to be preferred at O17 than O16 by 16 and 19 ($kJ/mol$) respectively. In their chemisorbed states, both ethene and isoprene preferred to form their alkoxide bond with O17 rather than O16 by 39 and 19 ($kJ/mol$) respectively. Stable carbenium intermediates were only found for isoprene, with protonation occurring on the primary carbon such that either a secondary or tertiary carbenium ion was formed. The tertiary carbenium was found to be 15 ($kJ/mol$) more stable than the secondary carbenium, however the subsequent tertiary alkoxide was 15 ($kJ/mol$) less stable than the secondary alkoxide. Overall, the two most thermodynamically favorable adsorption configurations; and therefore most likely to occupy the acid site, were: 1) *trans*-isoprene physisorbed at O17, and 2) chemisorbed ethene at O17. The favourablity between these two states depended upon the thermodynamic approximation used. In particular, physisorbed *trans*-isoprene can be 10 ($kJ/mol$) more favorable (Free Translator), or 7 ($kJ/mol$) less favourable (harmonic oscillator) than chemisorbed ethene. Despite being thermodynamically similar at these conditions, the physisorption of *trans*-isoprene was found to be the kinetically favored adsorbate.

The minimum energy path for chemisorption occurs through the formation of a

covalent alkoxide bond with an adjacent oxygen atom at the acid site. Therefore, the path for ethene chemisorption at O17 requires: 1) an initial proton migration between O17 to O16, 2) physisorption of ethene at O16, and 3) protonation of ethene with chemisorption at O17. The Gibbs free energy surface for this mechanism is represented by the pink colored path in Figure 3.10. The barrier for the proton to migrate from O17 to O16 is represented by the path connecting states "0" to "1" in Figure 3.10; it had an electronic DFT (non ZPE corrected) barrier of 69 ($kJ/mol$) and a Gibbs free energy barrier of 58 ($kJ/mol$). Sierka and Sauer have also investigated proton mobility in ZSM-5 by performing QM-POT calculations at the CCSD(T) level of theory.[137] Likewise, they considered the T7 O17 Brønsted site as the most energetically stable (labeled as O7 within their paper); and found the electronic barrier for proton migration to be 80.7 ($kJ/mol$), close to our predictions. Upon physisorption of ethene at O16, the barrier for protonation is represented by the path connecting states "5" to "4" within Figure 3.10. For this step, both the intrinsic Gibbs free energy and electronic DFT barriers were 64 ($kJ/mol$). Early ab-initio calculations using general zeolite models have shown intrinsic barriers for ethene chemisorption to be between 68-129 ($kJ/mol$).[138, 139, 140] More recently, Shen W.[135] and Chu et al.[141] have seperately investigated ethene dimerization within H-ZSM5 using DFT methods on variously sized cluster models. Both models had the aluminum atom located at the T12 position and the Brønsted proton residing at O24. Shen W. found the intrinsic energy barrier for the chemisorption of ethene from the physisorbed state to be 70-98 ($kJ/mol$); while Chu et al. have found barriers between 68-80 ($kJ/mol$). Work by Gleeson D.[142] using DFT based cluster models calculated an instrinsic barrier between 78-119 ($kJ/mol$). Based on these results, our calculated intrinsic barrier of 64 ($kJ/mol$) was slightly underpredicted. Nevertheless, our underprediction reinforces that ethene chemisorption will be kinetically limited

relative to *trans*-isoprene physisorption.

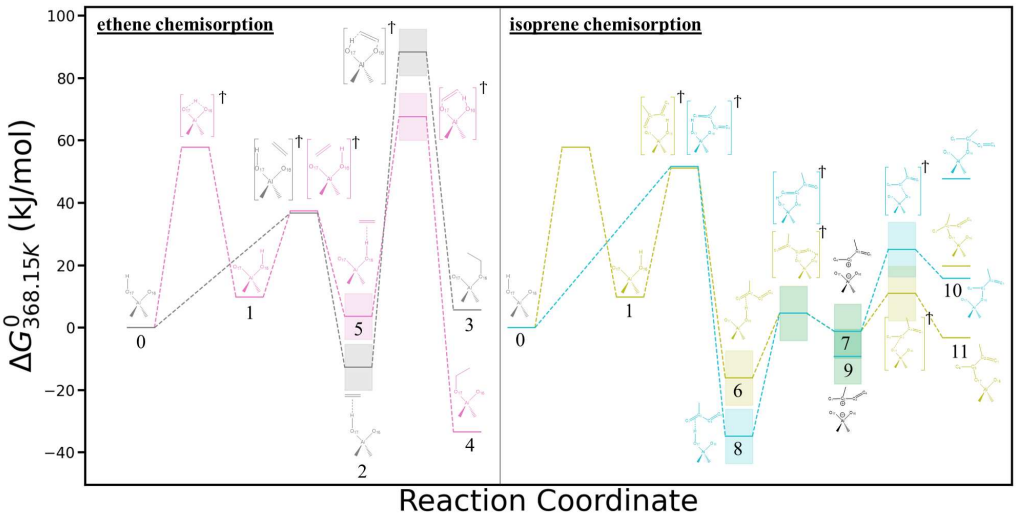


Figure 3.10: The Gibbs free energy surface at 368.15 K and 1 atm of the physisorption and subsequent chemisorption of ethene (Left) and isoprene (Right). The grey and turquoise path show olefin physisorption at O17 with chemisorption occurring at O16 for ethene and *trans*-isoprene respectively. The pink and olive paths show proton migration form O17 to O16 followed by physisorption of the olefin at O16 with subsequent chemisorption at O17 for ethene and *trans*-isoprene respectively. The numbering scheme for each state is as follows: 0) Brønsted site at O17, 1) Brønsted site at O16, 2) physisorbed ethene at O17, 3) chemisorbed ethene at O16, 4) chemisorbed ethene at O17, 5)physisorbed ethene at O16, 6) physisorbed *trans*-isoprene at O16, 7) secondary carbenium isoprene, 8) physisorbed *trans*-isoprene, 9) tertiary carbenium isoprene, 10) secondary alkoxide isoprene at O16, 11 secondary alkoxide isoprene at O17. Only the most stable physisorbed/chemisorbed states were shown, with the exception of tertiary isoprene alkoxides whose states are not connected, but are listed above states 10 and 11. The shaded regions represent the thermodynamic limits defined by the harmonic and Free Translator approximations, with the state representing an average between those two limits. The "†" superscript represents TSTs.

Given its small chemisorption barriers (relative to DA cycloaddition) and occupancy at the Brønsted site, we expect rapid equilibriation of *trans*-isoprene among its protonated states. As such, we reckon with a possible cyclization/dehydrogenation pathway for the formation of our C7 product. We based our mechanism on a low-energy route proposed for the cyclization between ethene and propene by Vandichel et al.;[132] which consisted of: 1) chain growth through oligomerization between a

chemisorbed olefin with a co-physisorbed olefin, 2) cyclization of the formed chain. The necessary chain to undergo cylization into our C7 product may only be generated through the oligomerization between a primary isoprene alkoxide with co-adsorbed ethene. This primary alkoxide is formed by the initial protonation of the secondary carbon of isoprene "$C_3$"; with the alkoxide bond formed between the primary "$C_4$" carbon of isoprene with the framework oxygen O16. We have found two transition states and a stable cationic cyclo-propyl intermediate along this path (see Figure 3.11). Although stable cyclopropane intermediates have been found in steps such as ethene methylation[127] and skeletal isomerization of alkenes[143, 144], their cationic forms have only been predicted as TSTs preceding ring closure or isomerization. However, such TSTs were not larger than di-methyl-cyclopropane and we rationalize the stability of our intermediate through the additional hydrogen bonding with the framework from its larger carbon number (C7 vs. C5).

The Gibbs free energy surface at our nominal conditions for the oligomerization/cyclization pathway to form our C7 product is shown in Figure 3.11. For comparison, we have also included the DA cycloaddition barrier, beginning with physisorbed *trans*-isoprene. It is evident, by Figure 3.11, that the DA cycloaddition barrier is not only kinetically favored; but is also thermodynamically preferred. In particular, the oligomeric intermediates along the cyclization path are almost equivalently less stable than physisorbed *trans*-isoprene, while the barriers for TST1 and TST2 with respect to physisorbed *trans*-isoprene are larger than DA cycloaddition. Based on these results we conclude that the DA cycloaddition is the dominant path toward C7 formation; and have disregarded the competitive oligomerization/cyclization pathway from our MKM.

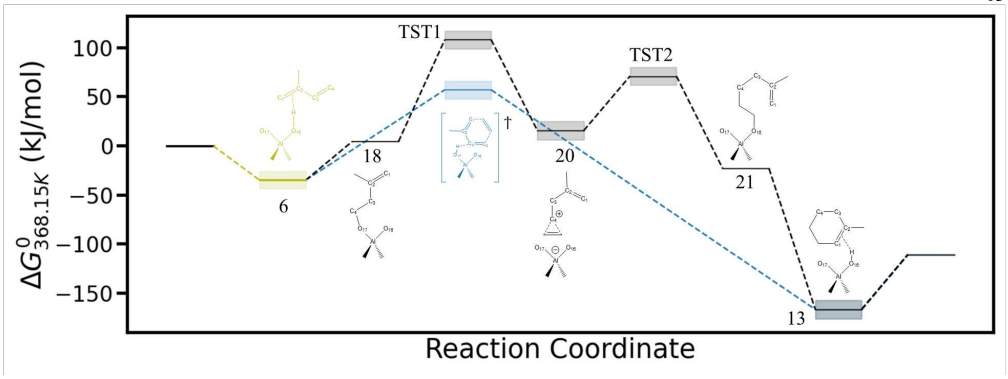


Figure 3.11: The Gibbs free energy surface at 368.15 K of isoprene oligomerizing with ethene to form the C7 product (Black) and the DA cyloaddition reaction outlined in Figure 3.6 (Blue). The shaded regions represent the thermodynamic limits defined by the harmonic and Free Translator approximations, with the state representing an average between those two limits.

### *3.3.4 Microkinetic modeling and sensitivity analysis*

We begin our analysis by first comparing our DFT based MKM against experimental results. Apart from furan chemistry, zeolite catalyzed DA reactions have been relatively unexplored. The most comparable experiment is that reported by Bernardon et al.[19], who investigated DA reactions between isoprene (diene) and methyl-acrylate (dienophile) within many acidic zeolites at 348.15 K and 1 atm. They have concluded that H-ZSM5 zeolites had the highest activity for DA reactions, achieving site time yields (STY) between 0.05-0.219 $(mmol - product/mmol - H + /hour)$ toward their most selective product (para-cycloadduct). Under the HO approximation with an identical temperature and reactant ratio (2:1 for dienophile:diene), our MKM predicted a STY of 0.01 $(mmol - C7/mmol - H + /hour)$, with C7 being most selective. Under the Free Translator approximation, the rates of product formation drop significantly; and the catalytic flux was an order of magnitude smaller than the gas-phase flux. The reason for this phenomenon is discussed later, but was ultimately due to the overprediction of the cycloadduct binding energies. To improve our catalytic performance while mitigating competitive oligomerization re-

actions, our preferred nominal condition included a higher temperature and larger dienophile:diene feed ratio. Our reaction conditions were chosen as 368.15 K at 1 atm with a reactant ratio of 4:1 (dienophile:diene). At these conditions we obtained a STY of 0.4 ($mmol - C7/mmol - H + /hour$), with the catalytic flux of isoprene consumption toward product formation being 3400% larger than the gas-phase flux; which accounted for less than 4% toward product formation. In a previous study[108], we concluded that the PBE(DFT-D2) functional overestimated the experimental heat of adsorption of ethene within H-ZSM5 by 20 kJ/mol at 195 K. To address how this discrepancy would impact our results, we ran our MKM under a destabilized reaction system where intermediates and TSTs which explicitly interact with the Brønsted acid site were destabilized by 20 [kJ/mol]. The resulting catalytic flux became 7000% larger than the gas-phase flux, owing largely to the increased desorption rates of our products.

We would like to briefly emphasize that Bernardon et al. achieved upwards of 90% selectivity toward their most favourable cycloadduct (the para- regioisomer), whereas under identical conditions and using an identical reactant diene (isoprene), we only obtained 61% toward our most favourable cycloadduct (C7), and our most favourable C10 product was the meta- regioisomer (C10-meta2). Such increased selectivity and preference toward the para- regioisomer by Bernardon et al. can be explained by the difference in functionality between our dienophiles. In particular, methyl-acrylate contains an electron withdrawing ester group, which is known to lower the HOMO/LUMO energy gap in normal electron demand DA reactions and increase the corresponding reaction rate. Additionally, such an explanation would support the larger STY observed by Bernardon et al.

The results of our MKM are shown in Table 3.3, which include reactant conversion, product selectivity/yield, and surface coverage. Two results were listed for each

quantity, corresponding to the HO and Free Translator approximations discussed in A.2 of the Appendix. Each case was run separately, keeping reactant conversion at approximately 1%. Only adsorbates with the most significant coverage were listed; but accounted for at least 99% of the surface species. The reaction network implemented within the MKM and its corresponding kinetic parameters are given within Figure A2 and Table A1 of the Appendix.

Under the HO approximation, the C7 product was most selective at 70% with respect to isoprene. Between the C10 products the meta isomers were favored, exhibiting at least 9% more selectivity relative to the para products. This preference is related to their intrinsic barriers being on average 7 ($kJ/mol$) smaller than for the para cycloadducts. Overall, the DA products dominated the surface, accounting for at least 93% of the coverage, with chemisorbed ethene and physisorbed *trans*-isoprene mostly accounting for the remainder. The most abundant surface intermediate (MASI) was the C10-para2 product, with a coverage of 60%. The C10-para2 product was also the most stable C10 cycloadduct, it had a DFT energy that was 11 ($kJ/mol$) more favourable than the other C10 products and a desorption barrier that was 11-27 ($kJ/mol$) larger than all other products.

Under the Free Translator approximation, there was a precipitous drop in reactant conversion and product yield among all cycloadducts. Furthermore, the selectivity toward C7 had substantially increased, while the selectivity for the C10 products had decreased. Surprisingly, the catalytic flux for the formation of all cycloadducts was reduced by two orders of magnitude relative to the HO approximation. For the C7 product in particular, the gas phase flux was an order of magnitude larger than its catalytic flux, implying that its selectivity is primarily the result of reactions occurring in the gas phase. Moreover, its smaller gas phase activation barrier (See A1 of the Appendix) relative to the formation of C10 products supports its large selectivity.

The coverage of ethene and isoprene was also reduced, accounting for less than 0.01% of the surface species. On the other hand, each cycloadduct experienced an increase in coverage, now accounting for over 99% of the surface species. Consequently, it was determined that the desorption of products was rate limiting, being directly caused by the enhanced stability accrued by the Free Translator approximation. As intended, the Free Translator approximation serves as an upper-bound for the entropic freedom of the adsorbates, but likely overestimates the stability of our products and underestimates the rate.

The molecular surface area for translational entropy used within this study (200 x 600 pm) was derived by Moor et al.[145] for C2-C8 n-alkanes. The surface area was calculated with their physisorbed configurations having been oriented along the straight channel; with the Brønsted proton interacting with one of the carbon atoms. Unlike n-alkanes, no orientation of our cyclic products along the straight channel can afford such freedom. Moreover, the presence of Brønsted-$\pi$ interactions for our products at the acid site (orientations shown in Figure 3.4) introduces enhanced stability relative to n-alkanes of an identical carbon number. We therefore presume that the actual entropy falls closer to the HO approximation; and have included a comparison among MKM results (See Table A1, Table A2, Table A3 in the Appendix) from the HO, Free Translator, and Free Translator/HO hybrid where only the product entropies were approximated by the HO. Ultimately, the results from the Free Translator/HO hyrbid were largely similar to the HO, and we, therefore, performed the remainder of our kinetic analysis considering only the HO approximation.

## Kinetic analysis and selectivity

Table 3.4 shows the apparent order (derived by equation A19 of the Appendix) and inferred rate expression for each DA cycloaddition reaction under the HO approxi-

| adsorbate | Conversion % | Yield % | Selectivity % | Fractional Coverage |
|---|---|---|---|---|
| ethene | 0.2(0.01) | - | - | $6.8 \times 10^{-2}(5.7 \times 10^{-7})$ |
| isoprene | 1.1(0.04) | - | - | $1.6 \times 10^{-3}(1.8 \times 10^{-6})$ |
| C7 | - | 0.8(0.03) | 69.9(97.7) | $2.0 \times 10^{-2}(6.2 \times 10^{-3})$ |
| C10-para1 | - | 0.01(<0.001) | 0.7(0.09) | $9.5 \times 10^{-3}(1.2 \times 10^{-2})$ |
| C10-para2 | - | 0.1(<0.001) | 10.3(0.79) | $6.0 \times 10^{-1}(6.6 \times 10^{-1})$ |
| C10-meta1 | - | 0.01(<0.001) | 1.0(0.08) | $4.1 \times 10^{-2}(5.1 \times 10^{-2})$ |
| C10-meta2 | - | 0.2(<0.001) | 18(1.4) | $2.6 \times 10^{-1}(2.8 \times 10^{-1})$ |

Table 3.3: The conversion, selectivity, and coverage per adsorbate quantified from the MKM. The isoprene conversion included both isomers (*cis*− and *trans*−). The coverage for ethene and isoprene included their physisorbed, chemisorbed, and carbenium states on O17 and O16, including both isomers of isoprene (*cis*− and *trans*−). Free Translator quantities are housed within parenthesis

mation. The formation of C7 was found to be first order in ethene, following from its preference to react with the physisorbed *trans*-isoprene rather than acting as a competitive adosorbate that would otherwise saturate the surface. Much to our surprise, the formation of C7 was found to be inhibited by isoprene, having a negative apparent order of approximately -0.5. As discussed in the next subsection, cycloadduct desorption was found to be rate limiting and ultimately the cause behind isoprene inhibiting C7 formation. The inferred rate expressions for C10 production were all essentially identical, being zero order with respective to ethene, and approximately half order with respect to isoprene. The C10 cycloadducts are formed by the reaction between two isoprene molecules, justifying the zeroth order in ethene. Moreover, physisorbed *trans*-isoprene was the preffered adsorbate at the acid site, therefore ethene adsorption had little to no effect as a competitive adsorbate. The half-order of isoprene in C10 production suggests that, much like in C7 production, the desorption of cycloadducts is also inhibitory.

Figure 3.12 shows: 1) the selectivity of C7 cycloadduct and 2) the ratio of product formation rates (C7/C10), as a function of the reactant feed ratio, as predicted by the model. For each plot, the catalytic flux was ensured to be at least an order of

magnitude larger than its respective gas phase flux. The selectivity toward C7 can be defined by the ratio of product formation to isoprene consumption ($r_{C7}/r_{C5}$), as defined by equation 3.2. At high ethene partial pressures, the consumption of isoprene by the formation of C10 products diminishes relative to C7 formation, driving the selectivity toward unity. This phenomena offers a reactor engineering solution for controlling the selectivity toward C7. The selectivity toward C10 is also defined by equation 3.2; and is justified by the plot of product formation rates (C7/C10) as a function of the reactant partial pressure ratio within Figure 3.12. In particular, the ratio of inferred rate expressions for C7 against C10 (see table 3.4), produces a linear function with slope $\frac{1}{2}(k_{C7}/k_{C10})$ and a y-intercept of 0. This is replicated within Figure 3.12, suggesting that the selectivity defined within equation 3.2 is valid for C10 formation.

| | $n_{C2}$ | $n_{C5}$ | Inferred rate expression |
|---|---|---|---|
| $r_{C7}$ | 0.93 | -0.54 | $\sim (k_{C7})p_{C2}p_{C5}^{-0.5}$ |
| $r_{C10P1}$ | -0.08 | 0.47 | $\sim (k_{C10P1})p_{C5}^{-0.5}$ |
| $r_{C10P2}$ | -0.08 | 0.47 | $\sim (k_{C10P2})p_{C5}^{-0.5}$ |
| $r_{C10M1}$ | -0.08 | 0.47 | $\sim (k_{C10M1})p_{C5}^{-0.5}$ |
| $r_{C10M2}$ | -0.08 | 0.47 | $\sim (k_{C10M2})p_{C5}^{-0.5}$ |

Table 3.4: The apparent order ($n_i$) under the HO approximation for each DA cycloaddition reaction with respect to ethene (C2) and isoprene (C5), as defined by equation A19 in the Appendix

$$S(C7/C5) = \frac{k_{C7}\frac{p_{C2}}{\sqrt{p_{C5}}}}{k_{C7}\frac{p_{C2}}{\sqrt{p_{C5}}} + 2k_{C10}\sqrt{p_{C5}}}$$

$$S(C10/C5) = \frac{2k_{C10}\sqrt{p_{C5}}}{k_{C7}\frac{p_{C2}}{\sqrt{p_{C5}}} + 2k_{C10}\sqrt{p_{C5}}} \tag{3.2}$$

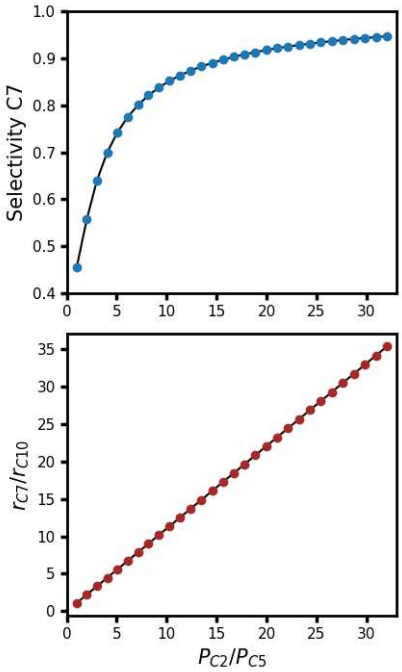


Figure 3.12: Plots of: 1) the Selectivity, predicted by the MKM, of C7, and 2) the ratio of C7 to C10 reaction rates, as a function of reactant ratio: (ethene/isoprene), otherwise (dienophile/diene).

## Global sensitivity analysis

The degree of rate control (DORC) defined by Campbell[146, 147] offers a robust approach toward finding the kinetically controlling steps in a reaction network. The notion behind this technique is to quantify the response of the overall reaction to a perturbation in the free energy of the TST or reaction intermediate and therefore the intrinsic barrier of that particular step. The outcome is a collection of quantities defining the enhancement/inhibition that each step within the network has on the overall rate. A limitation of this technique however, is that it only captures local sensitivities; since there are intrinsic errors in applying the DFT-D2 functional and uncertainty in calculating the entropy, we prefer a global estimate of sensitivities.

A variance-based global sensitivity analysis for catalytic reactions has recently been developed by Tian and Rangarajan[148], and we have applied their approach

to our system. Although the formulation of this method is outside the scope of this paper, we summarize the resulting quantities and how they can be used to infer kinetic relationships. In particular, the individual influence of an elementary step or reaction intermediate "$i$" on the overall rate is defined by "$S_i$", while the combined influence between "$i$" with elementary step or reaction intermediate "$j$" is defined by "$S_{ij}$". The "$S_{ij}$" variable encapsulates the nonlinearity of the MKM, as well as the extent to which the instantaneous change in the overall reaction rate with respect to parameter "$i$" (or "$j$") is dependent on the value of the other. In theory, the combined influence of every step and reaction intermediate within the network may be calculated ($S_{ijk...N}$), however this is typically not necessary and would require significant computational resources. Upon observing that the DA products dominate the surface coverage, we have limited our analysis to: 1) our catalyzed DA cycloaddition steps and the adsorption/desorption of their respective cycloadducts; 2) their adsorbed reactant/product intermediates (physisorbed ethene, isoprene, and each cycloadduct). The rate of isoprene consumption (limiting reactant) was defined as the overall rate, and the free energies of the TSTs and intermediates were perturbed to quantify the sensitivities. The perturbation range specified by the user has a significant impact on the resulting sensitivities, with small deviations from the nominal point encompassing local information (similar to DORC), while larger perturbations encompass global information, such as the correlation among parameters. It should be noted that while the DORC can be negative for some catalytic systems, the global sensitivities are always non-negative by definition.

Figure 3.13 shows the individual "$S_i$" and total "$S_i^T$" global sensitivities of the catalyzed DA cycloaddition reactions, their corresponding product adsorption steps, and their corresponding reaction intermediates under perturbation ranges from [-0.05, 0.05], [-5.0, 5.0], and [-7.5, 7.5] ($kJ/mol$) at 368.15 (K), 1 (atm) using the HO approx-

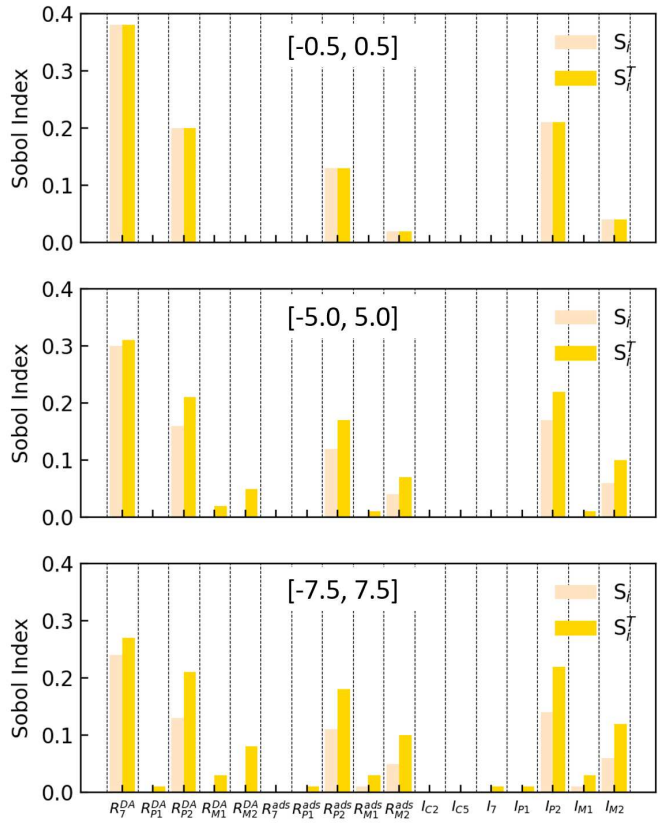


Figure 3.13: Global sensitivity values $S_i$ and $S_i^T$ for each catalyzed reaction step "$R$" and adsorbed intermediate "$I$", where the "DA" and "ads" superscripts correspond to the DA cycloaddition and corresponding product cycloadduct adsorption steps: C7 ($R_7$, $I_7$), C10-para1 ($R_{P1}$,$I_{P1}$) , C10-para2 ($R_{P2}$,$I_{P2}$), C10-meta1 ($R_{M1}$,$I_{M1}$), C10-meta2 ($R_{M2}$,$I_{M2}$), and adsorbed reactants ethene ($I_{C2}$), isoprene ($I_{C5}$) at 368.15 (K) and 1 (atm) under the HO approximation for perturbation ranges ([-0.5,0.5] to [-7.5,7.5] kJ/mol).

imation. The "$S_i^T$" variable captures the net influence of parameter "$i$", including it's contribution toward the other parameters; and therefore must always be larger than the individual contribution "$S_i$". Under small perturbations about the nominal point ([-0.5,0.5] ($kJ/mol$)), the sensitivities are qualitatively identical to the DORC. The overall reaction was most sensitive to the C7 DA cycloaddition step ($R_{C7}^{DA}$); this result is consistent with C7 being the most selective product and exhibiting the lowest apparent reaction barrier. Therefore, stabilizing the C7 TST would most influence the overall rate of this network. Following C7, the reaction rate was most sensitive to the

formation and desorption of C10-para2 ($R_{P2}^{DA}$, $R_{P2}^{ads}$, $I_{P2}$), which also exhibited the most exergonic adsorption step among the cycloadducts, and largest surface coverage (60%) among all adsorbed species. Among these parameters, the overall rate was most sensitive to the stability of the C10-para2 intermediate ($I_{P2}$). This result coincides with the precipitous drop in catalytic flux and upsurge in C10-para2 coverage (97.7%) under the Free Translator approximation, implicating that destabilizing the adsorbed C10-para2 product would increase the catalytic rate. The remaining sensitivities were marginal; and included the rate of adsorption of C10-meta2 ($R_{M2}^{ads}$) and it's adsorbed intermediate ($I_{M2}$). Similar to C10-para2, the rate was more sensitive to adsorbed C10-meta2, which was also the second most abundant surface intermediate (26%). Therefore in addition to C10-para2, the destabilization of the C10-meta2 adsorbate would increase the catalytic rate. As the range of allowed deviation is increased ([-5.0,5.0], [-7.5,7.5] ($kJ/mol$)), more parameters become important and the individual sensitivities of the C7, C10-para2, and C10-meta2 DA cycloaddition and adsorption steps/intermediates decrease, but the general rank of sensitivities remains unchanged. Moreover, we begin seeing that $S_i < S_i^T$, indicating that a combined contribution of multiple reactions steps and intermediate species affects the overall rate.

Figure 3.14 shows the heat map of $S_{ij}$, which can be thought of as a 2D correlation plot, where $S_{ij}$ quantifies the extent to which the combination of parameters "$i$" and "$j$" influence the overall rate. The average magnitude of $S_{ij}$ values progressively increase as the range is expanded from [-0.5, 0.5] to [-7.5, 7.5] ($kJ/mol$). This is consistent with $S_i$ decreasing relative to $S_i^T$ as the perturbation range is increased within Figure 3.13. The strongest correlations among the parameters are the DA cycloaddition step, adsorption step, and adsorbed intermediate of C10-para2 and of C10-meta2. Each of these parameters are kinetically related and consistent with the

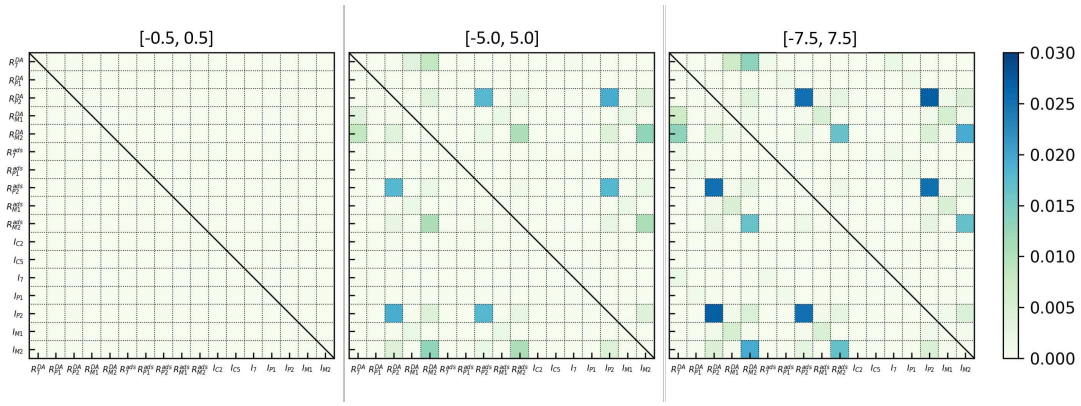


Figure 3.14: Heat maps of $S_{ij}$ values for each of the catalyzed DA reactions and corresponding adsorbed intermediates defined in Figure 3.13 at 368.15 (K) and 1 (atm) for perturbation ranges ([-0.5,0.5] to [-7.5,7.5] kJ/mol). Note, the map is symmetric, such that $S_{ij} = S_{ji}$ and $S_{ii}$ is not defined. As the parameter range is broadened, more $S_{ij}$ values become significant.

influence of C10-para2 and C10-meta2 on the overall rate. But apart from reinforcing our observations from Figure 3.13, the heatmap of Figure 3.14 shows that the formation of C7, formally the most influential independent parameter toward the reaction rate, is relatively uncorrelated with any species. This is also illustrated in Figure 3.13, where the difference between "$S_i^T$" and "$S_i$" are much smaller for C7 than for C10-para2 or C10-meta2. A possible explanation would be that C7 had the smallest apparent free energy barrier, such that it's formation and desorption rates were at least an order of magnitude larger than any other cycloadduct (See Table A3 of the Appendix). Therefore perturbations to C7 significantly impact the overall reaction (rate of isoprene consumption) by increasing/decreasing the rate of C7 formation, but do not significantly impact other steps/intermediates. Additionally, the formation of C7 was found to faintly correlate with the formation of C10-meta2, which exhibited the smallest apparent activation barrier and largest formation/desorption rate among the C10 products.

## 3.4 Conclusions

We performed a mechanistic study on the DA cycloaddition reactions between ethene and isoprene within H-ZSM5 using dispersion corrected DFT calculations. Temperature corrections were implemented using the HO and Free Translator approximations, offering a lower and upped bound on our MKM results respectively. The product cycloadducts produced a range of cyclic regio-isomers, which were the basis for studying how zeolite confinement would impact their stability and selectivity based on size and shape. The products included a smaller C7 cycloadduct produced from the coupling between ethene and isoprene, as well as four larger C10 cycloadducts produced by the coupling between two isoprene molecules. These corresponding C10 products were categorized based on their para- and meta- regiochemistry.

Our DFT results have shown adsorbate configurations and adsorption values consistent with available works. It was found that the DA cycloaddition reactions within H-ZSM5 were catalyzed with respect to their gas phase reactions; but with C10-para products having on average 7 ($kJ/mol$) larger intrinsic barriers than the C10-meta products. The driver for catalysis was determined to be the stability of the DA TSTs caused by their favourable van der Waals interactions with the framework (i.e confinement).

Oligomerization and cyclization mechanisms were also considered as potentially competitive pathways. In particular, we have investigated the chemisorption of our olefins via the formation of covalent alkoxide bonds with the oxygen atoms of the framework. Chemisorbed ethene was found to be thermodynamically competitive with physisorbed *trans*-isoprene; but it's intrinsic barrier rendered it kinetically unfavourable. A subsequent cyclization pathway for C7 formation was also investigated, beginning with olgiomerization ocurring between a primary isoprene alkoxide with coadsorbed ethene and concluding with dehydrogenation of a six-membered cyclic prod-

uct. However, this pathway was found to be both kinetically and thermodynamically unfavourable relative to DA cycloaddition.

Our DFT results were fed into a differential CSTR based MKM, with conditions replicated from an experimental work studying the DA reactions between isoprene and methyl-acrylate.[19] Our model was able to capture similar site time yields to experiments; but to enhance our catalytic flux we settled upon a larger reactant feed ratio and temperature. Under the Free Translator approximation, the gas phase flux for the formation of C7 was found to be an order of magnitude larger than it's respective catalytic flux. Under this approximation, the desorption of cycloadducts was found to be rate limiting, with the Free Translator approximation presumably under-estimating the entropic loss of adsorbed intermediates. Consequentially, combining the Free Translator approximation with the HO approximation for adsorbed cycloadducts produced largely similar MKM results to the HO approximation. Further analysis using the HO approximation has shown appreciable selectivity toward the C7 product (70%); and this approximation was used for the remainder of our analysis.

The apparent order for the formation of our C7 product was found to be first order in ethene but negative half order with respect to isoprene, suggesting that isoprene inhibits C7 formation. The apparent orders for formation of our C10 products were largely similar, all being zero order with respect to ethene; but half order with respect to isoprene. A global sensitivity analysis was implemented on our MKM, which showed that isoprene consumption (limiting reactant) was most sensitive to 1) the formation of C7, 2) the DA cycloaddition and adsorption steps of C10-para2 and C10-meta2 and their corresponding adsorbed state (the two most abundant surface intermediates). Surprisingly, the formation of C7 was largely uncorrelated with any other parameters, despite contributing the greatest individual influence on the

overall rate. It was posited that because C7 had the smallest apparent free energy barrier, with formation and desorption rates at least an order of magnitude larger than the other cycloadducts, that perturbations to the C7 TST resulted in significant consumption of isoprene but otherwise insignificant contributions to the other steps.

# CHAPTER 4

# ELUCIDATING ENTROPY DRIVEN STRUCTURE-PROPERTY RELATIONS USING NEURAL NETWORKS

## 4.1 Motivation

The adsorption entropy of molecules within microporous frameworks is a crucial quantity in our understanding of heterogeneous catalysis, hydrocarbon separations, and molecular storage. Traditionally, the entropy of an adosorbate has been illustrated through the notion of confinement, dictated by how well the molecule fits within the pores and active site of the framework.[2] The change in adsorption entropy "$\Delta s_{ads}$" can therefore be described as the loss in the degrees of freedom by the molecule upon a change of state from an unconstrained gas, to a confined adosorbate. Without the explicit formation of bonds with the framework structure, adsorbed molecules may retain the ability to partially translate, rotate, and vibrate within the framework. Naturally, such limitations are imposed by the size and shape of the pore. Molecules which retain translational motion will be constrained to the occupiable volume of the framework, while molecules whose length surpasses the diameter of the pore will be prevented from rotating about certain axis. Significant size/shape discrepancies between the framework topology and molecule can prevent adsorption altogether, while subtle differences may be exploited for a particular application, such as in the separation of alkanes,[75] and cracking of hydrocarbons.[149]

The adsorption entropy is a theoretical concept that can be experimentally measured by techniques such as gravimetry,[36] chromatographic adsorption,[34] and infrared operando spectroscopy.[37] Nevertheless, these techniques quickly become impractical for studying the many possible sorbate-framework combinations. Sim-

ulations offer an expedient route, being capable of predicting the entropy of adsorbed intermediates to within 10 J/mol-K error using a number of sophisticated approaches. These include computing anharmonic corrections to the quantum vibrational partition function,[38, 39] ab initio molecular dynamics simulations to calculate the adsorption free energy and thereby the entropy change upon adsorption,[40, 41] and configurational-bias Monte-Carlo simulations to generate entire adsorption isotherms.[42, 150] However, the use of such methods is computationally prohibitive; consequently, approximations are used as a compromise. For instance, knowledge of the physical behavior of the adsorbate may merit treating any mobility as purely vibrational, or invoking the hindered translator/rotor model.[44] However, such approximations are limited to the adsorbate-framework system. For example, the addition of two-dimensional free translations depend on a coverage term, which is particular to the dimensions of the active site of the framework and has been validated only for a subset of hydrocarbons.[151, 152, 110, 145]

The use of empirical correlations offers an expedient route in predicting sensible adsorption entropies without resorting to experiments or simulations. De Moor et al. [33] used ab initio simulations to show that the adsorption entropies for *n*-alkanes within acidic zeolites are linearly correlated with their carbon number. Campbell and Sellers compiled experimental alkane entropies on two-dimensional catalytic surfaces [5, 6]. The key finding of their work was that the ratio of the adsorbed-phase entropy to the gas-phase entropy was approximately two-thirds. Otyepková et al. calculated the adsorption entropy of a chemically diverse set of molecules adsorbed onto organic "van der Waals" materials using inverse gas chromatography and ab initio simulations [46, 47]. Their results showed a constant entropic loss of approximately 40 % relative to the gas-phase entropy. Likewise, Budi et al. calculated the adsorption entropy of a set of chemically diverse molecules adsorbed on mineral surfaces using

density functional theory [48]; and their data showed a strong linear dependence between the adsorbed-phase entropy and the gas-phase entropy.

The relationship between confinement and the adsorption entropy suggests that geometric descriptors may be suitable in developing empirical correlations. Denayer et al.[75] have demonstrated that the difference in rotational entropies was the driving force behind the zeolite separation of alkane/isoalkane mixtures, which correlated with the ratio the molecule's radius of gyration to the framework cavity's largest inscribed van der Waals radius. Dauenhauer and Abdelrhaman [7] compiled experimentally determined adsorption entropies for alkanes adsorbed in nine aluminosilicate zeolites. They showed that the entropic loss upon adsorption can be linearly correlated with the molecule's gas-phase translational and rotational entropies; and that the occupiable volume of a zeolite is a "universal" descriptor in predicting such losses. In a separate study[8], Rzepa et al. have qualitatively reproduced the trends found in Dauenhauer and Abdelrhamen's work for a larger set of molecules within pure-silica zeolites using Monte-Carlo simulations of the adsortion entropy at infinitely dilute conditions. The key findings of their work were: 1) the entropy ratio is remarkably uniform across the adsorbates for each framework, and 2) regardless of the metric used to characterize the zeolite adsorbent, the slope of the entropy relationship follows the same qualitative trend. The implication of their findings is that the entropic loss is effectively constant across adsorbates within materials with similar characteristic pore size and shape. This suggests that (except for a few prominent outliers) the adsorbate characteristics are relatively unimportant, and raises the possibility that models based primarily on adsorbent characteristics may be an effective tool for modeling the adsorbed-phase entropy.

Structure-entropy relationships within microporous frameworks offer rapid, yet sensible predictions of adsorption entropies, but constitute a challenge with a need

for discovering effective descriptors. Here, we build upon our previous work[8] by expanding our data set to 3965 molecule-zeolite systems, which includes 92 molecules and 89 framework combinations. We show that although the linear correlations between the adsortpion and gas phase entropies are largely preserved, their relationship with the framework's size metrics are much more complicated. To capture this non-linearity, we train a neural network model on a set of rapidly accessible/quantifiable geometric descriptors. We subsequently perform a local sensitivity analysis to identify the most prominent descriptors toward predicting the adsorption entropy, and rationalize their relevance through the lens of confinement.

## 4.2 Methods

### *4.2.1 Adsorption Entropy*

The following section describes the computational techniques that were used to calculate the adsorption entropy, $\Delta s^{\infty}_{ads}$, as derived in our previous work.[8] Our technique requires the selection of models for the zeolite adsorbents and adsorbate species and adaptation of a Monte Carlo integration technique (cf. Widom test particle method [64, 65]) to our adsorption system.

The adsorption system for our calculations is composed of a zeolitic adsorbent material and an adsorbate species. The adsorbents used here are in their pure-silicate forms (i.e., no aluminum substitutions) and listed in table A4 of the Appendix. The optimized structure of each silicate was obtained from Calero et al.[74], and then periodically replicated to form an adsorbent cell of adequate size (cf. discussion of the cutoff radii below). The adsorbent framework was kept entirely rigid. The adsorbate species used within this study are listed within table A5 and were selected on the basis of having an implicit hydrogen-based TraPPE force field.[69] The geometry of each

adsorbate was obtained by obtaining the lowest energy three-dimensional conformer from PubChem[153]. These molecular structures were used for the subsequent test particle insertions in the unaltered form, which allowed us to ignore energy terms related to bond bending, stretching, or rotation. The respective gas-phase entropies were calculated at the B3LYP level of theory with a cc-pVTZ basis set using Gaussian 09 software.[154] The zeolite frameworks were blocked according to the inaccessible voids characterized by Calero et al.[74] using methane as the probe molecule within frameworks of reduced oxygen radii (1.37 Å). These blockages were chosen based on our smallest adsorbate, methane, which is represented by a single sphere with Lennard-Jones $\sigma$ of 3.7 Åby the TraPPE forcefield. Additionally, reduced radii were implemented as an approximation of the framework's flexibility, which allowed for a more robust assessment of inaccessible sites.

As mentioned previously, the adsorbates were modeled using the TraPPE force field [49, 69] (i.e., Lennard-Jones parameters and point charges for constituent pseudo-atoms), while the adsorbents were modeled using the TraPPE-Zeo [70] force field. As specified in both TraPPE and TraPPE-Zeo, Lorentz–Berthelot combining rules were used to determine the cross-site Lennard-Jones parameters. The Ewald sum technique was used to compute the Coulomb energy of adsorbates whose TraPPE parameters include point charges, with consistent Ewald damping parameter $\alpha = 6.0/L_{\min}$ ($L_{\min}$ is the smallest side length of the simulation cell) and Fourier vectors with $k^2 < 27$. The calculation of the adsorption entropy (as defined by eqs 2.3 and 2.5) was done via Monte Carlo integration (equivalent to the Widom test particle method [64, 65]), in the same mode described by Sarkisov [58]. First, a position and orientation of the adsorbate is generated randomly, and the adsorbate–adsorbent potential energy $U_{\text{fs}}(\mathbf{r}, \boldsymbol{\psi})$ for that position/orientation is computed for the TraPPE-derived model. Then, the appropriate terms in eq 2.3 are accumulated, and the procedure is repeated

for $10^7$ random positions and orientations. This number of trials was satisfactory for numerical convergence for nearly all adsorbent/adsorbate combinations (Subsequent calculations at $10^8$ trials were performed for unconverged adsorbates), where the $\Delta s^{\infty}_{\text{ads}}$ measurements were considered converged when the change in the measured value associated with an order-of-magnitude increase in the number of trials was 1) less than 0.1R, 2) the associated isosteric heat was positive, and 3) the $\Delta s^{\infty}_{\text{ads}}$ was negative.The actual Monte Carlo integrations were done according to the methodology outlined in Chapter 2 Section 2.3.

### *4.2.2 Neural network model*

We have employed a multi-layer, feed-foreword, artificial neural network (NN) model to predict our simulated "$\Delta s^{\infty}_{ads}$" values using the "pytorch" package[155]. The loss function $L$ for training this model is given by:

$$L = \Sigma_i^N (w_i[y_i - \hat{y}_i])^2 \tag{4.1}$$

where $y$ denotes the simulated adsorption entropy $\Delta s^{\infty}_{ads}$ and $N$ denotes the number of data points (i.e the 3965 framework-molecule combinations). Furthermore, $\hat{y}$ denotes the predicted values of $\Delta s^{\infty}_{ads}$. The weights $w_i$ are used to scale the importance of the individual runs, and are inversely proportional to the convergence of $y$. To circumvent numerical overflow and limit the magnitude of $w_i$, a regularization value, $\lambda = 0.5$ was added to the denominator.

$$w_i = |\frac{1}{\Delta y_i^{conv} + \lambda}| \tag{4.2}$$

We posit that by including this weight, the model will be less influenced by the lesser converged data. As mentioned in section 4.2.1, the convergence criteria for our simulated adsorption entropy $\Delta s^{\infty}_{ads}$ was satisfied when the measured value associated with an order-of-magnitude increase in the number of trials was 1) less than 0.1R ($J/mol/K$), 2) the associated isosteric heat was positive, and 3) the $\Delta s^{\infty}_{\text{ads}}$ was negative.

$$\Delta y_i^{conv} = |y_i^{10^6_{trials}} - y_i^{10^7_{trials}}| \leq 0.1R \tag{4.3}$$

The convergence value of 0.1R ($J/mol/K$) was deemed suitable because it was lower than the mean absolute error (MAE) of all of our NN and linear models; and therefore limited the models from capturing numerical inaccuracies of the simulation while permitting a sizable data set of 3965 systems.

Three hidden layers were chosen for the NN (see Figure 4.1). The input layer consisted of 14 neurons corresponding to a set of 9 molecular and 5 framework descriptors, listed in table 4.1. The molecular descriptors were obtained using Rdkit[156], which were calculated from the three-dimensional conformers used in our simulations. Because the hydrogen atoms within the TraPPE forefield are implicit, descriptors which characterize the rotational aspects of the molecule would otherwise be inconsistent with the conformer used in our simulation. For example, methane is represented as a single sphere and therefore lacks any rotational inertia. To circumvent this issue, the hydrogen atoms of our conformer were removed prior to calculating the descriptors, but otherwise maintaining its hydrogen explicit molecular weight. Prior to calculating the framework descriptors, the number of Monte-Carlo trials and probe radius for the Zeo++ software was optimized by calculating the occupiable volume, occupiable surface area, and largest included sphere diameter for zeolite frameworks with

the largest and smallest occupiable volumes: IWV and JSW respectively (structures taken from IZA database[66]). Our rationale is that the framework with the largest occupiable volume will determine the minimum number of Monte-Carlo trials necessary for sampling a large space, while the framework with the smallest occupiable volume (and among the smallest cavity diameters) will determine the maximum radius of the probe. Calculations were run at probe radii of 0.5 to 2.0, in increments of 0.1 Å; and subsequently, Monte-Carlo trials from 25,000 to 500,000 samples (in increments of 25,000 trials) for each probe radius were ran. The smallest error occurred with a probe radius of 1.1 Åat 75,000 Monte-Carlo trials. At these conditions, the percent change with respect to the IZA database was: 3.3%/1.0% for the occupiable surface area, 0.71%/4.8% for the occupiable volume, and 4.7%/7.4% for the largest included sphere diameter of IWV/JSW respectively. The hidden layers of the NN consisted of 81, 71, 91 neurons respectively, and the output consisted of a single neuron corresponding to the adsorption entropy. The NN was trained on the absolute value of the adsorption entropies, and the rectified linear unit (ReLU) was used as the activation function to ensure this criteria. The data set was randomly split into 20% testing and 80% training sets. The ADAM optimizer[157], using the full training set as the batch, was implemented at a learning rate of $10^{-4}$ for a total number of 4000 epochs. This architecture of the neural network was determined after a systematic evaluation of over one-thousand structures with varying number of neurons per layer (1-100), activation functions (CELU, ELU, Hardshrink Hardsigmoid, Hardswish, Hardtanh, logSigmoid, RELU, ReLU6, RReLU, SELU, SIGMOID, Softplus, Softshrink Tanh, Tanhshrink), and learning rates ($10^{-3}, 10^{-4}, 10^{-5}$), to identify the best model in terms of the mean absolute error of the testing set.

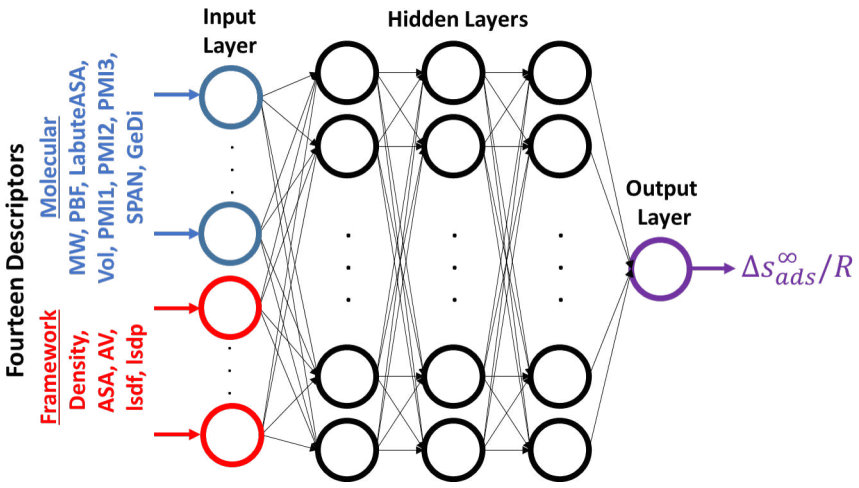


Figure 4.1: Structure of the artificial neural network model developed here. The output is the simulated adsorption entropy at infinite dilution $\Delta s^{\infty}_{ads}$. The input is an array of 14 descriptors, consisting of 9 molecular, and 5 framework descriptors. Color code: Molecular descriptors (blue), Framework descriptors (red), adsorption entropy (purple).

### *4.2.3 Local Linear Models*

Here, we derive the local linear models used to determine the most significant descriptors for our NN model predictions. Let $f_{nn}(\vec{x_i})$ be our NN model, expressed as:

$$f_{nn}(\vec{x_i}) = \hat{y}_i \tag{4.4}$$

where "$\hat{y}_i$", corresponds to the output and represents a single prediction of the absolute value of the adsorption entropy $|\Delta s^{\infty}_{ads}/R|$ for a particular adsorbate-framework system "$i$". Vector $\vec{x_i} = \{x_{i,1}, x_{i,2}, ..., x_{i,N}\}$ of size $N$ denotes our molecular and framework descriptors (listed in table 4.1). The local sensitivity "$S_{x_{i,j}}$" of $f_{nn}(\vec{x_i})$ to a small change in $x_{i,j}$ can be approximated by:

$$S_{x_{i,j}} = \frac{\Delta \hat{y_{i,j}}}{\Delta x_{i,j}}|_{x_{i,z\neq j}} \tag{4.5}$$

where "$j$" represents a particular descriptor of system $i$ and $\Delta \hat{y_{i,j}}$ represents the predicted response to $\Delta x_{i,j}$. We would like to emphasize that by virtue of the universal approximation theorem[158], finite neural networks may only approximate smooth continuous functions, and therefore are not analytically differentiable. Equation 4.5 may be viewed as an approximation of the partial derivative of $f_{nn}(\vec{x_i})$ with respect to $x_{i,j}$, which holds the more general form:

$$\begin{aligned} \Delta \hat{y_i} = \frac{\Delta \hat{y_{i,1}}}{\Delta x_{i,1}}|_{x_{i,j\neq 1}} \Delta x_{i,1} + \frac{\Delta \hat{y_{i,2}}}{\Delta x_{i,2}}|_{x_{i,j\neq 2}} \Delta x_{i,2} + ... + \frac{\Delta \hat{y_{i,N}}}{\Delta x_{i,N}}|_{x_{i,j\neq N}} \Delta x_{i,N} \\ \Delta \hat{y_i} = S_{x_{i,1}} \Delta x_{i,1} + S_{x_{i,2}} \Delta x_{i,2} + ... + S_{x_{i,N}} \Delta x_{i,N} \end{aligned} \tag{4.6}$$

The sensitivities in equation 4.6 measure the individual influence of each descriptor on the prediction, and can be determined by performing a linear least squares regression on $\Delta \vec{x_i}$. This is done by exploring the local behavior of $f_{nn}(\vec{x_i})$ about the vicinity of system $i$ by sampling a perturbation $\Delta \vec{x_i} = \{\Delta x_{i,1}, \Delta x_{i,2}, ..., \Delta x_{i,N}\}$ around $\vec{x_i}$ such that $\Delta \vec{x_i} \in I^N = [-a, a]$ with the prediction

$$f_{nn}(\vec{x_i} + \Delta \vec{x_i}) = \hat{y_i} + \Delta \hat{y_i} \tag{4.7}$$

Sampling "$M$" instances around the nominal point $\vec{x_i}$ for system $i$ gives

$$\boldsymbol{\Delta X_i} = \begin{bmatrix} \Delta \vec{x_{i,1}} \\ \vdots \\ \Delta \vec{x_{i,M}} \end{bmatrix}, \boldsymbol{\Delta Y_i} = \begin{bmatrix} \Delta \hat{y}_{i,1} \\ \vdots \\ \Delta \hat{y}_{i,M} \end{bmatrix} \tag{4.8}$$

where $\boldsymbol{X_i}$ and $\hat{\boldsymbol{Y}}_{\boldsymbol{i}}$ correspond to the $M$ perturbations of $\vec{x_i}$ and their predicted deviations $\Delta\hat{y}_i$ respectively. For small perturbations ($|a| \approx 0$) the relationship between $\boldsymbol{X_i}$ and $\hat{\boldsymbol{Y}}_{\boldsymbol{i}}$ is approximately linear

$$\hat{\boldsymbol{Y}}_{\boldsymbol{i}} \approx \boldsymbol{X_i}\boldsymbol{\beta_i} + \boldsymbol{\epsilon_i} \tag{4.9}$$

with weights $\boldsymbol{\beta_i} = \{\beta_{i,1}, \beta_{i,2}, ..., \beta_{i,N}\}$ and intercept $\boldsymbol{\epsilon_i}$. We emphasize that the intercept for this linear relationship is necessarily 0, because $\Delta\vec{x_i} = 0$ will result in $\Delta\hat{y}_i = 0$. To capture this linearity, we apply the least absolute shrinkage and selection operator (LASSO), which extends the standard linear least-squares regression by including a penalty value "$\lambda$" into the objective function of the linear regression. Although the derivation is outside the scope of this work[159], the theory may be illustrated by the following equation, where the coefficients $\boldsymbol{\beta_i}$ of system $i$ are determined by

$$\boldsymbol{\beta_i} = \underset{\boldsymbol{\beta_i}}{argmin}\{\frac{1}{M}\|\hat{\boldsymbol{Y}}_{\boldsymbol{i}} - \boldsymbol{X_i}\boldsymbol{\beta_i}\|_2^2 + \lambda\|\boldsymbol{\beta_i}\|_1\} \tag{4.10}$$

The penalty $\lambda$ controls the extent of sparsity, and is a hyperparameter of the model (Note, $\lambda = 0$ returns the definition of the standard least squares regression). Large values of $\lambda$ drive insignificant features to zero but diminish the accuracy of the model, while smaller values of $\lambda$ may be incapable of discarding truly insignificant

features. Therefore the penalty $\lambda$ is system specific, and must be tuned to balance between model accuracy and model compactness. Our motivation behind applying the LASSO was to minimize the cumulative influence of correlated descriptors on $|\Delta s^{\infty}_{ads}/R|$. *As a consequence, the resulting weights* $\boldsymbol{\beta_i}$ *of the LASSO are not identical to the sensitivities* $S_{x_{i,j}}$ *defined in equation 4.6. Rather, they are regularized, sparse, representations of* $S_{x_{i,j}}$.

In this work our perturbation space was $-0.01 \leq a \leq 0.01$, the number of samples were $M = 4096$, the penalty for the LASSO was $\lambda = 1 \times 10^{-6}$.

## 4.3 Results and Discussion

Figure 4.2 shows the distribution of the simulated adsorption entropies $|\Delta s^{\infty}_{ads}/R|$ for the 3965 molecule-framework systems. The mean, maximum, and minimum adsorption entropies of the distribution were 7.0R, 14.2R, 1.2R (J/mol/K) respectively. The system which exhibited the smallest entropic loss of 1.2R (J/mol/K) was methane within zeolite RWY, as shown in Figure 4.4 a) and b). Methane is the smallest molecule within our dataset, represented by a single sphere with Leonnard Jones diameter of 3.7 Å. The characteristics of framework RWY are among the largest within our data set, exhibiting the second largest lsd_p (diameter of the largest sphere along free path) of 15.1 Åand the largest accessible volume of 0.66 $cm^3/g$. Systems with significant entropic losses (below two standard deviations from the mean) consisted almost exclusively of small adsorbates: methane, ethane, ethene and perfluoroethane within large pore frameworks: RWY, FAU, SAO, and TSC (each of which have lsd_p $\geq$ 7.7 Å), with the exception of propane and propene within framework RWY.

The system which exhibited the largest entropic loss of 14.2R (J/mol/K) was ethyl-methacrylate within framework SZR, as shown in Figure 4.4 c) and d). Ethyl-methacrylate was not the largest adsorbate within our dataset, exhibiting a geometric

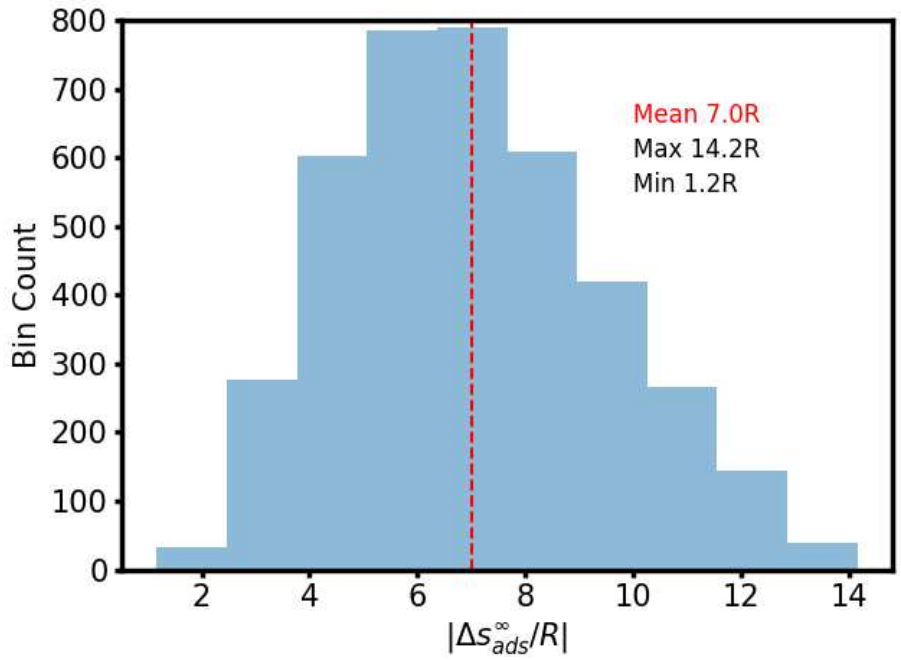


Figure 4.2: Histogram of the 3965 simulated adsorption entropies $|\Delta s^{\infty}_{ads}/R|$. The mean, max and min of the distribution were 7.0R, 14.2R, 1.2R (J/mol/K) respectively.

diameter (maximum distance between two atoms, "GeDi") of 6.2 Å, while the largest adsorbate, n-perfluorodecane, exhibited 11.8 Å. Moreover, framework SZR had an lsd_p of 5.3 Åand an accessible volume of 0.043 $cm^3/g$, both of which were larger than the smallest framework, CAS, which exhibited 0.01 $cm^3/g$ and 3.8 Åfor the accessible volume and lsd_p respectively. We would like to highlight that not every molecule is capable of adsorbing within every framework, and in retrospect, ethyl-methacrylate within SZR was simply the limit for the largest adsorbate within the smallest framework per our convergence threshold of 0.1R (J/mol/K). This system reflects the prevalent characteristics of large entropic loss systems in general: long molecules within tight channels. Figure 4.3 illustrates this trend, where we plot two metrics of framework size (accessible volume and lsd_p) against the GeDi of the adsorbate with the corresponding entropic loss represented by the blue/red color gradient. Additional systems with similar entropic loss (above two standard deviations from

the mean) include: pentanal (5.0 Å), n-perfluorohexane (7.1 Å), 2-ethylhexylacrylate (8.9 Å), and pentane-1,5-diol (6.3 Å) within frameworks with lsd_p and accessible volumes: NPT (5.7Å/0.15 $cm^3/g$), MFI (5.7 Å/0.063 $cm^3/g$), MOR (6.0 Å/0.064 $cm^3/g$), ATN (5.3 Å/0.034 $cm^3/g$) respectively.

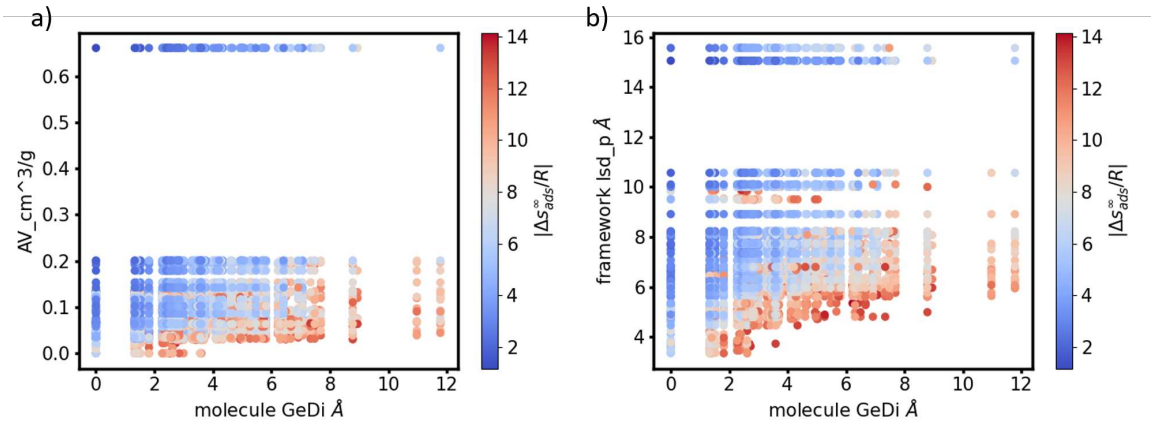


Figure 4.3: Largest distance between two atoms of an adsorbate (GeDi Å) plotted against a) the accessible volume of the framework (AV_cm3̂/g) and b) the diameter of the largest sphere along a free sphere path (lds_p Å) of the framework with corresponding adsorption entropies $|\Delta s^{\infty}_{ads}/R|$ represented by a color gradient (blue/red smallest/largest). Both metrics of framework size, AV_cm3̂/g and lsd_p follow the same trend with respect to adsorbate GeDi: long molecules within tight channels correspond to large entropic loss.

### *4.3.1 Limitations of Linear Correlations*

The standard gas-phase entropy of a molecule "$S^0_{gas}$" was shown to be an excellent descriptor in predicting the entropy of an adsorbate "$S^0_{ads}$" on two-dimensional catalytic, mineral, and van der Waals surfaces[60, 5, 6, 46, 47, 48]. Each of these relationships showed strong linear correlations of the form $S^0_{ads} = \eta S^0_{gas} + b_0$, spanning as much as 50R of entropy space with a standard deviation of as little as 2R (where R is the universal gas constant).[5] These linear relationships were also shown to hold for industrially relevant hydrocarbons within nine distinct aluminosilicate zeolites[7], with slopes $\eta$ depending on the relative size of the framework. In our pre-

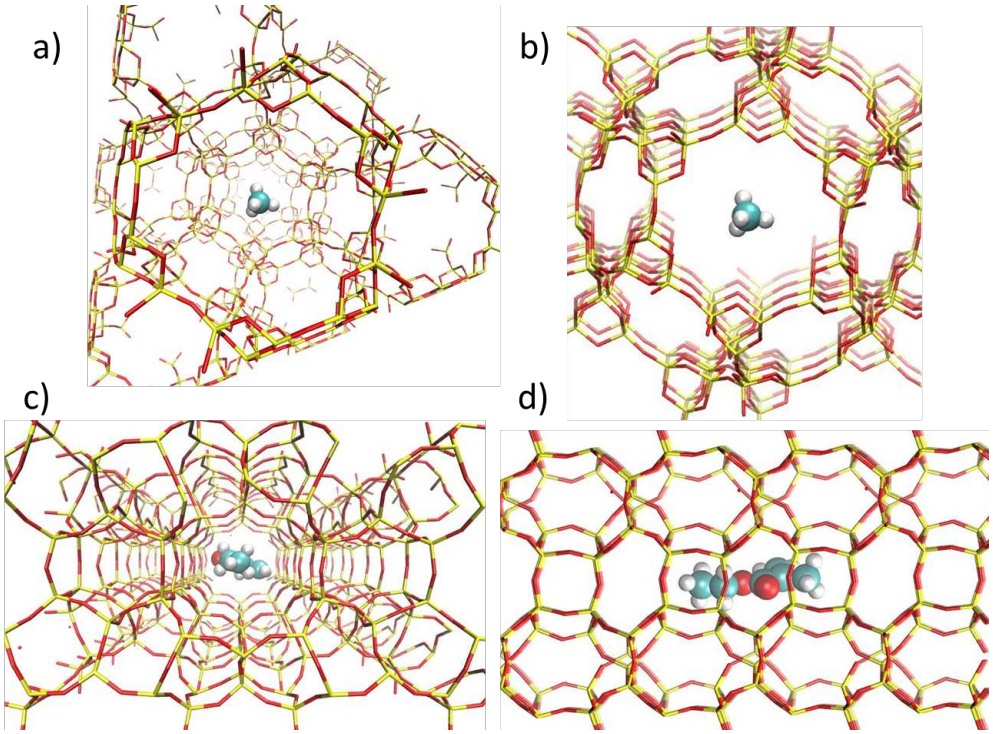


Figure 4.4: Systems corresponding to the maximum and minimum $|\Delta s^{\infty}_{ads}/R|$ of our dataset. a) & b) Two perspectives of methane within framework RWY, the system with the smallest entropic loss $|\Delta s^{\infty}_{ads}/R| = 1.2R$ (J/mol/K) c) & d) Two perspectives of ethyl-methacrylate within framework SZR, the system with the largest entropic loss $|\Delta s^{\infty}_{ads}/R| = 14.2R$ (J/mol/K). Colors represent: carbon(blue), oxygen(red), silicon(yellow), hydrogen(white).

vious work[8], we have qualitatively reproduced these linear trends for a larger set of molecules within pure-silica zeolites using the simulations outlined in this work. The implication of our findings was that the entropic loss was effectively constant across adsorbates within frameworks of similar size characteristics, and that the slopes of their linear relationships $\eta$ followed the same qualitative trend against various framework size metrics. Herein we revisit these conclusions from the perspective our larger data set, and show the limitations of these linear relationships.

Equivalent to our previous work[8], we have performed linear least squares regression of the adsorabte entropies "$s^{\infty}_{ads}/R$" with the form $s^{\infty}_{ads}/R = \eta_i s^{0}_{gas}/R$ for each framework "$i$" in our data set. In Figure 4.5 a), we show the coefficient of determi-

nation (R2) for each regression against two metrics corresponding to the framework size: 1) the occupiable volume in units of $Å^3$ and 2) the diameter of the largest accessible sphere along a free sphere path (lsd_p) Å. Additionally, Figure 4.5 b) and c) show regressions of framework LOV, and AET, which illustrate systems with poor and excellent fits, having R2 values of 0.25 and 1.0 respectively (Their corresponding placements within Figure 4.5 a) have been emphasized with cyan borders). To avoid cluttering the data, we have omitted significant outliers from Figure 4.5 a), which had at most three data points in the regression and consisted of frameworks: SOS, SIV, AEN, LTJ, PHI, and APC with respective R2 values of -18.7, -11.2, -5.0, -3.5, -1.6, and -0.5. *The key aspect of Figure 4.5 a) is that linear correlations hold for many adsorbent systems, but break down for frameworks with small features.* In general, systems with smaller R2 values also consisted of less data points for the regression, which was a consequence of their smaller framework features being unable to accommodate many adsorbates.

In Figure 4.6, we plot the slope "$\eta$" of our least squares regressions against the framework size metrics from Figure 4.5 a). For comparison, we have included the frameworks considered from our previous work: FAU, LTA, MOR, MFI, FER[8], where we performed an identical analysis. Despite applying framework blockages and utilizing energy-optimized structures by Calero et al.[74] (as opposed to using IZA structures previously), the same qualitative trend as in our previous work was reproduced in Figure 4.6. In particular, beginning at the frameworks with larger pore characteristics, $\eta$ begins to slowly decrease with decreasing pore size before precipitously collapsing in the vicinity of $\eta = 0.85$ toward $\eta = 0.74$. As evident by the other datapoints within Figure 4.6, this trend was not followed by the many frameworks regardless of their goodness of fit (R2). *The important result shown in Figure 4.6 is that the relationships between $\eta$ and the framework size metrics are*

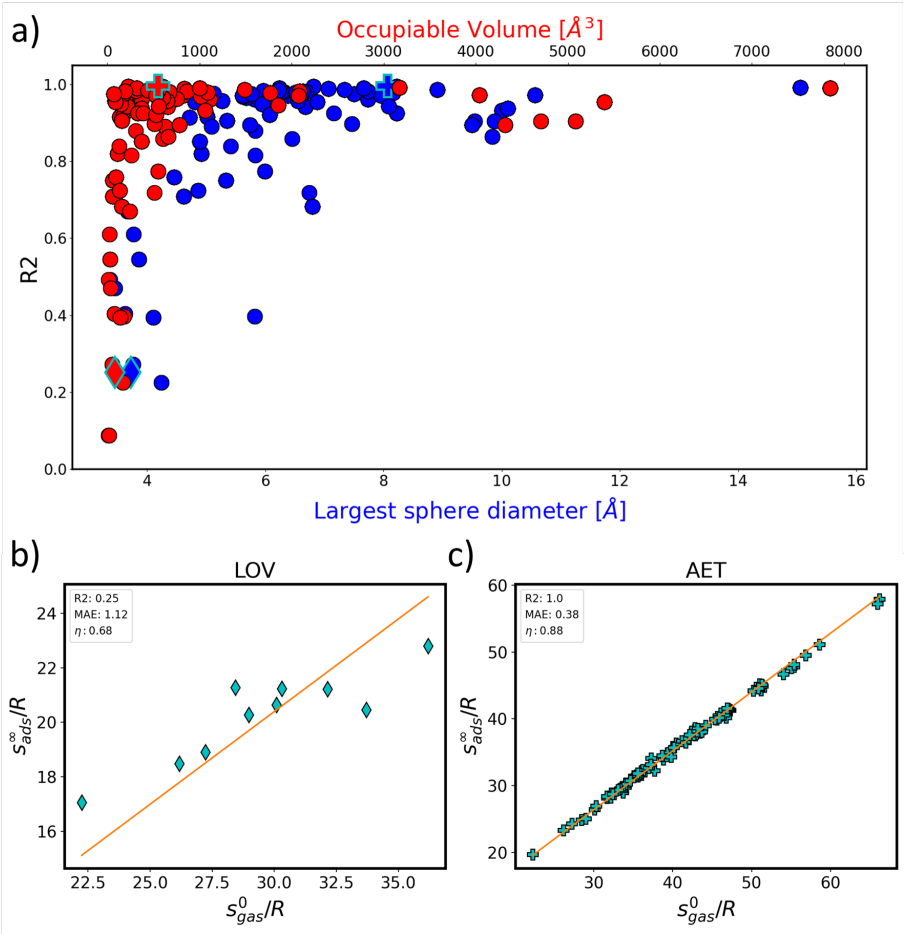


Figure 4.5: a) Coefficient of determination (R2) corresponding to linear least squares regression of the form $s^{\infty}_{ads}/R = \eta_i s^{0}_{gas}/R$ plotted against 1) the occupiable volume in units of Å$^3$ (shown in red) and 2) the lsd_p Å(shown in blue) for each framework "$i$". This figure excludes frameworks: SOS, SIV, AEN, LTJ, PHI, and APC with respective R2 values of -18.7, -11.2, -5.0, -3.5, -1.6, and -0.5. Frameworks LOV and AET have been emphasized with diamond and cross markers with cyan borders respectively. b) Linear least squares regression fit (orange) with corresponding data points (cyan) for framework LOV, representative of a poor fit (R2 = 0.25). c) Linear least squares regression fit (orange) with corresponding data points (cyan) for framework AET, representative of an excellent fit (R2 = 1.0).

*much more complicated than previously anticipated; and therefore a more predictable model is necessary to learn their relationships.* In the following section, we discuss employing an artificial neural network (NN) to predict the adsorption entropy using geometric descriptors of the adsorbate-framework system.

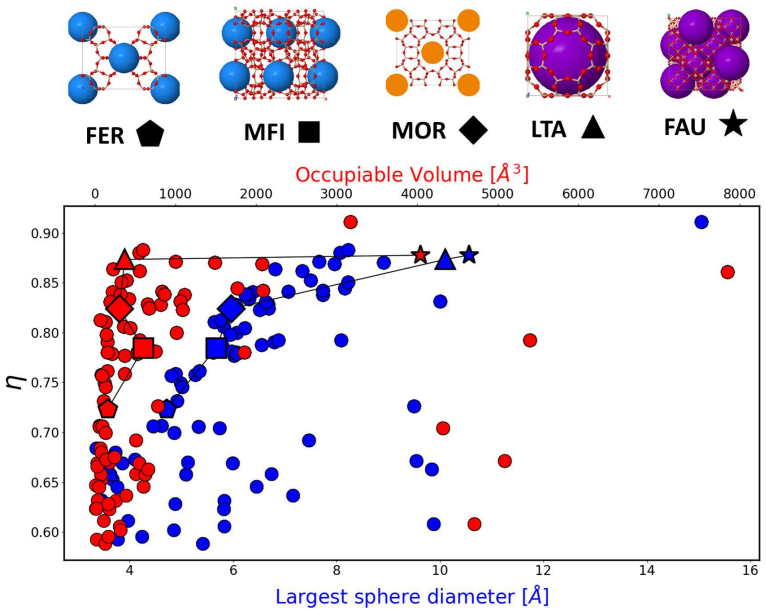


Figure 4.6: Slopes $\eta$ of the linear least squares regression from Figure 4.5 plotted against 1) the occupiable volume in units of Å$^3$ (shown in red) and 2) the lsd_p Å(shown in blue). The trends for frameworks: FAU, LTA, MOR, MFI, FER (represented by the star, triangle, diamond, square, and pentagon markers respectively) are connected by a black line.

### *4.3.2 Neural network predictions and descriptors*

The molecular and framework descriptors applied in the NN model are listed in Table 4.1, and were based primarily on the relationship between "$\Delta s^{\infty}_{ads}$" and the geometry of the adsorbate and framework. Herein, we describe these descriptors and our rationale for choosing them.

The molecular weight is a non-derivable, fundamental, physical quantity with a presence in: 1) the statistical mechanical expression of the translational entropy for

**Molecular Descriptors**

| Abbreviation | Description | Units |
|---|---|---|
| MW | Molecular Weight. | g/mol |
| PBF | Plane of Best Fit. | Å |
| LabuteASA | Labute's Approximate Surface Area. | $Å^2$ |
| Vol | Van der Waals Volume. | $Å^3$ |
| PMI1 | Princple Moment of Inertia about first axis. | $Å^2$amu |
| PMI2 | Princple Moment of Inertia about secound axis. | $Å^2$amu |
| PMI3 | Princple Moment of Inertia about third axis. | $Å^2$amu |
| SPAN | Radius of smallest sphere completely enclsoing all atoms of the molecule from the center of mass. | Å |
| GeDi | Largest distance between two atoms. | Å |

**Framework Descriptors**

| Abbreviation | Description | Units |
|---|---|---|
| Density | Number of tetrahedral sites per 1000 $Å^3$ of unit cell volume. | $Å^{-3}$ |
| ASA_m^2/g | Surface area accessible to the center of a spherical probe of a defined radius normalized by unit cell mass. | $m^2/g$ |
| AV_cm^3/g | Volume accessible to the center of a spherical probe of a defined radius normalized by unit cell mass. | $cm^3/g$ |
| lsd_f | Diameter of largest free sphere. | Å |
| lsd_p | Diameter of largest sphere along free sphere path. | Å |

Table 4.1: Descriptors used for training the artificial neural network, with corresponding abbreviations, descriptions, and units.

an ideal gas, 2) the hindered/free translator entropy approximation for adsorbates within protonic zeolites[151, 152, 110, 145, 44], 3) is related to the number of heavy atoms (non-hydrogen) of the molecule, which has been shown to linearly correlate with the adsorption entropy of alkanes within protonic zeolites.[33] Under standard conditions, the translational entropy is the most significant contribution toward the absolute entropy of a molecule; and predictably experiences the most significant loss upon adsorption.

The rotational moments of inertia make for excellent descriptors toward distinguishing among spherical, linear, and non linear adsorbates. Namely, molecules represented by a single sphere exhibit no rotational inertia, while linear and non-linear molecules rotate about two and three axes respectively. Collectively, the principle moments provide additional information about the shape of the molecule. For example, the ratio of the principle moments (i.e $I_1/I_3$ and $I_2/I_3$) describes the molecule as being 'rod-like', 'sphere-like', 'disc-like', or 'envelope-like'.[160, 161] More complicated expressions, such as the "Molecular eccentricity"[162], "Inertial Shape Factor"[161] and "Spherocity Index"[163] provide additional information about the shape of the molecule. The rotational freedom of a molecule is generally the second

largest contributor to the standard ideal-gas entropy of the adsorbate, and the principle moments of inertia are explicitly present in the rotational entropy equation. The principle moments are also found within the hindered/free rotor approximation for alkanes within protonic zeolites and organic molecules on two-dimensional catalytic surfaces.[44, 113, 123]

The "Labute Approximate Surface Area" descriptor is an efficient approximation of the surface area of a molecule based on its covalent radius.[164] To complement the surface area, we have also included the volume of a molecule based on a grid-encoding of the molecule's shape.[156] Our motivation behind applying these descriptors is: 1) they provide a general metric for the relative shape and size of the molecule, 2) they can be further used to create interpretable dimensionless quantities (i.e the ratio of framework accessible volume to the volume of an adsorbate provides a rough estimate of the maximum adsorption capacity).

The "Plane of Best Fit" (PBF) descriptor is the average distance of the atoms of a molecule from a flat, two-dimensional plane, optimized to intersect the maximum number of its atoms.[165] Mathematically, the PBF has a range of $[0, \infty)$, with small values describing "2D rod" or "disc-like character" (i.e PBF of ethene and toluene $< 6.8 \times 10^{-8}$ Å) and large values describing "sphere-like 3D" molecules (i.e PBF of sarin $> 0.5$ Å).

As mentioned previously, Dauenhauer and Abdelrahman have identified the occupiable volume of protonic zeolites as a "universal" descriptor toward predicting the adsorption entropy of alkanes.[7] In addition to the occupiable volume, their expression for the adsorption entropy depended upon the "critical volume", which was an adjustable parameter defined as the volume at which all rotational freedom of the adsorbate is lost. Consequently, we have decided to incorporate: 1) the radius of the smallest sphere enclosing all atoms of the molecule from its center of mass

"SPAN", and 2) the largest distance between two atoms in the molecule "Geometric Diameter".[161] Although the SPAN correlates with the GeDi ($2 * SPAN \approx GeDi$), this linear relationship breaks down for long chain molecules which provides the model with information to distinguish long molecules. Moreover there is no mathematical, one-to-one relationship between the two descriptors.

Figure 4.7 shows the histogram distributions of our molecular descriptors encompassing our data set. To assess the generality of our descriptors, and thereby the applicability of our NN model to foreign molecules, we have also included the histogram distributions for an equivalent number of quasi-randomly selected (92) molecules from the PubChem database.[153] We consider our data set of molecules from PubChem as "quasi-random", because we have refrained from considering molecular isotopes, radicals, and charged species.-which are outside the scope of this work. To avoid biasing our histograms toward significantly larger molecules (which predominate PubChem), we also preserved the heavy-atom (non hydrogen) distributions of our native data-set within our PubChem sampling. Both histogram distributions are plotted within Figure 4.7 for each molecular descriptor, including their calculated coefficient of intersection "$K$".[166] The significant overlap between our native and PubChem distributions shown in Figure 4.7 suggests that our data set of molecules and their corresponding descriptors resemble the broader amalgam of organic molecules; and therefore our NN model should be capable in adequately predicting their adsorption entropy. The histograms corresponding to our dataset of molecules show a bias toward medium-to-small (GeDi similar to 1,3-butadiene), flat molecules (flatter than diethyl-ether) with mass on the order of n-heptane ($\approx$ 100 g/mol). Additionally, for each descriptor there appear to be a small number of significant outliers, which were determined to mainly consist of fluorocarbons (i.e n-perfluorodecane), with the exception of the SPAN which also included acrylates (i.e n-octyl-acrylate).

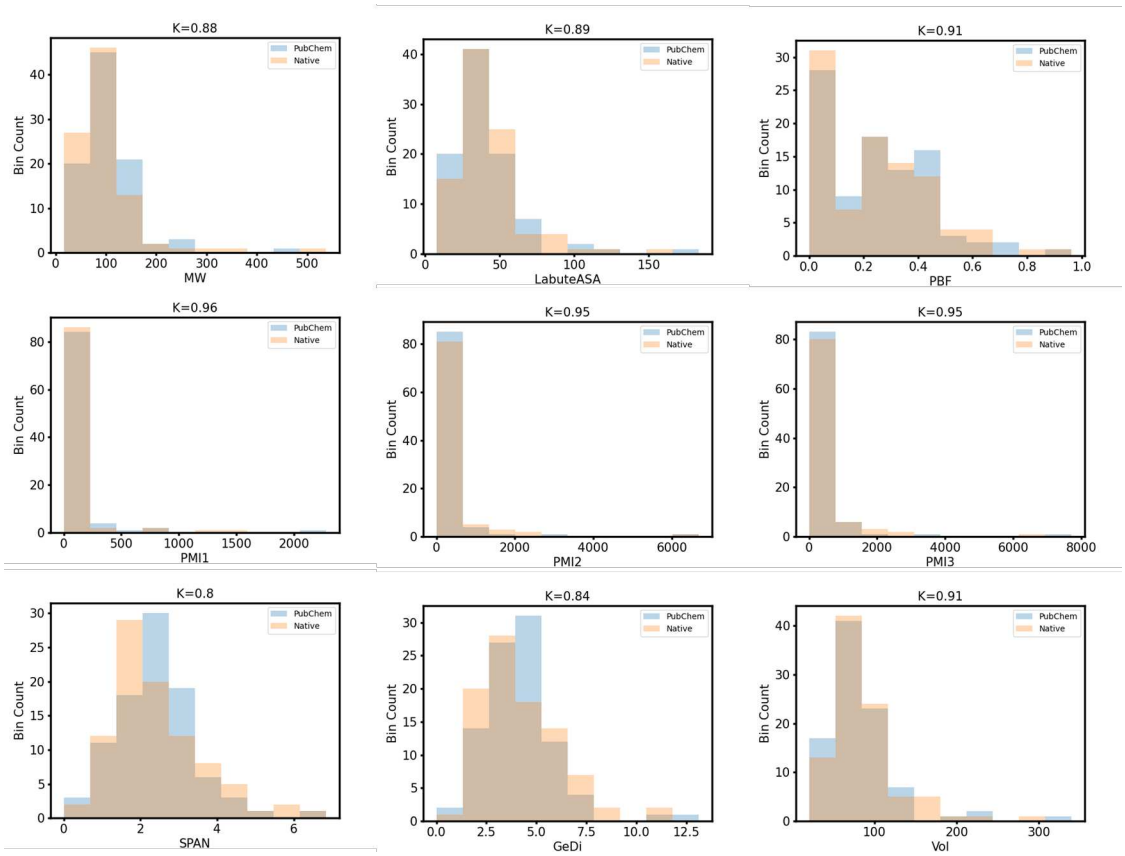


Figure 4.7: Histogram distributions of the molecular descriptors for 1) our native data set (orange) and randomly chosen PubChem molecules (blue), with corresponding histogram intersections "$0 \leq K \leq 1.0$".

As mentioned previously, our framework structures and their blockages were taken from Calero et al.[74], who have implemented the Zeo++ software[167] in determining the inaccessible regions of each structure. To remain consistent with this source, our descriptors were also calculated using Zeo++. We have omitted all descriptors defining the inaccessible characteristics of our frameworks to minimize the correlation among descriptors (i.e the accessible surface area was found to inversely correlate with many non-accessible descriptors). Moreover, these inaccessible regions have been blocked from our simulations. Zeo++ provides the framework accessible surface area and volume in absolute units (i.e $\text{Å}^2$), or normalized by the unit cell volume ($m^2/cm^3$) or mass ($m^2/g$). Additionally, the user has the option of calculating

"accessible" (space available to the center of spherical probe) or "pore occupiable" (space accessible the entire volume of the spherical probe) descriptors.[168] To further minimize descriptor correlation, we have limited our framework descriptors to: 1) the framework "Density", which is the number of tetrahedral sites per 1000 $Å^3$ of unit cell volume, and has been shown to correlate with pore size, 2) the accessible surface area normalized by the unit-cell mass "ASA_m$\hat{2}$/g", 3) the accessible volume normalized by the unit cell mass "AV_cm$\hat{3}$/g", 4) the diameter of the largest sphere along a free sphere path lsd_p, and 5) the diameter of the largest free sphere lsd_f. Figure 4.8 shows the histogram distributions of our framework descriptors within our dataset. Analogous to the analysis of our molecular descriptors, we assessed the generality of our framework descriptors by including histogram distributions for an equal number of quasi-randomly selected (89) frameworks from 1) the IZA database [66], and 2) predicted Crystallography Open Database (PCOD)[169] (a database of hypothtical zeolite frameworks). We consider our sampling as being quasi-random because we have omitted frameworks which lacked any accessible volume (see section 4.2.2). All three histogram distributions are plotted within Figure 4.8 for each framework descriptor, including the coefficient of intersection $K$ for each data set. The overlap between our native and IZA distributions were marginally better relative to the overlap with the PCOD distributions, but were in general worse than the intersections among the molecule distributions. Our native data set (consisting of only orthorhombic frameworks) is therefore less general of the broader space of zeolite frameworks and consequently more limited in predictability. The outliers within the histogram of our native dataset consisted primarily of frameworks similar to RWY, which exhibited an AV_cm$\hat{3}$/g of 0.66, an ASA_m$\hat{2}$/g of 2874.75, a lsd_f of 6.3, and a lsd_p of 15.1. The majority of our data set consisted of frameworks with descriptors similar to MFI: AV_cm$\hat{3}$/g of 0.063, an ASA_m$\hat{2}$/g of 817.70, a lsd_f of 4.7, and a lsd_p

of 5.67.

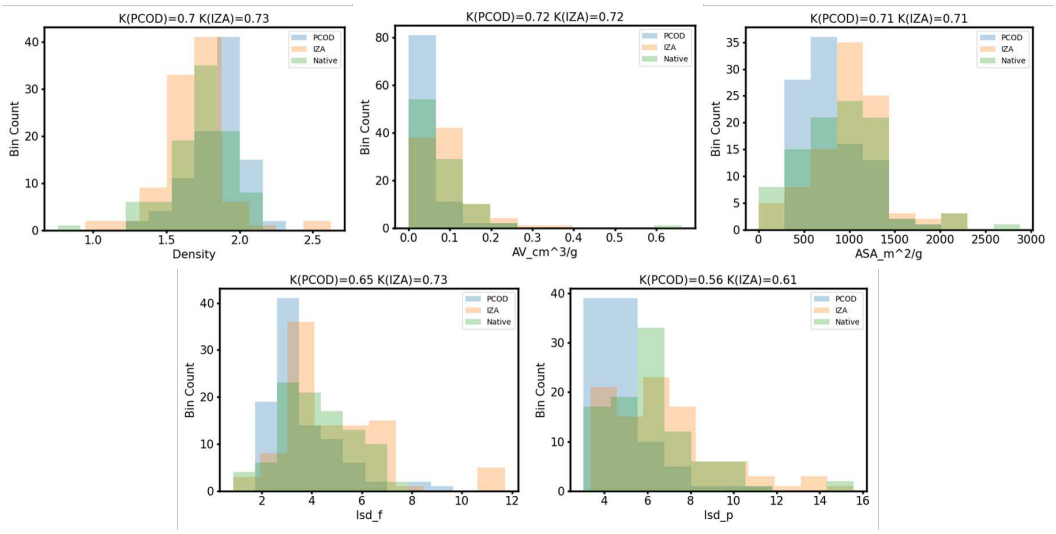


Figure 4.8: Histogram distributions of the framework descriptors for 1) our native data set (green) and randomly chosen hypothetical zeolite frameworks from the PCOD database (blue), and randomly chosen frameworks from the IZA database (orange) with corresponding histogram intersections "$0 \leq K \leq 1.0$".

The predictions for the test and training sets of our optimized NN model are shown in Figure 4.9, with test/training set RMSE, MAE, and R2 of 0.88R/0.5R, 0.61R/0.35R, and 0.87/0.95 respectively. The model was trained on an 80/20 train/test random split, and the model details are provided within section 4.2.2. We observe that most of the deviations from the $y = x$ line in the test and training set occur for large entropic loss $10R \leq |\Delta s^{\infty}_{ads}/R|$ systems, implicating that the model had difficulty learning that region. Systems with the largest residuals showed no clear pattern of outliers (apart from significant entropic loss), and included: RWY/cyclooctadecane (6.5R), CON/2-methyl-1-propanethiol (5.5R), NPT/perfluoroethane (4.4R), LTA/1-octene (4.3R), AWW/n-pentane (3.5R), and BOG/2-hydroxyethyl-acrylate (3.3R). We would like to emphasize that the NN consistently under-predicted this region among training thousands of models. The distribution of $|\Delta s^{\infty}_{ads}/R|$ in Figure 4.2 shows that only $\approx 450$ or 10% of data points comprise a range of $10R \leq |\Delta s^{\infty}_{ads}/R| \leq$

14.2$R$ data. Implying that the sparsity within this region may have affected the model performance. Additionally, we may have missed an important descriptor for characterizing large entropic loss systems. Nevertheless, our model is capable of predicting $|\Delta s^{\infty}_{ads}/R|$ within a MAE of 0.61R, which outperformed over $\approx$ 80% of the linear models within section 4.3.1 with $0.3R \leq MAE \leq 1.6R$.

In the following section, we apply local linear models to rank the importance of our descriptors in predicting the adsorption entropy.

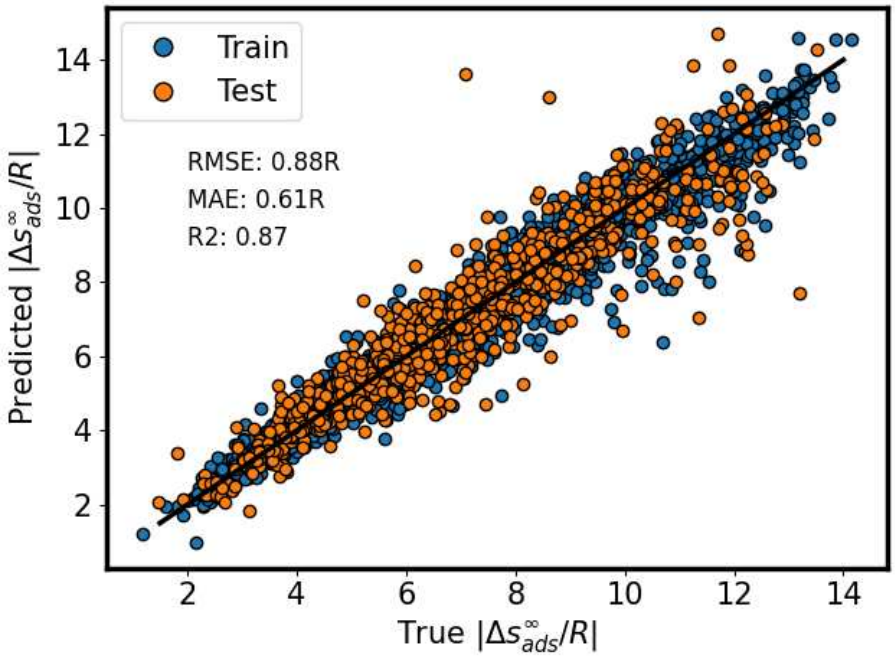


Figure 4.9: Parity plot of the NN predicted and true $|\Delta s^{\infty}_{ads}/R|$ values. The training (blue) and test (orange) sets show that the model has difficulty learning systems with large entropic losses. The RMSE, MAE, and R2 correspond to the test set. The black line represents $y = x$.

### *4.3.3 Local linear models*

Despite their exceptional predicative capabilities, NN models remain mostly opaque (i.e difficult to interpret); and understanding the reasons behind their predictions is challenging. Such difficulties impede establishing *trust* in the model, which is imper-

ative if one plans to initiate further research based on their predictions. For example, we should be skeptical if hypothetically the foremost descriptor of our NN were to be the framework's code (an otherwise random three letter sequence). The demand for trust in opaque models has stimulated research in interpretable machine learning, which introduced many novel interpretation methods.[170] Some of which include: 1) "SHapley Additive exPlanations" (SHAP), which computes the contribution of each feature toward the model's prediction using coalitional game theory,[171] 2) partial dependence plots[172] (PDP), which shows whether the relationship between a feature and model target is linear, monotonic, or more complex, and 3) Local Interpretable Model-Agnostic Explanations, which applies a highly interpretable surrogate model to capture the local importance of features surrounding a prediction.[173] Herein, we adopt a unique approach based on the LIME method to provide an interpretable explanation for the local predictions of our NN model. Although several other methods may have performed comparatively well, the implementation and theory behind our approach are more straightforward (See section 4.2.3).

Figure 4.10 shows the LASSO weights "$<\beta_i>$" for each descriptor across all systems "$i$" averaged across 10 different NN models. *The most important aspect of Figure 4.10 is that the framework occupiable volume is the single most significant descriptor in predicting the adsorption entropy.* This result coincides with the work of Dauenhauer and Abdelrahman[7], who considered the occupiable volume as "a universal descriptor" in predicting the entropy of adsorbed molecules in aluminosilicate zeolite frameworks. Each $<\beta_i>$ also reveals the general relationship between each descriptor with the entropic loss. Negative weights correspond to inverse relationships, which suggest that as: LabuteASA, PBF, Vol, AV_cm$\hat{3}$/g, lsd_f, and lsd_p descrease, the entropic loss increases (the molecule becomes more confined). Likewise, an increase in: MW, PMI(1,2,3), SPAN, GeDI, and density increase the entropic loss. These

relationships show that in addition to determining the significance of each descriptor, the natural relationship between confinement and entropic loss was captured.

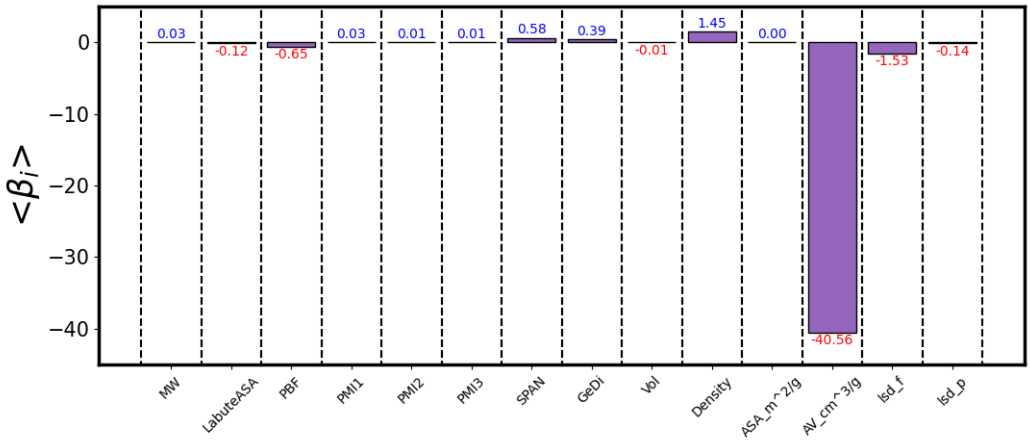


Figure 4.10: Average of the LASSO weights "$<\beta_i>$" for each descriptor across all systems "$i$" for 10 different models. Red/blue values represent negative/positive weights respectively.

## 4.4 Conclusions

In this Chapter, we have built upon our work from Chapter 2 by expanding our data set to include 3965 adsorption entropies, consisting of 89 orthorhombic zeolite frameworks and 92 adsorbates. We have shown that this expanded data set continues to follow the general relationship between adsorbate confinement and entropic loss "$|\Delta s^{\infty}_{ads}|$". We have repeated our analysis of the linear correlations between $s^{\infty}_{ads}$ vs. $s^{0}_{gas}$ of Chapter 2 for our expanded data set and concluded that although these linear trends were largely held, they failed for frameworks with small features (-18.7 $\leq$ R2 $\leq$ 1.0). The relationship between the slope "$\eta$" of the linear regressions against the framework occupiable volume and largest sphere diameter were essentially reproduced for the five frameworks considered in Figure 2.4 of Chapter 2. However, this relationship proved to be much more complicated for the majority of frameworks than previously anticipated (See Figure 4.6). To capture this relationship, we instead

trained a neural network model using fourteen purely geometric descriptors of the adsorbate-framework system to predict $|\Delta s^{\infty}_{ads}|$. Our optimized model was capable of predicting $|\Delta s^{\infty}_{ads}|$ with a mean absolute error of 0.61R. To understand the sensitivity, and thereby the significance of each descriptor in predicting $|\Delta s^{\infty}_{ads}|$, we built linear models on the local perturbations of the neural network descriptors. Our results show that the framework descriptors were more significant in predicting $|\Delta s^{\infty}_{ads}|$ than the adsorbate descriptors, reflecting the conclusion inferred from Figure 2.3 of Chapter 2, where the entropic ratio of each adsorbate "$s^{\infty}_{ads}/s^{0}_{gas}$" clustered about the framework rather than the molecule type. The most significant framework descriptor was the accessible volume, which is closely related to the "Occupiable Volume" deemed to be a universal descriptor in the prediction of adsorption entropies for molecules within aluminosilicate zeolites.[7] Finally, assessing the most significant geometric descriptors in the prediction of $|\Delta s^{\infty}_{ads}|$ paves the way for determining highly interpretable, analytical expressions for the prediction of $|s^{\infty}_{ads}|$; and is the objective of a future work.

# CHAPTER 5

# FUTURE OUTLOOK

## 5.1 Machine intelligent framework selection for Diels Alder promotion

In Chapter 3, we applied a combination of Density Functional Theory (DFT) and microkinetic modeling (MKM) to study the kinetics of [4+2] Diels-Alder reactions between ethene and isoprene within the aluminosilicate zeolite H-ZSM5. Our choice of reactants provided an array of variously shaped (meta- vs. para-) and sized (C7 vs. C10) transition states/products whose corresponding reactions may occur uncatalyzed. This allowed us to gauge the extent of confinement relative to their uncatalyzed gas phase reactions; and infer discriminating factors based on the specie's size/shape. The conclusions of that work were: 1) the reactions were mainly catalyzed by the dispersive interactions with the zeolite framework rather than the Brønsted proton at the acid site, 2) reactions forming the meta products were most catalyzed relative to their respective uncatalyzed gas-phase reactions, 3) the overall reaction (consumption of isoprene) was inhibited by the formation of the C10-para2 and C10-meta2 products, owing largely to their high stability at the active site which limited their desorption and therefore dominated the coverage. Because catalysis was largely driven by the framework structure rather than the Brønsted acid at the active site, the Diels-Alder reactions should be capable of being catalyzed using all-silica zeolite frameworks, and modeled using the techniques applied in Chapter 3. Specifically, one may implement DFT to determine the ground state energies of the reaction system, subsequently build a MKM to predict reaction rates, and perform a global sensitivity analysis to determine rate controlling steps and/or intermediates. The results establish a nominal point for optimizing the reaction which then becomes contingent upon

selecting a framework that destabilizes rate inhibiting intermediates/reactions while stabilizing rate facilitating intermediates/reactions. However, the number of candidate frameworks is enormous, with hundreds having been synthesized[3] and billions proposed on geometric and energetic considerations.[174] In this section, we describe how the neural network model (NN) of Chapter 4 can be leveraged to intelligently select frameworks for the promotion of the Diels-Alder reactions from Chapter 3 within all-silica zeolites.

Transition state theory defines the rate of an elementary reaction step by the frequency of reactants crossing a free energy barrier. The barrier height is dictated by the relative free energy difference between the transition state and the reactant(s), which are assumed to be quasi-equilibriated.[175] Consider the following catalytic reaction:

$$\begin{aligned} &1)\ A + S \rightleftharpoons AS \\ &2)\ AS + B \rightleftharpoons CS \\ &3)\ C + S \rightleftharpoons CS \end{aligned} \tag{5.1}$$

where "$A$" and "$B$" are the reactants, "$C$" is the resulting product, and "$S$" is an unoccupied adsorption site within a heterogeneous catalyst. Adsorbed reactants and products are represented by "$AS$" and "$CS$" respectively. For an ideal gas system, the activity coefficients of the corresponding species are unity, this allows for the rate of an elementary step to be expressed in terms of the reactant partial pressure and/or coverage, and their respective equilibrium constant with the transition state. Under reaction conditions, the adsorption/desorption steps are typically barrierless relative to the product formation step, which also suggests that product formation is

approximately irreversible. This is defined by the following expressions:

$$
\begin{aligned}
r_1 &= 0,\ K_1 = \frac{[AS]}{P_A[S]} \\
r_2 &= \nu K^{\ddagger}[AS]P_B,\ K^{\ddagger} = \frac{[ABS^{\ddagger}]}{[AS]P_B} \\
r_3 &= 0,\ K_3 = \frac{[CS]}{P_C[S]}
\end{aligned}
\tag{5.2}
$$

where "$r_i$" and "$K_i$ are the rate and equilibrium constant of reaction step "$i$", "$\nu$" is a frequency factor, "$P_j$" is the partial pressure of species "$j$" with concentrations and transitions states denoted by brackets, and "‡" respectively. Equation 5.2 shows that the equilibrium constants are the primary quantities necessary in determining the reaction rates. In principle, the reaction system may be optimized if these equilibrium constants are determined across a database of potential catalysts.

The Henry's constant provides the concentration of an adsorbate within an adsorbent (zeolite framework) at a corresponding partial pressure within the low-loading regime. Formally, the Henry's constant for gas adsorption is defined by the following equation:

$$
K_H(T) = \frac{1}{\rho_S V_S} \lim_{p \to 0} \left[ \frac{< N_{ads} >}{p} \right] \tag{5.3}
$$

Where "$K_H$" is the Henry's constant, "T" is the temperature, "$\rho_S = 1/V_P$" is the adsorbent density ($V_P$ is the specific pore volume), "$V_S$" is the bulk volume, "$p$" is the bulk pressure, and "$<N_{ads}>$" is the average amount of adsorbate. Theoretically, $K_H$ provides an approximation to the equilibrium constants defined in equation 5.2. We consider this relation "theoretical" because the $K_H$ of a transition state has

no physical basis, but is otherwise necessary in the determination of reaction rates. The $K_H$ of stable species on the other hand, is well established both experimentally and computationally. We consider $K_H$ an approximation because it is measured without the chemical transformation of the adsorbate, which ultimately influences the equilibrium concentrations.

In reality, the limit of zero partial pressure within equation 5.3 cannot be reached, and the thermodynamic quantities associated with Henry's constant are calculated at a reference pressure. Namely, equation 5.4 may be expanded to include the adsorption entropy and isosteric heat as:

$$\frac{\Delta s^0_{ads}}{R} = \frac{\Delta h^0_{ads}}{RT} + \ln\left(\frac{K_H p^0}{\Gamma^\infty}\right) \tag{5.4}$$

where "$\Delta s^0_{ads}$" is the adsorption entropy, "$\Delta h^0_{ads}$" is the isosteric heat, "$R$" is the ideal gas constant, and "$p^0$" is the reference pressure such that the loading is far below the maximum "$\Gamma^\infty$". The adsorption entropies calculated in Chapter 2 and Chapter 4, were also extracted from their corresponding Henry adsorption constants, which were the primary quantities computed in the Monte-Carlo simulations and exhibit a similar form to equation 5.4:

$$\frac{\Delta s^\infty}{R} = \frac{\Delta h^\infty}{RT} + \ln\left(\frac{RTK^\infty_H}{V_P}\right) + 1 \tag{5.5}$$

We would like to emphasize that although equations 5.4 and 5.5 appear similar, $K^\infty_H$ was calculated within the canonical ensemble at the theoretical reference state of infinite dilution for an adsorbate without internal degrees of freedom. It is certainly invalid to apply it quantitatively, but it nevertheless captures the dominant physical

effects of confinement, which can be qualitatively useful. We instead propose that $K_H^\infty$ may be applied as a surrogate to $K_H$, which can efficiently be used to scan a database of candidate catalysts with the intent to optimize a given reaction system.

Despite the efficiency of calculating $K_H^\infty$, it remains implausible to scour through billions of hypothetical frameworks. Moreover, it is unlikely that the forcefields developed for stable intermediates will be applicable to transition state structures. But based on the conclusions of Chapter 4, we demonstrated that the adsorption entropy may be predicted within 0.6R MAE using a NN trained entirely on geometric descriptors of the adsorbate/framework system. We therefore propose that a model of similar performance may be built to predict the corresponding $K_H^\infty$. Ultimately, applying geometric descriptors circumvents the issue of simulating transition state structures. Although we admit that the model was trained without any such structures, transition states which are geometrically similar to their reactants/products typically exhibit similar relative stabilities (a geometric interpretation of the Bell-Evans-Polanyi relationship).

In conclusion, we posit that the Diels-Alder reactions, discussed in Chapter 3, may be catalyzed within all-silica zeolite frameworks based on the conclusion that these reactions were primarily catalyzed by the framework of H-ZSM5 rather than the Brønsted acid. The kinetic analysis also discussed in Chapter 3 may be applied to determine the rate controlling intermediates and reaction steps within MFI (all-silica analog of H-ZSM5). We propose that this reaction may be optimized by qualitatively predicting the Henry adsorption constants of the rate controlling species among a database of all-silica zeolite frameworks. Namely, an optimal framework would exhibit selective Henry's constants of the rate facilitating species, and unfavorable constants of rate limiting species. Based on the success in predicting the adsorption entropies using a NN trained on the geometric descriptors of the adsorbate/framework system

in Chapter 2, we propose training a similar NN to predict the Henry's constants (the primary quantity calculated in the simulations of Chapter 2). The geometric relationship between transition states with their reactants/products and corresponding stability suggests that the NN may be capable in also predicting transition state structures. The application of a NN makes scanning large databases of zeolite frameworks feasible. Although the deductions made in this section were not formally derived, they follow from the conjectures of each Chapter.

## 5.2 Covalent Organic Frameworks

Covalent organic frameworks (COFs) are rapidly emerging as a novel class of crystalline, multi-dimensional porous materials with unprecedented designability. They arise from the bonding of organic "building-units" which can produce highly tunable skeletons and meso/micro-pores by reticular chemistry. It should be noted that while the number of as yet successful chemistries appear to be few, the number of possible materials that can be formed is still very large. Depending on the choice of the building blocks and its structure, a variety of COF topologies can be formed; and the composition of COFs can be further tuned at essentially any stage of the synthesis process.[176, 177] The dimensions of the pores in a COF can also be tuned by varying the size of the linker even while preserving the fucntional groups and the shape/symmetry of the building blocks. Thus, COFs can have nano-, micro-, or mesoporosity; COFs can also be designed to have multiple types of pores. They generally offer robust thermal and mechanical stability, which make them suitable candidates for catalyzing organic reactions.[178] Moreover, pore surface engineering enables the integration of functional groups into their channels, which may further enhance the reactivity/selectivity of organic reactions. In particular, COF's are capable in accommodating various Brønsted/Lewis acids within their channels. For example,

sulfonic acid COFs with the form COF-SO3Hx have already been synthesized; and have shown to be successful in the catalysis of fructose dehydration into key intermediates for biorefineries.[179] An example of their structure and building units is illustrated in Figure 5.1, which includes the two-dimensional, sulfonated microporous COF named "TFPDABA".[179] Apart from TFP-DABA, the domain of acidic functional groups to be included within COFs is broad; and includes varying degrees of acid-strength and size.

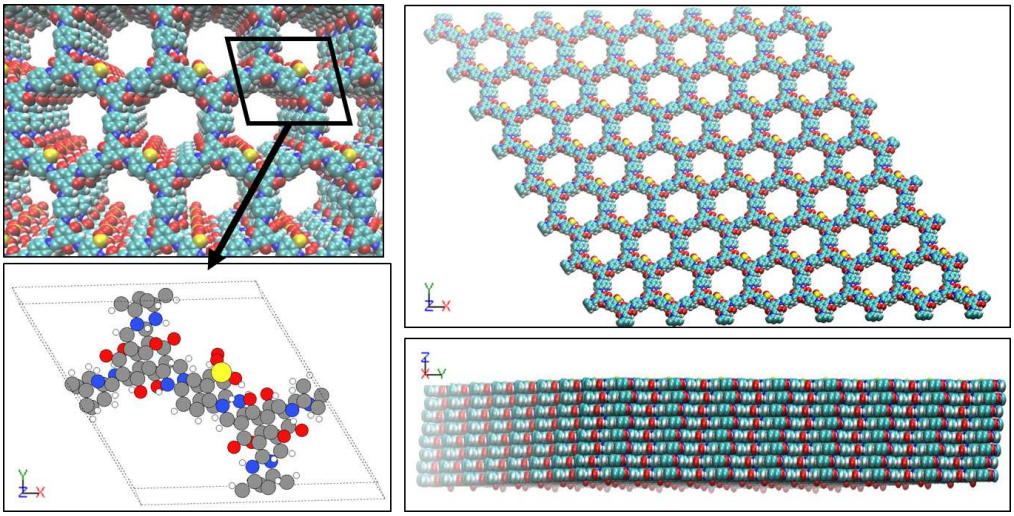

Figure 5.1: An illustration of a two-dimensional, microporous, sulfonated covalent organic framework named "TFPDABA5". The building units consist of 1,3,5-triformylphloroglucinol (TFP) and 2,5-diaminobenzenesulfonic acid (DABA). The colors in the lower left hand image: white, grey, red, blue and yellow represent hydrogen, carbon, oxygen, nitrogen, and sulphur atoms respectively. The pore diameter is 14 Å.

An understanding of the framework's acid-site (strength, coordination, etc.) is an especially valuable quantity for optimizing catalytic reactions. However, the realm of possible combinations among frameworks and acidic functional groups make such calculations experimentally non-trivial; but they may be readily computed using density functional theory (DFT). DFT has proven itself as a valuable method in quantum chemistry, owing to it's astonishing accuracy with experimental results. Modern software such as the Vienna ab-initio Simulation Package (VASP)[102] provides struc-

tural minimization algorithms (i.e steepest descent method) in tandem with DFT to iteratively locate energy minima; which provides accurate energetic and structural information about the system.

Ultimately, these simulations can be used to understand and quantify the nature of acid functionalized COFs with the implications of these results to assist in the catalysis of organic reactions. Although there are many definitions of acidity (Brønsted and Lewis acidity to name a few), hydrogen bonded interaction theory has shown to describe the effect of Brønsted sites for alkane adsorption within other acidic frameworks quite well.[180] This theory suggests that the binding of probe molecules (typically alkanes) to a gas phase Brønsted acid (and other cationic proton donors) can be viewed as a hydrogen-bonded interaction.[181] Analogously, the additional heats of adsorption of alkanes on acidic frameworks (such as zeolites) can be related to the strength of hydrogen bonding between the alkane and the Brønsted proton of the framework. In the complete absence of a solvent, the strength of the hydrogen bond is the proton affinity, but this interaction is moderated in the condensed phase by solvent interactions. Further, while proton affinity intrinsically captures the stability of a three-center-two-electron complex that a proton can form with two neighboring carbon atoms of the alkane and the included bond,[182] the interactions of an alkane with a Brønsted proton does not lead to such a strong covalently bonded complex.[183, 184, 33], Therefore the "acid affinity", $\Delta E_{acid-affinity}$, is instead suggested as an alternative descriptor of the alkane-Brønsted site interaction is suggested. The acid affinity is the energy of complexation for an acid-alkane complex, as given in equation 5.6.

$$\Delta E_{acid-affinity} = E_{probe-acid} - E_{acid} - E_{probe} \tag{5.6}$$

Here, $E_{probe-acid}$ is the energy of the hydrogen-bonded, probe-acid complex, $E_{acid}$ is the energy of the acid in the gas phase, and Eprobe is the energy of the probe molecule in the gas phase. Because the acid affinity of an alkane is expected to depend on the strength of the acid, the optimal choice of acid to explore hydrogen bonding would be one with a deprotonation energy (DPE) equal to that of the acidic framework. It should be noted that the particular choice of acid is, in principle, not critical as any acid could be used as reference. However, one must ensure that the acid is not so weak that the corresponding acid affinity of most alkanes is negligibly small, it is not so strong to completely protonate the probe molecule, and it is not bulky enough to lead to steric effects with larger or branched alkanes.

# REFERENCES


[1] Karolina A. Tarach, Joaquin Martinez-Triguero, Fernando Rey, and Kinga Góra-Marek. Hydrothermal stability and catalytic performance of desilicated highly siliceous zeolites zsm-5. *Journal of Catalysis*, 339:256–269, 2016.

[2] Eric G. Derouane. *Confinement Effects in Sorption and Catalysis by Zeolites*, pages 225–239. Springer US, Boston, MA, 1990.

[3] Database of zeolite structures. `http://www.iza-structure.org/databases/`, 2012. Accessed: 2020-12-07.

[4] R. Pophale, P. A. Cheeseman, and M. W. Deem. A database of new zeolite-like materials. *Phys. Chem. Chem. Phys.*, 13:12407–12412, 2011.

[5] C. T. Campbell and J. R. V. Sellers. The entropies of adsorbed molecules. *J. Am. Chem. Soc.*, 134(43):18109–18115, 2012.

[6] C. T. Campbell and J. R. V. Sellers. Enthalpies and entropies of adsorption on well-defined oxide surfaces: Experimental measurements. *Chem. Rev.*, 113(6):4106–4135, 2013.

[7] P. J. Dauenhauer and O. A. Abdelrahman. A universal descriptor for the entropy of adsorbed molecules in confined spaces. *ACS Cent. Sci.*, 4(9):1235–1243, 2018.

[8] Christopher Rzepa, Daniel W. Siderius, Harold W. Hatch, Vincent K. Shen, Srinivas Rangarajan, and Jeetain Mittal. Computational investigation of correlations in adsorbate entropy for pure-silica zeolite adsorbents. *The Journal of Physical Chemistry C*, 124(30):16350–16361, 2020.

[9] Michito Yoshizawa, Masazumi Tamura, and Makoto Fujita. Diels-alder in aqueous molecular hosts: Unusual regioselectivity and efficient catalysis. *Science*, 312(5771):251–254, 2006.

[10] Hisashi Yamamoto and Kentaro Futatsugi. "designer acids": Combined acid catalysis for asymmetric synthesis. *Angewandte Chemie International Edition*, 44(13):1924–1942, 2005.

[11] Albert Wassermann. 127. homogeneous catalysis of diene syntheses. a new type of third-order reaction. *J. Chem. Soc.*, pages 618–621, 1942.

[12] Peter Yates and Philip Eaton. Acceleration of the diels-alder reaction by aluminum chloride. *Journal of the American Chemical Society*, 82(16):4436–4437, 1960.

[13] Roderigh Y. Rohling, Evgeny Uslamin, Bart Zijlstra, Ionut C. Tranca, Ivo A. W. Filot, Emiel J. M. Hensen, and Evgeny A. Pidko. An active alkali-exchanged faujasite catalyst for p-xylene production via the one-pot diels–alder cycloaddition/dehydration reaction of 2,5-dimethylfuran with ethylene. *ACS Catalysis*, 8(2):760–769, 2018.

[14] Roderigh Y. Rohling, Ionut C. Tranca, Emiel J. M. Hensen, and Evgeny A. Pidko. Electronic structure analysis of the diels–alder cycloaddition catalyzed by alkali-exchanged faujasites. *The Journal of Physical Chemistry C*, 122(26):14733–14743, 2018.

[15] Roderigh Y. Rohling, Ionut C. Tranca, Emiel J. M. Hensen, and Evgeny A. Pidko. Mechanistic insight into the [4 + 2] diels–alder cycloaddition over first row d-block cation-exchanged faujasites. *ACS Catalysis*, 9(1):376–391, 2019.

[16] Nima Nikbin, Phuong T. Do, Stavros Caratzoulas, Raul F. Lobo, Paul J. Dauenhauer, and Dionisios G. Vlachos. A dft study of the acid-catalyzed conversion of 2,5-dimethylfuran and ethylene to p-xylene. *Journal of Catalysis*, 297:35–43, 2013.

[17] C. Luke Williams, Chun-Chih Chang, Phuong Do, Nima Nikbin, Stavros Caratzoulas, Dionisios G. Vlachos, Raul F. Lobo, Wei Fan, and Paul J. Dauenhauer. Cycloaddition of biomass-derived furans for catalytic production of renewable p-xylene. *ACS Catalysis*, 2(6):935–939, 2012.

[18] Phuong T. M. Do, Jesse R. McAtee, Donald A. Watson, and Raul F. Lobo. Elucidation of diels–alder reaction network of 2,5-dimethylfuran and ethylene on hy zeolite catalyst. *ACS Catalysis*, 3(1):41–46, 2013. PMID: 23316419.

[19] Claire Bernardon, Benoît Louis, Valérie Bénéteau, and Patrick Pale. Diels–alder reaction between isoprene and methyl acrylate over different zeolites: Influence of pore topology and acidity. *ChemPlusChem*, 78(9):1134–1141, 2013.

[20] Kevin Maik Jablonka, Daniele Ongari, Seyed Mohamad Moosavi, and Berend Smit. Big-data science in porous materials: Materials genomics and machine learning. *Chemical Reviews*, 120(16):8066–8129, 2020. PMID: 32520531.

[21] Kamal Choudhary, Taner Yildirim, Daniel W. Siderius, A. Gilad Kusne, Austin McDannald, and Diana L. Ortiz-Montalvo. Graph neural network predictions of metal organic framework co2 adsorption properties. *Computational Materials Science*, 210:111388, 2022.

[22] C. M. Simon, R. Mercado, S. K. Schnell, B. Smit, and M. Haranczyk. What are the best materials to separate a xenon/krypton mixture? *Chem. Mater.*, 27(12):4459–4475, 2015.

[23] A. Corma. From microporous to mesoporous molecular sieve materials and their use in catalysis. *Chem. Rev.*, 97(6):2373–2420, 1997.

[24] A. Corma. Inorganic solid acids and their use in acid-catalyzed hydrocarbon reactions. *Chem. Rev.*, 95(3):559–614, 1995.

[25] C. M. Simon, J. Kim, L-C. Lin, R. L. Martin, M. Haranczyk, and B. Smit. Optimizing nanoporous materials for gas storage. *Phys. Chem. Chem. Phys.*, 16:5499–5513, 2014.

[26] J. Kim, A. Maiti, L-C. Lin, J. K. Stolaroff, B. Smit, and R. D. Aines. New materials for methane capture from dilute and medium-concentration sources. *Nat. Commun.*, 4(1):1694, 2013.

[27] Y. Lee, Senja D. Barthel, P. Dłotko, S. M. Moosavi, K. Hess, and B. Smit. High-throughput screening approach for nanoporous materials genome using topological data analysis: Application to zeolites. *J. Chem. Theory Comput.*, 14(8):4427–4437, 2018.

[28] W. Sun, L-C. Lin, X. Peng, and B. Smit. Computational screening of porous metal-organic frameworks and zeolites for the removal of SO2 and NOx from flue gases. *AIChE J.*, 60(6):2314–2323, 2014.

[29] E. Haldoupis, S. Nair, and D. S. Sholl. Pore size analysis of $> 250\,000$ hypothetical zeolites. *Phys. Chem. Chem. Phys.*, 13:5053–5060, 2011.

[30] M. W. Deem, R. Pophale, P. A. Cheeseman, and D. J. Earl. Computational discovery of new zeolite-like materials. *J. Phys. Chem. C*, 113(51):21353–21360, 2009.

[31] M. E. Davis. Ordered porous materials for emerging applications. *Nature*, 417(6891):813–821, 2002.

[32] A. G. Slater and A. I. Cooper. Function-led design of new porous materials. *Science*, 348(6238):aaa8075, 2015.

[33] B. A. De Moor, M-F. Reyniers, O. C. Gobin, J. A. Lercher, and G. B. Marin. Adsorption of C2–C8 n-alkanes in zeolites. *J. Phys. Chem. C*, 115(4):1204–1219, 2011.

[34] J. F. Denayer, G. V. Baron, J. A. Martens, and P. A. Jacobs. Chromatographic study of adsorption of n-alkanes on zeolites at high temperatures. *J. Phys. Chem. B*, 102(17):3077–3081, 1998.

[35] J. F. Denayer, W. Souverijns, P. A. Jacobs, J. A. Martens, and G. V. Baron. High-temperature low-pressure adsorption of branched C5–C8 alkanes on Zeolite Beta, ZSM-5, ZSM-22, Zeolite Y, and Mordenite. *J. Phys. Chem. B*, 102(23):4588–4597, 1998.

[36] G. M. Piccini, M. Alessio, J. Sauer, Y. Zhi, Y. Liu, R. Kolvenbach, A. Jentys, and J. A. Lercher. Accurate adsorption thermodynamics of small alkanes in zeolites. ab initio theory and experiment for H-chabazite. *J. Phys. Chem. C*, 119(11):6128–6137, 2015.

[37] H. Li, S. A. Kadam, A. Vimont, R. F. Wormsbecher, and A. Travert. Monomolecular cracking rates of light alkanes over zeolites determined by IR operando spectroscopy. *ACS Catal*, 6(7):4536–4548, 2016.

[38] A. Beste. One-dimensional anharmonic oscillator: Quantum versus classical vibrational partition functions. *Chem. Phys. Lett.*, 493(1):200–205, 2010.

[39] G. M. Piccini and J. Sauer. Effect of anharmonicity on adsorption thermodynamics. *J. Chem. Theory Comput.*, 10(6):2479–2487, 2014.

[40] K. Alexopoulos, M-S. Lee, Y. Liu, Y. Zhi, Y. Liu, M-F. Reyniers, G. B. Marin, V-A. Glezakou, R. Rousseau, and J. A. Lercher. Anharmonicity and confinement in zeolites: Structure, spectroscopy, and adsorption free energy of ethanol in H-ZSM-5. *J. Phys. Chem. C*, 120(13):7172–7182, 2016.

[41] H. Li, C. Paolucci, and W. F. Schneider. Zeolite adsorption free energies from ab initio potentials of mean force. *J. Chem. Theory Comput.*, 14(2):929–938, 2018.

[42] M. Jørgensen, L. Chen, and H. Grönbeck. Monte carlo potential energy sampling for molecular entropy in zeolites. *J. Phys. Chem. C*, 122(35):20351–20357, 2018.

[43] B. A. De Moor, M-F. Reyniers, O. C. Gobin, J. A. Lercher, and G. B. Marin. Adsorption of C2–C8 n-alkanes in zeolites. *J. Phys. Chem. C*, 115(4):1204–1219, 2011.

[44] L. H. Sprowl, C. T. Campbell, and L. Árnadóttir. Hindered translator and hindered rotor models for adsorbates: Partition functions and entropies. *J. Phys. Chem. C*, 120(18):9719–9731, 2016.

[45] R. Krishna, B. Smit, and S. Calero. Entropy effects during sorption of alkanes in zeolites. *Chem. Soc. Rev.*, 31:185–194, 2002.

[46] Eva Otyepková, Petr Lazar, Klára Čépe, Ondřej Tomanec, and Michal Otyepka. Organic adsorbates have higher affinities to fluorographene than to graphene. *Appl. Mater. Today*, 5:142–149, 2016.

[47] František Karlický, Eva Otyepková, Rabindrana Lo, Michal Pitoňák, Petr Jurečka, Martin Pykal, Pavel Hobza, and Michal Otyepka. Adsorption of organic molecules to van der waals materials: Comparison of fluorographene and fluorographite with graphene and graphite. *J. Chem. Theory Comput.*, 13(3):1328–1340, 2017.

[48] A. Budi, S. L. S. Stipp, and M. P. Andersson. Calculation of entropy of adsorption for small molecules on mineral surfaces. *J. Phys. Chem. C*, 122(15):8236–8243, 2018.

[49] B. L. Eggimann, A. J. Sunnarborg, H. D. Stern, A. P. Bliss, and J. I. Siepmann. An online parameter and property database for the TraPPE force field. *Mol. Sim.*, 40(1-3):101–105, 2014.

[50] F. Eder, , and J. A. Lercher. Alkane sorption in molecular sieves: The contribution of ordering, intermolecular interactions, and sorption on Brønsted acid sites. *Zeolites*, 18(1):75–81, 1997.

[51] F. Eder, M. Stockenhuber, and J. A. Lercher. Brønsted acid site and pore controlled siting of alkane sorption in acidic molecular sieves. *J. Phys. Chem. B*, 101(27):5414–5419, 1997.

[52] L. D. Gelb, K. E. Gubbins, R. Radhakrishnan, and M. Sliwinska-Bartkowiak. Phase separation in confined systems. *Rep. Prog. Phys.*, 62(12):1573–1659, 1999.

[53] M. Thommes, K. Kaneko, A. V. Neimark, J. P. Olivier, F. Rodriguez-Reinoso, J. Rouquerol, and K. S. W Sing. Physisorption of gases, with special reference to the evaluation of surface area and pore size distribution. *Pure Appl. Chem.*, 87:1051–1069, 2015.

[54] A. L. Myers. Thermodynamics of adsorption in porous materials. *AIChE J.*, 48(1):145–160, 2002.

[55] A. L. Myers. Characterization of nanopores by standard enthalpy and entropy of adsorption of probe molecules. *Colloids Surf. A*, 241(1):9–14, 2004.

[56] A. L. Myers and P. A. Monson. Physical adsorption of gases: the case for absolute adsorption as the basis for thermodynamic analysis. *Adsorption*, 20(4):591–622, 2014.

[57] S. Brandani, E. Mangano, and L. Sarkisov. Net, excess and absolute adsorption and adsorption of helium. *Adsorption*, 22(2):261–276, Feb 2016.

[58] L. Sarkisov. Toward rational design of metal–organic frameworks for sensing applications: Efficient calculation of adsorption characteristics in zero loading regime. *J. Phys. Chem. C*, 116(4):3025–3033, 2012.

[59] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids*. Clarendon, New York, 1987.

[60] Charles T. Campbell, Lynza H. Sprowl, and Líney Árnadóttirr. Equilibrium constants and rate constants for adsorbates: Two-dimensional (2d) ideal gas,

2d ideal lattice gas, and ideal hindered translator models. *J. Phys. Chem. C*, 120(19):10283–10297, 2016.

[61] T. L. Hill. *An Introduction to Statistical Thermodynamics*. Addison-Wesley, Reading, MA, 2nd edition, 1962.

[62] D. A. McQuarrie. *Statistical Thermodynamics*. University Science Books, Mill Valley, 1973.

[63] R. D. Johnson III, editor. *NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101.* National Institute of Standards and Technology, Gaithersburg, MD, 20899, Release 20, August 2019.

[64] B. Widom. Some topics in the theory of fluids. *J. Chem. Phys.*, 39(11):2808–2812, 1963.

[65] B. Widom. Potential-Distribution theory and the statistical mechanics of fluids. *J. Phys. Chem.*, 86:869–872, 1982.

[66] Ch. Baerlocher and L. B. McCusker. Database of zeolite structures, 2020. Accessed 23 January 2020.

[67] N. M. O'Boyle, Michael Banck, Craig A. James, Chris Morley, Tim Vandermeersch, and Geoffrey R. Hutchison. Open Babel: An open chemical toolbox. *J. Cheminf.*, 3(1):33, 2011.

[68] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case. Development and testing of a general amber force field. *J. Comput. Chem.*, 25(9):1157–1174, 2004.

[69] J. I. Siepmann Group. TraPPE Database, 2017. Accessed 23 January 2020.

[70] P. Bai, M. Tsapatsis, and J. I. Siepmann. TraPPE-zeo: Transferable potentials for phase equilibria force field for all-silica zeolites. *J. Phys. Chem. C*, 117(46):24375–24387, 2013.

[71] H. W. Hatch, N. A. Mahynski, and V. K. Shen. FEASST: Free Energy and Advanced Sampling Simulation Toolkit. *J. Res. Natl. Inst. Stand. Technol.*, 123:123004, 2018.

[72] E. L. First, C. E. Gounaris, J. Wei, and C. A. Floudas. Computational characterization of zeolite porous networks: An automated approach. *Phys. Chem. Chem. Phys.*, 13:17339–17358, 2011.

[73] M. M. J. Treacy and M. D. Foster. Packing sticky hard spheres into rigid zeolite frameworks. *Micropor. and Mesopor. Mater.*, 118(1):106–114, 2009.

[74] Paula Gómez-Álvarez, A. Rabdel Ruiz-Salvador, Said Hamad, and Sofia Calero. Importance of blocking inaccessible voids on modeling zeolite adsorption: Revisited. *J. Phys. Chem. C*, 121(8):4462–4470, 2017.

[75] J. F. M. Denayer, R. A. Ocakoglu, I. C. Arik, C. E. A. Kirschhock, J. A. Martens, and G. V. Baron. Rotational entropy driven separation of alkane/isoalkane mixtures in zeolite cages. *Angew. Chem. Int. Ed* ., 44(3):400–403, 2005.

[76] A. Corma. Application of zeolites in fluid catalytic cracking and related processes. In P.A. Jacobs and R.A. van Santen, editors, *Zeolites: Facts, Figures, Future Part A - Proceedings of the 8th International Zeolite Conference*, volume 49 of *Studies in Surface Science and Catalysis*, pages 49–67. Elsevier, 1989.

[77] A. Corma. Inorganic solid acids and their use in acid-catalyzed hydrocarbon reactions. *Chem. Rev.*, 95(3):559–614, 1995.

[78] Zhu Jie, Cui Yu, Chen Yuanjun, Zhou Huaqun, Yao Wang, and Fei Wei. Recent researches on process from methanol to olefins. *Journal of Chemical Industry and Engineering (China)*, 61, 07 2010.

[79] Samia Ilias and Aditya Bhan. Mechanism of the catalytic conversion of methanol to hydrocarbons. *ACS Catalysis*, 3(1):18–31, 2013.

[80] Vasant R. Choudhary, Suman K. Jana, and B. P. Kiran. Alkylation of benzene by benzyl chloride over h-zsm-5 zeolite with its framework al completely or partially substituted by fe or ga. *Catalysis Letters*, 59(2):217–219, Jun 1999.

[81] M. Linares, C. Vargas, A. García, C. Ochoa-Hernández, J. Čejka, R. A. García-Muñoz, and D. P. Serrano. Effect of hierarchical porosity in beta zeolites on the beckmann rearrangement of oximes. *Catal. Sci. Technol.*, 7:181–190, 2017.

[82] Sina Sartipi, Kshitij Parashar, Michiel Makkee, Jorge Gascon, and Freek Kapteijn. Breaking the fischer–tropsch synthesis selectivity: direct conversion of syngas to gasoline over hierarchical co/h-zsm-5 catalysts. *Catal. Sci. Technol.*, 3:572–575, 2013.

[83] Li-Hua Chen, Ming-Hui Sun, Zhao Wang, Weimin Yang, Zaiku Xie, and Bao-Lian Su. Hierarchically structured zeolites: From design to application. *Chemical Reviews*, 120(20):11194–11294, 2020. PMID: 32915551.

[84] Berend Smit and Theo L. M. Maesen. Towards a molecular understanding of shape selectivity. *Nature*, 451(7179):671–678, Feb 2008.

[85] Aditya Bhan and Enrique Iglesia. A link between reactivity and local structure in acid catalysis on zeolites. *Accounts of Chemical Research*, 41(4):559–567, 2008. PMID: 18278876.

[86] Gina Noh, Zhichen Shi, Stacey I. Zones, and Enrique Iglesia. Isomerization and beta-scission reactions of alkanes on bifunctional metal-acid catalysts: Consequences of confinement and diffusional constraints on reactivity and selectivity. *Journal of Catalysis*, 368:389–410, 2018.

[87] Steven P. Crossley, Daniel E. Resasco, and Gary L. Haller. Clarifying the multiple roles of confinement in zeolites: From stabilization of transition states to modification of internal diffusion rates. *Journal of Catalysis*, 372:382–387, 2019.

[88] Yilang Liu, Leelavathi Annamalai, and Prashant Deshlahra. Effects of lattice o atom coordination and pore confinement on selectivity limitations for ethane oxidative dehydrogenation catalyzed by vanadium-oxo species. *The Journal of Physical Chemistry C*, 123(46):28168–28191, 2019.

[89] Leelavathi Annamalai, Yilang Liu, Sopuruchukwu Ezenwa, Yanliu Dang, Steven L Suib, and Prashant Deshlahra. Influence of tight confinement on selective oxidative dehydrogenation of ethane on movtenb mixed oxides. *ACS Catalysis*, 8(8):7051–7067, 2018.

[90] Harshul S. Khanna, Anne Mirich, Adam Twombly, Yanliu Dang, Alireza Shirazi-Amin, Inosh Perera, Yilang Liu, Luisa F. Posada, Prashant Deshlahra, and Steven L. Suib. Assessment of micropore accessibility for hydrocarbon oxidation in manganese oxide sieves. *Applied Catalysis A: General*, 635:118557, 2022.

[91] Debdutta Chakraborty and Pratim Kumar Chattaraj. Confinement induced thermodynamic and kinetic facilitation of some diels–alder reactions inside a cb[7] cavitand. *Journal of Computational Chemistry*, 39(3):151–160, 2018.

[92] Jacques-Alexis Funel and Stefan Abele. Industrial applications of the diels–alder reaction. *Angewandte Chemie International Edition*, 52(14):3822–3863, 2013.

[93] Amy E. Settle, Laura Berstis, Nicholas A. Rorrer, Yuriy Roman-Leshkóv, Gregg T. Beckham, Ryan M. Richards, and Derek R. Vardon. Heterogeneous diels–alder catalysis for biomass-derived aromatic compounds. *Green Chem.*, 19:3468–3492, 2017.

[94] Kendall N. Houk, Javier Gonzalez, and Yi Li. Pericyclic reaction transition states: Passions and punctilios, 1935-1995. *Accounts of Chemical Research*, 28(2):81–90, 1995.

[95] Avelino Corma, Sara Iborra, and Alexandra Velty. Chemical routes for the transformation of biomass into chemicals. *Chemical Reviews*, 107(6):2411–2502, 2007. PMID: 17535020.

[96] Nima Nikbin, Shuting Feng, Stavros Caratzoulas, and Dionisios G. Vlachos. p-xylene formation by dehydrative aromatization of a diels–alder product in

lewis and brønsted acidic zeolites. *The Journal of Physical Chemistry C*, 118(42):24415–24424, 2014.

[97] M.A. den Hollander, M. Wissink, M. Makkee, and J.A. Moulijn. Gasoline conversion: reactivity towards cracking with equilibrated fcc and zsm-5 catalysts. *Applied Catalysis A: General*, 223(1):85 – 102, 2002.

[98] Xiangxue Zhu, Shenglin Liu, Yueqin Song, and Longya Xu. Catalytic cracking of c4 alkenes to propene and ethene: Influences of zeolites pore structures and si/al2 ratios. *Applied Catalysis A: General*, 288(1):134 – 142, 2005.

[99] J.S Buchanan. The chemistry of olefins production by zsm-5 addition to catalytic cracking units. *Catalysis Today*, 55(3):207 – 212, 2000.

[100] Arian Ghorbanpour, Jeffrey D. Rimer, and Lars C. Grabow. Periodic, vdw-corrected density functional theory investigation of the effect of al siting in h-zsm-5 on chemisorption properties and site-specific acidity. *Catalysis Communications*, 52:98–102, 2014.

[101] Yiwei Zhang, JingYe Yu, Yu-Hao Yeh, Raymond. J. Gorte, Srinivas Rangarajan, and Manos Mavrikakis. An adsorption study of ch4 on zsm-5, mor, and zsm-12 zeolites. *The Journal of Physical Chemistry C*, 119(52):28970–28978, 2015.

[102] G. Kresse and J. Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science*, 6(1):15 – 50, 1996.

[103] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54:11169–11186, Oct 1996.

[104] G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B*, 59:1758–1775, Jan 1999.

[105] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865–3868, Oct 1996.

[106] Stefan Grimme. Semiempirical gga-type density functional constructed with a long-range dispersion correction. *Journal of Computational Chemistry*, 27(15):1787–1799, 2006.

[107] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu. *The Journal of Chemical Physics*, 132(15):154104, 2010.

[108] Yu-Hao Yeh, Christopher Rzepa, Srinivas Rangarajan, and Raymond J. Gorte. Influence of brønsted-acid and cation-exchange sites on ethene adsorption in zsm-5. *Microporous and Mesoporous Materials*, 284:336 – 342, 2019.

[109] Graeme Henkelman, Blas P. Uberuaga, and Hannes Jónsson. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *The Journal of Chemical Physics*, 113(22):9901–9904, 2000.

[110] Bart A. De Moor, An Ghysels, Marie-Françoise Reyniers, Veronique Van Speybroeck, Michel Waroquier, and Guy B. Marin. Normal mode analysis in zeolites: Toward an efficient calculation of adsorption entropies. *Journal of Chemical Theory and Computation*, 7(4):1090–1101, 2011. PMID: 26606357.

[111] Kristof De Wispelaere, Louis Vanduyfhuys, and Veronique Van Speybroeck. Chapter 6 - entropy contributions to transition state modeling. In C. Richard A. Catlow, Veronique Van Speybroeck, and Rutger A. van Santen, editors, *Modelling and Simulation in the Science of Micro- and Meso-Porous Materials*, pages 189–228. Elsevier, 2018.

[112] Toshihide Baba, Norito Komatsu, Yoshio Ono, and Hisashi Sugisawa. Mobility of the acidic protons in hzsm5 as studied by variable temperature 1h mas nmr. *The Journal of Physical Chemistry B*, 102(5):804–808, 1998.

[113] Cuong M. Nguyen, Bart A. De Moor, Marie-Françoise Reyniers, and Guy B. Marin. Physisorption and chemisorption of linear alkenes in zeolites: A combined qm-pot(mp2//b3lyp:gulp)–statistical thermodynamics study. *The Journal of Physical Chemistry C*, 115(48):23831–23847, 2011.

[114] V.B. Kazansky. Adsorbed carbocations as transition states in heterogeneous acid catalyzed transformations of hydrocarbons. *Catalysis Today*, 51(3):419 – 434, 1999.

[115] James F. Haw, John B. Nicholas, Teng Xu, Larry W. Beck, and David B. Ferguson. Physical organic chemistry of solid acids: Lessons from in situ nmr and theoretical chemistry. *Accounts of Chemical Research*, 29(6):259–267, 1996.

[116] Alexandre Tkatchenko and Matthias Scheffler. Accurate molecular van der waals interactions from ground-state electron density and free-atom reference data. *Phys. Rev. Lett.*, 102:073005, Feb 2009.

[117] Stephan N. Steinmann and Clemence Corminboeuf. A generalized-gradient approximation exchange hole model for dispersion coefficients. *The Journal of Chemical Physics*, 134(4):044117, 2011.

[118] Stephan N. Steinmann and Clemence Corminboeuf. Comprehensive benchmarking of a density-dependent dispersion correction. *Journal of Chemical Theory and Computation*, 7(11):3567–3577, 2011. PMID: 26598255.

[119] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist. Van der waals density functional for general geometries. *Phys. Rev. Lett.*, 92:246401, Jun 2004.

[120] Guillermo Román-Pérez and José M. Soler. Efficient implementation of a van der waals density functional: Application to double-wall carbon nanotubes. *Phys. Rev. Lett.*, 103:096102, Aug 2009.

[121] Kyuho Lee, Éamonn D. Murray, Lingzhu Kong, Bengt I. Lundqvist, and David C. Langreth. Higher-accuracy van der waals density functional. *Phys. Rev. B*, 82:081101, Aug 2010.

[122] Aditya Bhan, Yogesh V. Joshi, W. Nicholas Delgass, and Kendall T. Thomson. Dft investigation of alkoxide formation from olefins in h-zsm-5. *The Journal of Physical Chemistry B*, 107(38):10476–10487, 2003.

[123] Cuong M. Nguyen, Bart A. De Moor, Marie-Françoise Reyniers, and Guy B. Marin. Isobutene protonation in h-fau, h-mor, h-zsm-5, and h-zsm-22. *The Journal of Physical Chemistry C*, 116(34):18236–18249, 2012.

[124] Ville Nieminen, Marek Sierka, Dmitry Yu. Murzin, and Joachim Sauer. Stabilities of c3–c5 alkoxide species inside h-fer zeolite: a hybrid qm/mm study. *Journal of Catalysis*, 231(2):393 – 404, 2005.

[125] P. Cnudde, K. De Wispelaere, J. Van der Mynsbrugge, M. Waroquier, and V. Van Speybroeck. Effect of temperature and branching on the nature and stability of alkene cracking intermediates in h-zsm-5. *Journal of Catalysis*, 345:53 – 69, 2017.

[126] Alexander G. Stepanov, Mikhail V. Luzgin, Vyacheslav N. Romannikov, Vladimir N. Sidelnikov, and Eugenii A. Paukshtis. The nature, structure, and composition of adsorbed hydrocarbon products of ambient temperature oligomerization of ethylene on acidic zeolite h-zsm-5. *Journal of Catalysis*, 178(2):466–477, 1998.

[127] M. N. Mazar, S. Al-Hashimi, A. Bhan, and M. Cococcioni. Methylation of ethene by surface methoxides: A periodic pbe+d study across zeolites. *The Journal of Physical Chemistry C*, 116(36):19385–19395, 2012.

[128] Alexander G. Stepanov, Mikhail V. Luzgin, Vyacheslav N. Romannikov, Vladimir N. Sidelnikov, and Eugenii A. Paukshtis. The nature, structure, and composition of adsorbed hydrocarbon products of ambient temperature oligomerization of ethylene on acidic zeolite h-zsm-5. *Journal of Catalysis*, 178(2):466 – 477, 1998.

[129] M. C. Grady and Raymond J. Gorte. Adsorption of 2-propanol and propene on h-zsm-5: evidence for stable carbenium ion formation. *The Journal of Physical Chemistry*, 89(7):1305–1308, 1985.

[130] Sergio Vernuccio, Elizabeth E. Bickel, Rajamani Gounder, and Linda J. Broadbelt. Microkinetic model of propylene oligomerization on brønsted acidic zeolites at low conversion. *ACS Catalysis*, 9(10):8996–9008, 2019.

[131] Aditya Bhan, Shuo-Huan Hsu, Gary Blau, James M. Caruthers, Venkat Venkatasubramanian, and W. Nicholas Delgass. Microkinetic modeling of propane aromatization over hzsm-5. *Journal of Catalysis*, 235(1):35–51, 2005.

[132] Matthias Vandichel, David Lesthaeghe, Jeroen Van der Mynsbrugge, Michel Waroquier, and Veronique Van Speybroeck. Assembly of cyclic hydrocarbons from ethene and propene in acid zeolite catalysis to produce active catalytic sites for mto conversion. *Journal of Catalysis*, 271(1):67–78, 2010.

[133] Yogesh V. Joshi and Kendall T. Thomson. Embedded cluster (qm/mm) investigation of c6 diene cyclization in hzsm-5. *Journal of Catalysis*, 230(2):440 – 463, 2005.

[134] Stian Svelle, Stein Kolboe, and Ole Swang. Theoretical investigation of the dimerization of linear alkenes catalyzed by acidic zeolites. *The Journal of Physical Chemistry B*, 108(9):2953–2962, 2004.

[135] Wanling Shen. A theoretical study of confinement effect of zeolite on the ethylene dimerization reaction. *Microporous and Mesoporous Materials*, 247:136–144, 2017.

[136] Mark N. Mazar, Saleh Al-Hashimi, Matteo Cococcioni, and Aditya Bhan. beta-scission of olefins on acidic zeolites: A periodic pbe-d study in h-zsm-5. *The Journal of Physical Chemistry C*, 117(45):23609–23620, 2013.

[137] Marek Sierka and Joachim Sauer. Proton mobility in chabazite, faujasite, and zsm-5 zeolite catalysts. comparison based on ab initio calculations. *The Journal of Physical Chemistry B*, 105(8):1603–1613, 2001.

[138] E. M. Evleth, E. Kassab, H. Jessri, M. Allavena, L. Montero, and L. R. Sierra. Calculation of the reaction of ethylene, propene, and acetylene on zeolite models. *The Journal of Physical Chemistry*, 100(27):11368–11374, 1996.

[139] A.M Rigby and M.V Frash. Ab initio calculations on the mechanisms of hydrocarbon conversion in zeolites: Skeletal isomerisation and olefin chemisorption. *Journal of Molecular Catalysis A: Chemical*, 126(1):61–72, 1997.

[140] V.B. Kazansky. Adsorbed carbocations as transition states in heterogeneous acid catalyzed transformations of hydrocarbons. *Catalysis Today*, 51(3):419–434, 1999.

[141] Yueying Chu, Bing Han, Anmin Zheng, and Feng Deng. Influence of acid strength and confinement effect on the ethylene dimerization reaction over solid

acid catalysts: A theoretical calculation study. *The Journal of Physical Chemistry C*, 116(23):12687–12695, 2012.

[142] Duangkamol Gleeson. Application of qm simulations and multivariate analysis in the study of alkene reactivity in the zeolite h-zsm5. *Journal of Chemometrics*, 22(6):372–377, 2008.

[143] M. Boronat and A. Corma. Are carbenium and carbonium ions reaction intermediates in zeolite-catalyzed reactions? *Applied Catalysis A: General*, 336(1):2–10, 2008. Special Issue Dedicated to Professor Jean Sommer on his 70th Birthday.

[144] T. Demuth, X. Rozanska, L. Benco, J. Hafner, R.A. van Santen, and H. Toulhoat. Catalytic isomerization of 2-pentene in h-zsm-22—a dft investigation. *Journal of Catalysis*, 214(1):68–77, 2003.

[145] Cuong M. Nguyen, Bart A. De Moor, Marie-Françoise Reyniers, and Guy B. Marin. Physisorption and chemisorption of linear alkenes in zeolites: A combined qm-pot(mp2//b3lyp:gulp)–statistical thermodynamics study. *The Journal of Physical Chemistry C*, 115(48):23831–23847, 2011.

[146] Charles T. Campbell. The degree of rate control: A powerful tool for catalysis research. *ACS Catalysis*, 7(4):2770–2779, 2017.

[147] Carsten Stegelmann, Anders Andreasen, and Charles T. Campbell. Degree of rate control: How much the energies of intermediates and transition states control rates. *Journal of the American Chemical Society*, 131(23):8077–8082, 2009. PMID: 19341242.

[148] Huijie Tian and Srinivas Rangarajan. Computing a global degree of rate control for catalytic systems. *ACS Catalysis*, 10(22):13535–13542, 2020.

[149] Aditya Bhan, Rajamani Gounder, Josef Macht, and Enrique Iglesia. Entropy considerations in monomolecular cracking of alkanes on acidic zeolites. *Journal of Catalysis*, 253(1):221–224, 2008.

[150] A. Janda, B. Vlaisavljevich, B. Smit, L-C. Lin, and A. T. Bell. Effects of pore and cage topology on the thermodynamics of n-alkane adsorption at Brønsted protons in zeolites at high temperature. *J. Phys. Chem. C*, 121(3):1618–1638, 2017.

[151] Bart A. De Moor, Marie-Françoise Reyniers, Marek Sierka, Joachim Sauer, and Guy B. Marin. Physisorption and chemisorption of hydrocarbons in h-fau using qm-pot(mp2//b3lyp) calculations. *The Journal of Physical Chemistry C*, 112(31):11796–11812, 2008.

[152] Bart A. De Moor, Marie-Françoise Reyniers, and Guy B. Marin. Physisorption and chemisorption of alkanes and alkenes in h-fau: a combined ab initio–statistical thermodynamics study. *Phys. Chem. Chem. Phys.*, 11:2939–2958, 2009.

[153] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang, and E. E. Bolton. *PubChem in 2021: new data content and improved web interfaces.* Nucleic Acids Res., 49(D1), D1388–D1395, 2019.

[154] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox. Gaussian˜16 Revision B.01, 2016. Gaussian Inc. Wallingford CT.

[155] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019.

[156] RDKit: Open-source cheminformatics. `http://www.rdkit.org`. [Version; 2020.09.3].

[157] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2014.

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

[159] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning: with Applications in R.* Springer, 2013.

[160] Wolfgang H. B. Sauer and Matthias K. Schwarz. Molecular shape diversity of combinatorial libraries: A prerequisite for broad bioactivity. *Journal of Chemical Information and Computer Sciences*, 43(3):987–1003, 2003. PMID: 12767158.

[161] Roberto Todeschini and Viviana Consonni. *Descriptors from Molecular Geometry*, chapter VIII.2, pages 1004–1033. John Wiley & Sons, Ltd, 2003.

[162] Gustavo A. Arteca. *Molecular Shape Descriptors*, pages 191–253. John Wiley & Sons, Ltd, 1996.

[163] A. Baumgärtner. Shapes of flexible vesicles at constant volume. *The Journal of Chemical Physics*, 98(9):7496–7501, 1993.

[164] Paul Labute. A widely applicable set of descriptors. *Journal of Molecular Graphics and Modelling*, 18(4):464–477, 2000.

[165] Nicholas C. Firth, Nathan Brown, and Julian Blagg. Plane of best fit: A novel method to characterize the three-dimensionality of molecules. *Journal of Chemical Information and Modeling*, 52(10):2516–2525, 2012. PMID: 23009689.

[166] Michael J. Swain and Dana H. Ballard. Color indexing. *International Journal of Computer Vision*, 7(1):11–32, Nov 1991.

[167] Thomas F. Willems, Chris H. Rycroft, Michaeel Kazi, Juan C. Meza, and Maciej Haranczyk. Algorithms and tools for high-throughput geometry-based analysis of crystalline porous materials. *Microporous and Mesoporous Materials*, 149(1):134–141, 2012.

[168] Daniele Ongari, Peter G. Boyd, Senja Barthel, Matthew Witman, Maciej Haranczyk, and Berend Smit. Accurate characterization of the pore volume in microporous crystalline materials. *Langmuir*, 33(51):14529–14538, 2017. PMID: 28636815.

[169] Armel Le Bail. Inorganic structure prediction with *grinsp*. *Journal of Applied Crystallography*, 38(2):389–395, Apr 2005.

[170] Christoph Molnar. *Interpretable Machine Learning*. 2 edition, 2022.

[171] Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

[172] Jerome H. Friedman. Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5):1189 – 1232, 2001.

[173] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. "why should i trust you?": Explaining the predictions of any classifier, 2016.

[174] Ramdas Pophale, Phillip A. Cheeseman, and Michael W. Deem. A database of new zeolite-like materials. *Phys. Chem. Chem. Phys.*, 13:12407–12412, 2011.

[175] Henry Eyring. The activated complex in chemical reactions. *The Journal of Chemical Physics*, 3(2):107–115, 1935.

[176] Christian S. Diercks and Omar M. Yaghi. The atom, the molecule, and the covalent organic framework. *Science*, 355(6328):eaal1585, 2017.

[177] Peter J. Waller, Steven J. Lyle, Thomas M. Osborn Popp, Christian S. Diercks, Jeffrey A. Reimer, and Omar M. Yaghi. Chemical conversion of linkages in covalent organic frameworks. *Journal of the American Chemical Society*, 138(48):15519–15522, 2016. PMID: 27934009.

[178] Jia Guo and Donglin Jiang. Covalent organic frameworks for heterogeneous catalysis: Principle, current status, and challenges. *ACS Central Science*, 6(6):869–879, 2020. PMID: 32607434.

[179] Yongwu Peng, Zhigang Hu, Yongjun Gao, Daqiang Yuan, Zixi Kang, Yuhong Qian, Ning Yan, and Dan Zhao. Synthesis of a sulfonated two-dimensional covalent organic framework as an efficient solid acid catalyst for biobased chemical conversion. *ChemSusChem*, 8(19):3208–3212, 2015.

[180] Yu-Hao Yeh, Raymond J. Gorte, Srinivas Rangarajan, and Manos Mavrikakis. Adsorption of small alkanes on zsm-5 zeolites: Influence of brønsted sites. *The Journal of Physical Chemistry C*, 120(22):12132–12138, 2016.

[181] Solveig Gaarn Olesen and Steen Hammerum. Hydrogen bonding to alkanes: Computational evidence. *The Journal of Physical Chemistry A*, 113(27):7940–7944, 2009. PMID: 19522536.

[182] Jose Walkimar de M. Carneiro, Paul v. R. Schleyer, Martin Saunders, Richard Remington, Henry F. Schaefer, Arvi Rauk, and Theodore S. Sorensen. Protonated ethane. a theoretical investigation of c2h7+ structures and energies. *Journal of the American Chemical Society*, 116(8):3483–3493, 1994.

[183] Florian Göltl and Jürgen Hafner. Modelling the adsorption of short alkanes in protonated chabazite: The impact of dispersion forces and temperature. *Microporous and Mesoporous Materials*, 166:176–184, 2013. 'Innovations in Zeolites and ordered Porous Solids' Selected Papers and Invited Lectures from the 5th International FEZA Conference (FEZA-2011) July 3-7, 2011.

[184] Christian Tuma and Joachim Sauer. Quantum chemical ab initio prediction of proton exchange barriers between ch4 and different h-zeolites. *The Journal of Chemical Physics*, 143(10):102810, 2015.

[185] D. W. Siderius, N. A. Mahynski, and Vincent K. Shen. Effects of ideality assumptions in calculations of the isosteric heat of adsorption for single-component adsorbates. *J. Chem. Phys.*, 2020. in preparation.

[186] T. Vuong and P. A. Monson. Monte carlo simulation studies of heats of adsorption in heterogeneous solids. *Langmuir*, 12:5425–5432, 1996.

[187] Y-H. Yeh, C. Rzepa, S. Rangarajan, and R. J. Gorte. Influence of brønsted-acid and cation-exchange sites on ethene adsorption in ZSM-5. *Microporous Mesoporous Mater.*, 284:336–342, 2019.

[188] B. Smit and T. L. M. Maesen. Towards a molecular understanding of shape selectivity. *Nature*, 451(7179):671–678, 2008.

[189] M. E. Davis. Shape selective catalysis in industrial applications by n. y. chen, william e. garwood, and francis g. dwyer, marcel dekker, new york, 1989, 320 pp. $99.75. *AIChE J.*, 37(2):310–310, 1991.

[190] John Towns, Timothy Cockerill, Maytal Dahan, Ian Foster, Kelly Gaither, Andrew Grimshaw, Victor Hazlewood, Scott Lathrop, Dave Lifka, Gregory D. Peterson, Ralph Roskies, J. Ray Scott, and Nancy Wilkins-Diehr. XSEDE: Accelerating Scientific Discovery. *Comput. Sci. Eng.*, 16(6):62–74, 2014.

[191] M. T. Aronson, R. J. Gorte, W. E. Farneth, and David White. Carbon-13 nmr identification of intermediates formed by 2-methyl-2-propanol adsorption in h-zsm-5. *Journal of the American Chemical Society*, 111(3):840–846, 1989.

[192] James F. Haw, John B. Nicholas, Weiguo Song, Feng Deng, Zhike Wang, Teng Xu, and Catherine S. Heneghan. Roles for cyclopentenyl cations in the synthesis of hydrocarbons from methanol on zeolite catalyst hzsm-5. *Journal of the American Chemical Society*, 122(19):4763–4775, 2000.

[193] Teng Xu and James F. Haw. Cyclopentenyl carbenium ion formation in acidic zeolites: An in situ nmr study of cyclic precursors. *Journal of the American Chemical Society*, 116(17):7753–7759, 1994.

[194] Teng Xu and James F. Haw. Nmr observation of indanyl carbenium ion intermediates in the reactions of hydrocarbons on acidic zeolites. *Journal of the American Chemical Society*, 116(22):10188–10195, 1994.

[195] James F. Haw. Zeolite acid strength and reaction mechanisms in catalysis. *Phys. Chem. Chem. Phys.*, 4:5431–5441, 2002.

[196] James F. Haw, Benny R. Richardson, Irene S. Oshiro, Noel D. Lazo, and James A. Speed. Reactions of propene on zeolite hy catalyst studied by in situ variable temperature solid-state nuclear magnetic resonance spectroscopy. *Journal of the American Chemical Society*, 111(6):2052–2058, 1989.

[197] Brett R. Beno, K. N. Houk, and Daniel A. Singleton. Synchronous or asynchronous? an "experimental" transition state from a direct comparison of experimental and theoretical kinetic isotope effects for a diels-alder reaction. *Journal of the American Chemical Society*, 118(41):9984–9985, 1996.

[198] Alexander M. Niziolek, Onur Onel, Yannis A. Guzman, and Christodoulos A. Floudas. Biomass-based production of benzene, toluene, and xylenes via methanol: Process synthesis and deterministic global optimization. *Energy & Fuels*, 30(6):4970–4998, 2016.

[199] Junming Sun and Yong Wang. Recent advances in catalytic conversion of ethanol to chemicals. *ACS Catalysis*, 4(4):1078–1090, 2014.

[200] U.s. billion-ton update: Biomass supply for a bioenergy and bioproducts industry. `https://www.energy.gov/eere/bioenergy/downloads/us-billion-ton-update-biomass-supply-bioenergy-and-bioproducts-industry`, 2016. Accessed: 2020-12-07.

[201] Annual energy outlook 2012. `https://www.eia.gov/outlooks/aeo/`, 2012. Accessed: 2020-12-07.

[202] George W. Huber, Sara Iborra, and Avelino Corma. Synthesis of transportation fuels from biomass: Chemistry, catalysts, and engineering. *Chemical Reviews*, 106(9):4044–4098, 2006. PMID: 16967928.

[203] Daniel A. Singleton, Steven R. Merrigan, Brett R. Beno, and K.N. Houk. Isotope effects for lewis acid catalyzed diels-alder reactions. the experimental transition state. *Tetrahedron Letters*, 40(32):5817 – 5821, 1999.

[204] Kendall N. Houk, Javier Gonzalez, and Yi Li. Pericyclic reaction transition states: Passions and punctilios, 1935-1995. *Accounts of Chemical Research*, 28(2):81–90, 1995.

[205] Enrico Catizzone, Alfredo Aloise, Emanuele Giglio, Giorgia Ferrarelli, Micaela Bianco, Massimo Migliori, and Girolamo Giordano. Mfi vs. fer zeolite during methanol dehydration to dimethyl ether: The crystal size plays a key role. *Catalysis Communications*, 149:106214, 2021.

[206] Weili Dai, Chuanming Wang, Xianfeng Yi, Anmin Zheng, Landong Li, Guangjun Wu, Naijia Guan, Zaiku Xie, Michael Dyballa, and Michael Hunger. Identification of tert-butyl cations in zeolite h-zsm-5: Evidence from nmr spectroscopy and dft calculations. *Angewandte Chemie International Edition*, 54(30):8783–8786, 2015.

[207] M. Boronat, P. M. Viruela, and A. Corma. Reaction intermediates in acid catalysis by zeolites: Prediction of the relative tendency to form alkoxides or carbocations as a function of hydrocarbon nature and active site structure.

*Journal of the American Chemical Society*, 126(10):3300–3309, 2004. PMID: 15012161.

[208] Christian Tuma, Torsten Kerber, and Joachim Sauer. The tert-butyl cation in h-zeolites: Deprotonation to isobutene and conversion into surface alkoxides. *Angewandte Chemie International Edition*, 49(27):4678–4680, 2010.

[209] Christian Tuma and Joachim Sauer. Protonated isobutene in zeolites: tert-butyl cation or alkoxide? *Angewandte Chemie International Edition*, 44(30):4769–4771, 2005.

[210] Nilton Rosenbach, Alex P.A. dos Santos, Marcelo Franco, and Claudio J.A. Mota. The tert-butyl cation on zeolite y: A theoretical and experimental study. *Chemical Physics Letters*, 485(1):124 – 128, 2010.

[211] K. Weissermel and H.J. Arpe. *Industrial Organic Chemistry*. Wiley, 2008.

[212] Lidan Ye, Xiaomei Lv, and Hongwei Yu. Engineering microbes for isoprene production. *Metabolic Engineering*, 38:125 – 138, 2016.

[213] Thomas D Sharkey and Sansun Yeh. Isoprene emission from plants. *Annual Review of Plant Physiology and Plant Molecular Biology*, 52(1):407–436, 2001. PMID: 11337404.

[214] Minbo Yang, Xueyu Tian, and Fengqi You. Manufacturing ethylene from wet shale gas and biomass: Comparative technoeconomic analysis and environmental life cycle assessment. *Industrial & Engineering Chemistry Research*, 57(17):5980–5998, 2018.

[215] Yongjin Lee, Senja D. Barthel, Paweł Dłotko, Seyed Mohamad Moosavi, Kathryn Hess, and Berend Smit. High-throughput screening approach for nanoporous materials genome using topological data analysis: Application to zeolites. *Journal of Chemical Theory and Computation*, 14(8):4427–4437, 2018. PMID: 29986145.

[216] John D. Sherman. Synthetic zeolites and other microporous oxide molecular sieves. *Proceedings of the National Academy of Sciences*, 96(7):3471–3478, 1999.

[217] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, Nov 1964.

[218] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, Nov 1965.

[219] C. Stampfl, S. Schwegmann, H. Over, M. Scheffler, and G. Ertl. Structure and stability of a high-coverage $(1 \times 1)$ oxygen phase on ru(0001). *Phys. Rev. Lett.*, 77:3371–3374, Oct 1996.

[220] Kenneth C. Hass, William F. Schneider, Alessandro Curioni, and Wanda Andreoni. The chemistry of water on alumina surfaces: Reaction dynamics from first principles. *Science*, 282(5387):265–268, 1998.

[221] Ali Alavi, Peijun Hu, Thierry Deutsch, Pier Luigi Silvestrelli, and Jürg Hutter. Co oxidation on pt(111): An ab initio density functional theory study. *Phys. Rev. Lett.*, 80:3650–3653, Apr 1998.

[222] C. Stampfl and C. G. Van de Walle. Density-functional calculations for iii-v nitrides using the local-density approximation and the generalized gradient approximation. *Phys. Rev. B*, 59:5521–5535, Feb 1999.

[223] Mark Saeys, Marie-Françoise Reyniers, Guy B. Marin, and Matthew Neurock. Density functional study of benzene adsorption on pt(111). *The Journal of Physical Chemistry B*, 106(30):7489–7498, 2002.

[224] Srinivas Rangarajan, Christos T. Maravelias, and Manos Mavrikakis. Sequential-optimization-based framework for robust modeling and design of heterogeneous catalytic systems. *The Journal of Physical Chemistry C*, 121(46):25847–25863, 2017.

[225] Jia Guo and Donglin Jiang. Covalent organic frameworks for heterogeneous catalysis: Principle, current status, and challenges. *ACS Central Science*, 6(6):869–879, 2020.

[226] Xuesong Ding, Jia Guo, Xiao Feng, Yoshihito Honsho, Jingdong Guo, Shu Seki, Phornphimon Maitarad, Akinori Saeki, Shigeru Nagase, and Donglin Jiang. Synthesis of metallophthalocyanine covalent organic frameworks that exhibit high carrier mobility and photoconductivity. *Angewandte Chemie International Edition*, 50(6):1289–1293, 2011.

[227] Pierre Kuhn, Markus Antonietti, and Arne Thomas. Porous, covalent triazine-based frameworks prepared by ionothermal synthesis. *Angewandte Chemie International Edition*, 47(18):3450–3453, 2008.

[228] Yang Wu, Hong Xu, Xiong Chen, Jia Gao, and Donglin Jiang. A $\pi$-electronic covalent organic framework catalyst: $\pi$-walls as catalytic beds for diels–alder reactions under ambient conditions. *Chem. Commun.*, 51:10096–10098, 2015.

[229] Xing Han, Qingchun Xia, Jinjing Huang, Yan Liu, Chunxia Tan, and Yong Cui. Chiral covalent organic frameworks with high chemical stability for heterogeneous asymmetric catalysis. *Journal of the American Chemical Society*, 139(25):8693–8697, 2017. PMID: 28595384.

[230] Adrien P. Côté, Annabelle I. Benin, Nathan W. Ockwig, Michael O'Keeffe, Adam J. Matzger, and Omar M. Yaghi. Porous, crystalline, covalent organic frameworks. *Science*, 310(5751):1166–1170, 2005.

[231] Qi Sun, Yongquan Tang, Briana Aguila, Sai Wang, Feng-Shou Xiao, Praveen K. Thallapally, Abdullah M. Al-Enizi, Ayman Nafady, and Shengqian Ma. Reaction environment modification in covalent organic frameworks for catalytic performance enhancement. *Angewandte Chemie International Edition*, 58(26):8670–8675, 2019.

[232] Alessandro Laio and Michele Parrinello. Escaping free-energy minima. *Proceedings of the National Academy of Sciences*, 99(20):12562–12566, 2002.

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

[234] Samuel L. C. Moors, Kristof De Wispelaere, Jeroen Van der Mynsbrugge, Michel Waroquier, and Veronique Van Speybroeck. Molecular dynamics kinetic study on the zeolite-catalyzed benzene methylation in zsm-5. *ACS Catalysis*, 3(11):2556–2567, 2013.

[235] Colin D. Abernethy, Gareth M. Codd, Mark D. Spicer, and Michelle K. Taylor. A highly stable N-heterocyclic carbene complex of trichloro-oxo-vanadium(V) displaying novel Cl—C(carbene) bonding interactions. *J. Am. Chem. Soc.*, 125(5):1128–1129, 2003.

[236] Anthony J. Arduengo, III, H. V. Rasika Dias, Richard L. Harlow, and Michael Kline. Electronic stabilization of nucleophilic carbenes. *J. Am. Chem. Soc.*, 114(14):5530–5534, 1992.

[237] Anthony J. Arduengo, III, Siegfried F. Gamper, Joseph C. Calabrese, and Fredric Davidson. Low-coordinate carbene complexes of nickel(0) and platinum(0). *J. Am. Chem. Soc.*, 116(10):4391–4394, 1994.

[238] Leah N. Appelhans, Daniele Zuccaccia, Anes Kovacevic, Anthony R. Chianese, John R. Miecznikowski, Aleco Macchioni, Eric Clot, Odile Eisenstein, and Robert H. Crabtree. An anion-dependent switch in selectivity results from a change of C—H activation mechanism in the reaction of an imidazolium salt with $IrH_5(PPh_3)_2$. *J. Am. Chem. Soc.*, 127(46):16299–16311, 2005.

[239] Anne M. Coghill and Lorrin R. Garson, editors. *The ACS Style Guide*. Oxford University Press, Inc. and The American Chemical Society, New York, 3 edition, 2006.

[240] Frank Albert Cotton, Geoffrery Wilkinson, Carlos A. Murillio, and Manfred Bochmann. *Advanced Inorganic Chemistry*. Wiley, Chichester, 6 edition, 1999.

[241] Angel Abarca, Pilar Gómez-Sal, Avelino Martín, Miguel Mena, Josep María Poblet, and Carlos Yélamos. Ammonolysis of

mono(pentamethylcyclopentadienyl) titanium(IV) derivatives. *Inorg. Chem.*, 39(4):642–651, 2000.

[242] Christopher Rzepa, Daniel W. Siderius, Harold W. Hatch, Vincent K. Shen, Srinivas Rangarajan, and Jeetain Mittal. Computational investigation of correlations in adsorbate entropy for pure-silica zeolite adsorbents. *The Journal of Physical Chemistry C*, 124(30):16350–16361, 2020.

[243] Berend Smit. Molecular simulations of zeolites: Adsorption, diffusion, and shape selectivity. *Chemical Reviews*, 108(10):4125–4184, 2008. PMID: 18817356.

[244] J. Hafner, L. Benco, and T. Bučko. Acid-based catalysis in zeolites investigated by density-functional methods. *Topics in Catalysis*, 37(1):41–54, Mar 2006.

[245] Jiří Klimeš and Angelos Michaelides. Perspective: Advances and challenges in treating van der waals dispersion forces in density functional theory. *The Journal of Chemical Physics*, 137(12):120901, 2012.

[246] T.J.Gricus Kofke and R.J. Gorte. A temperature-programmed desorption study of olefin oligomerization h-zsm-5. *Journal of Catalysis*, 115(1):233–243, 1989.

[247] Yogesh V. Joshi, Aditya Bhan, and Kendall T. Thomson. Dft-based reaction pathway analysis of hexadiene cyclization via carbenium ion intermediates: Mechanistic study of light alkane aromatization catalysis. *The Journal of Physical Chemistry B*, 108(3):971–980, 2004.

[248] J.S. Buchanan, J.G. Santiesteban, and W.O. Haag. Mechanistic considerations in acid-catalyzed cracking of olefins. *Journal of Catalysis*, 158(1):279–287, 1996.

[249] Aditya Bhan, Rajamani Gounder, Josef Macht, and Enrique Iglesia. Entropy considerations in monomolecular cracking of alkanes on acidic zeolites. *Journal of Catalysis*, 253(1):221–224, 2008.

[250] Jing Luo, Balaaji V. Bhaskar, Yu-Hao Yeh, and Raymond J. Gorte. n-hexane cracking at high pressures on h-zsm-5, h-bea, h-mor, and usy for endothermic reforming. *Applied Catalysis A: General*, 478:228–233, 2014.

[251] Rajamani Gounder and Enrique Iglesia. The roles of entropy and enthalpy in stabilizing ion-pairs at transition states in zeolite acid catalysis. *Accounts of Chemical Research*, 45(2):229–238, 2012. PMID: 21870839.

[252] Haowei Peng, Zeng-Hui Yang, John P. Perdew, and Jianwei Sun. Versatile van der waals density functional based on a meta-generalized gradient approximation. *Phys. Rev. X*, 6:041005, Oct 2016.

[253] Aron J. Cohen, Paula Mori-Sánchez, and Weitao Yang. Insights into current limitations of density functional theory. *Science*, 321(5890):792–794, 2008.

[254] Abhishek Gumidyala, Bin Wang, and Steven Crossley. Direct carbon-carbon coupling of furanics with acetic acid over brønsted zeolites. *Science Advances*, 2(9):e1601072, 2016.

[255] Sereina Riniker and Gregory A. Landrum. Better informed distance geometry: Using what we know to improve conformation generation. *Journal of Chemical Information and Modeling*, 55(12):2562–2574, 2015. PMID: 26575315.

# CHAPTER 6

# APPENDIX

## A.1 Derivation of the adsorption entropy at dilute conditions

In the following appendix, we derive the entropy of adsorption in the case of infinite dilution and show how it may be related to familiar thermodynamic quantities (cf. eq 2.5). To begin, we define our adsorption entropy in the limit of infinite dilution via

$$\Delta s_{\text{ads}}^{\infty} = \lim_{N/V \to 0} \left[ \left( \frac{\partial S}{\partial N} \right)_{V,T,ads} - \left( \frac{\partial S}{\partial N} \right)_{V,T,gas} \right] \tag{A1}$$

$$= S_{\text{ads}}(N=1,V,T) - S_{\text{gas}}(N=1,V,T) \tag{A2}$$

where the subscripts "ads" and "gas" labels identify the adsorbed (confined) and gas (bulk) phases, respectively. Given the constraints of a fixed number of molecules ($N$), volume ($V$), and temperature ($T$), it is natural to continue working in the canonical ensemble. (As in the main text, lowercase symbols identify intensive molecular or molar properties, whereas the uppercase indicates an extensive property.) For the molecules that we examine in this work (e.g., rigid molecules with no internal degrees of freedom), the canonical partition function $Q$ may be expressed as

$$Q(N,V,T) = Q_{\text{trans}}(N,V,T)\, Q_{\text{rot}}(N,V,T)\, Z(N,V,T) \tag{A3}$$

where $Q_{\text{trans}}$ and $Q_{\text{rot}}$ are the translational and rotational partition functions (which are unaffected by confinement) and $Z$ is the configurational partition function. The translational and rotational partition functions are available in the usual statistical mechanics texts [61, 62], but are ultimately unimportant for the derivation that follows.

Continuing, the configurational partition function may be written as

$$Z\left(N,V,T\right)=\frac{1}{\Omega^{N}}\int_{V}d\mathbf{r}^{N}\int_{\psi}d\boldsymbol{\psi}^{N}\exp\left[-\beta U\left(\mathbf{r}^{N},\boldsymbol{\psi}^{N}\right)\right] \tag{A4}$$

where $\beta=1/k_{\mathrm{B}}T$ and $U\left(\mathbf{r}^{N},\boldsymbol{\psi}^{N}\right)$ is the potential energy (no kinetic contributions) of a molecule at position $\mathbf{r}$ and with Euler angles (orientation) $\boldsymbol{\psi}$. $\Omega=8\pi^{2}V$ is the "volume" of the configurational space (i.e., it includes both positional and orientational degrees of freedom).

Using the bridge function for the Helmholtz free energy, $F=-k_{\mathrm{B}}T\ln Q$, the entropy is given by

$$\begin{aligned}S\left(N,V,T\right)/k_{\mathrm{B}}&=\frac{3}{2}N+\beta E_{\mathrm{rot}}+\beta\left\langle U\right\rangle\left(N,V,T\right)\\&+\ln Q_{\mathrm{trans}}\left(N,V,T\right)+\ln Q_{\mathrm{rot}}\left(N,V,T\right)+\ln Z\left(N,V,T\right)\end{aligned} \tag{A5}$$

where $\langle U\rangle$ is the ensemble average potential energy,

$$\left\langle U\right\rangle\left(N,V,T\right)=\frac{1}{\Omega^{N}}\frac{\int_{V}\int_{\psi}U\left(\mathbf{r},\boldsymbol{\psi}\right)\exp\left[-\beta U\left(\mathbf{r},\boldsymbol{\psi}\right)\right]d\mathbf{r}^{N}d\boldsymbol{\psi}^{N}}{Z\left(N,V,T\right)}, \tag{A6}$$

and $E_{\mathrm{rot}}$ is the molecule-specific rotational kinetic energy. We note that the kinetic energy and partition functions for the translational and rotational degrees of freedom in eq A6 are constants independent of the adsorbent characteristics.

In the limit of $N=1$, the adsorption entropy (eq 2.5) may thus be written as

$$\frac{\Delta s_{\mathrm{ads}}^{\infty}}{k_{\mathrm{B}}}=\beta\left\langle U_{\mathrm{fs}}\right\rangle\left(1,V,T\right)+\ln Z_{\mathrm{ads}}\left(1,V,T\right) \tag{A7}$$

In the equation above, we have taken advantage of a number of simplifications. First, the volumes of the gas and adsorbed phases are taken to be identical; $Q_{\mathrm{trans}}$ is therefore identical for the two phases. Second, $Q_{\mathrm{rot}}$ is, by nature, identical in both

phases. Third, we have exploited the fact that for a single adsorbed molecule the potential energy $U$ is just the adsorbate–adsorbent potential energy $U_{\mathrm{fs}}$. Finally, $\langle U \rangle = 0$ and $Z_{\mathrm{gas}} = 1$ for a single, rigid gas molecule in isolation. Furthermore, the use of eqs A4 and A6 allows us to express the infinite dilution adsorption entropy for a rigid molecule as

$$\begin{aligned} \frac{\Delta s_{\mathrm{ads}}^{\infty}}{k_{\mathrm{B}}} = & \frac{\int_V \int_\psi \beta U_{\mathrm{fs}}(\mathbf{r}, \boldsymbol{\psi}) \exp\left[-\beta U_{\mathrm{fs}}(\mathbf{r}, \boldsymbol{\psi})\right] d\mathbf{r} d\boldsymbol{\psi}}{\int_V \int_\psi \exp\left[-\beta U_{\mathrm{fs}}(\mathbf{r}, \boldsymbol{\psi})\right] d\mathbf{r} d\boldsymbol{\psi}} \\ & + \ln\left[\frac{1}{\Omega} \int_V \int_\psi \exp\left[-\beta U_{\mathrm{fs}}(\mathbf{r}, \boldsymbol{\psi})\right] d\mathbf{r} d\boldsymbol{\psi}\right] \end{aligned} \tag{A8}$$

At this point, it is helpful to convert the integrals in eq A8 to spatial averages, consistent with the approach of Sarkisov in ref [58]:

$$\frac{\Delta s_{\mathrm{ads}}^{\infty}}{k_{\mathrm{B}}} = \frac{\langle U_{\mathrm{fs}}(\mathbf{r}, \boldsymbol{\psi}) \exp\left[-\beta U_{\mathrm{fs}}(\mathbf{r}, \boldsymbol{\psi})\right] \rangle}{\langle \exp\left[-\beta U_{\mathrm{fs}}(\mathbf{r}, \boldsymbol{\psi})\right] \rangle} + \ln \langle \exp\left[-\beta U_{\mathrm{fs}}(\mathbf{r}, \boldsymbol{\psi})\right] \rangle \tag{A9}$$

where the $\langle\rangle$ brackets now indicate (not ensemble) averages over the positional and orientational degrees of freedom. Finally, using Sarkisov's expressions for the Henry's law constant and infinite dilution enthalpy of the adsorption of rigid molecules [58] in eq 2.3, the adsorption entropy reduces to the form shown in eq 2.5 (subject to appropriate conversions of energy quantities from molecular to molar basis).

## A.2 The free translator and harmonic oscillator approximations

For molecules that were strongly adsorbed, all modes were assumed to be vibrational with the entropy calculated through it's standard statistical mechanical expression:

$$S^0_{vib} = R\Sigma_i[\frac{h\nu_i}{k_BT(\exp(\frac{h\nu_i}{k_BT}-1))} - \ln(1-\exp(\frac{-h\nu_i}{k_BT}))] \tag{A10}$$

Weakly adsorbed systems, particularly those without a direct bond with the acid site, have been shown to exhibit some translational freedom under ambient conditions. As a consequence, the harmonic oscillator approximation would greatly underestimate the entropy of such a loosely bound state. It has been shown that a better approximation includes decoupling the adsorbate's modes into vibrations and two-dimensional free translations about an area commensurate to the pore of MFI (200 pm x 600 pm).[151, 152, 110, 145] This includes removing the two smallest wavenumbers from the vibrational entropy and adding a translational contribution to the entropy given by the following expression:

$$\begin{aligned} S^0_{trans} &= R[\ln[(\frac{2\pi M k_B T}{h^2})\frac{A^0}{N_A}] + 2] \\ S^0_{ads} &= S^0_{trans} + S^0_{vib} \end{aligned} \tag{A11}$$

These approximations are defined as "Free translator" and “Harmonic Oscillator” within the manuscript respectively.

The enthalpies were calculated by taking the sum of the DFT-calculated ground state electronic energy, the zero point vibrational energy, and temperature contributions:

$$\begin{aligned} \Delta H(T) &= \Delta E_{DFT} + \Delta E_{ZPVE} + \Delta H_{10\longrightarrow T} \\ \Delta E_{ZPVE} &= \frac{1}{2}\Sigma_{i=1}^{modes} h\nu_i \\ \Delta H_{10\longrightarrow T} &= \int_{10}^{T} C_P dT \end{aligned} \tag{A12}$$

The temperature corrections were obtained through numerical integration of the temperature dependent heat capacity from a reference temperature (10K). The heat capacity was quantified through fitting the Shomate equation to a set of entropy values computed along the incremental temperatures. Finally, the standard Gibbs Free Energy was derived from it's classical definition:

$$\Delta G^0(T) = \Delta H^0(T) - T\Delta S^0(T) \tag{A13}$$

## A.3 Thermodynamic approximation of the rate of adsorption/desorption

We have approximated the thermodynamics of the adsorption/desorption steps by constructing a series of pseudo-transition states which have the entropy of their corresponding adsorbate under the "Free Translator" approximation, and the enthalpy of their ideal-gas state. Our rationale is that these transition states should feel the effects of confinement (loss in partial translational and complete rotational entropy) but not yet be stabilized by the framework's van der Waals forces. Moreover, such an approximation will ensure that the adsorption/desorption steps are quasi-equilibrated; which is a common assumption. The thermodynamic quantities of these steps are mathematically defined as:

$$\begin{aligned}
\Delta S^{0,\ddagger}_{i,ads} &= S^0_{i^*,2D} - (S^0_{i,gas} + S^0_{H+}) \\
\Delta H^{0,\ddagger}_{i,ads} &= (H^0_{i,gas} - H^0_{i,gas}) = 0 \\
\Delta S^{0,\ddagger}_{i,des} &= S^0_{i^*,2D} - (S^0_{i^*}) \\
\Delta H^{0,\ddagger}_{i,des} &= (H^0_{i,gas} - H^0_{i^*})
\end{aligned} \tag{A14}$$

## A.4 Derivation of the microkinetic model

An upper and lower bound on each kinetic parameter was derived based on the thermodynamic approximations outlined within section A.2. The equilibrium constant for each step "i" was defined by the expression:

$$K_i(T) = \exp(-\frac{\Delta G_i^0}{RT}) = \exp(-\frac{\Delta H_i^0 - T\Delta S_i^0}{RT}) \tag{A15}$$

Rate coefficients for each elementary step were obtained using transition state theory and were calculated according to the expression:

$$k_i(T) = \frac{k_B T}{h} \exp(\frac{\Delta S_i^{0,\ddagger}}{R}) \exp(-\frac{\Delta H_i^{0,\ddagger}}{RT}) \tag{A16}$$

An isothermal-isobaric, ideal, continuously stirred tank reactor (CSTR) under differential conditions with ideal gas streams was used to model our reaction system:

$$\begin{aligned} \frac{dF_i}{dt} = F_{i,0} - F_i + \Omega_{sites} W(\Sigma_j^{steps^*} \nu_{ij} r_j) + V_{gas}(\Sigma_j^{steps} \nu_{ij} r_j) \\ \Sigma_j^{steps^*} \nu_{ij} r_j = \frac{d\Theta_i}{dt},\ where\ \Theta_i = \frac{N_i^*}{N_{sites}} \\ \Sigma_i \Theta_i + \Theta^* = 1 \\ p_i = p_{tot} \frac{F_i}{\Sigma_{i \in Gas} F_i} \end{aligned} \tag{A17}$$

Where $F_i$ represents the molar flow rate of species "i" $(mol/s)$. The generation/consumption terms were segregated between reactions occurring within the gas-phase, given by $\Sigma_j^{steps} \nu_{ij} r_j$, and within the catalyst, given by $\Sigma_j^{steps^*} \nu_{ij} r_j$. The index "j", represents the elementary steps which contain species "i" and each $\nu_{ij}$ represents their respective stoichiometric coefficient within that step. The reaction rates within the adsorbed phase were given in terms of fractional surface coverage

($mol/mol_{H+}$), which was represented by the variable $\Theta_i$ and constrained by the site balance. The molar quantity of total acid sites per gram of catalyst was defined by $\Omega_{sites}$ ($mol_{H+}/g-cat$) and represented a unit cell containing 47 Si/Al. The weight of catalyst was defined by $W$ ($g-cat$). The volume of the gas phase, $V_g$, was set equal to the volume occupied by the catalyst mass. The isobaric condition of the reactor was satisfied by including an inert, which ensured that the sum of ideal-gas partial pressures of the reactants and products and inert was 1 ($atm$). Equation A17 was written for each species, which produced a system of differential equations that have been solved simultaneously using an in-house python code. The volumetric flow rate of our reactor was 1800.0 ($mL/hr$), with a catalyst mass of 0.1 ($g$).

The conversion ($\xi$), selectivity ($S$), and yield ($Y$) are defined by equation A18, where $S(D/R)$ and $Y(D/R)$ represent the selectivity and yield of a product "D" with respect to the consumption of a reactant "R". Mathematically, the selectivity must sum to unity; and the reactant "R" was chosen as isoprene unless specified otherwise.

$$\begin{aligned} \xi_i &= \frac{F_{i,0} - F_i}{F_{i,0}} \\ S(D/R) &= \frac{-\nu_R F_D}{\nu_D \xi_R F_{R,0}} \\ \Sigma_{i=1}^{rxns} S(D_i/R) &= 1 \\ Y(D/R) &= \xi_R S(D/R) \end{aligned} \tag{A18}$$

The apparent order "$n_i$" is defined by equation A19. The MKM was run across five partial pressures, corresponding to multiples of 0.95, 0.98, 1.0, 1.02, 1.05 of the nominal pressure (1 atm). The corresponding $ln(r_i)$ were then fit to the pressures using least squares regression. The apparent order was then calculated by taking the product of the slope of the fit with the nominal pressure of 1 atm.

$$n_i = p_i \frac{\partial ln(r_i)}{\partial p_i} \tag{A19}$$

The activation energy "$\Delta E_{act}$" is defined by equation A20, where "R" and "T" represent the universal gas constant and temperature respectively. The reaction temperature was varied at ± 25 K about the nominal condition, fitting the corresponding $ln(r_i)$ to the temperature using least squares regression. The activation energy was then calculated as the product between the slope of the fit with $RT^2$, where the temperature was our nominal temperature of 368.15 K. The reaction was the consumption of our limiting reactant, isoprene ($C_5H_8$).

$$\Delta E_{act} = RT^2 \frac{\partial ln(r_{C_5H_8})}{\partial T} \tag{A20}$$

## A.5 Results of the microkinetic model

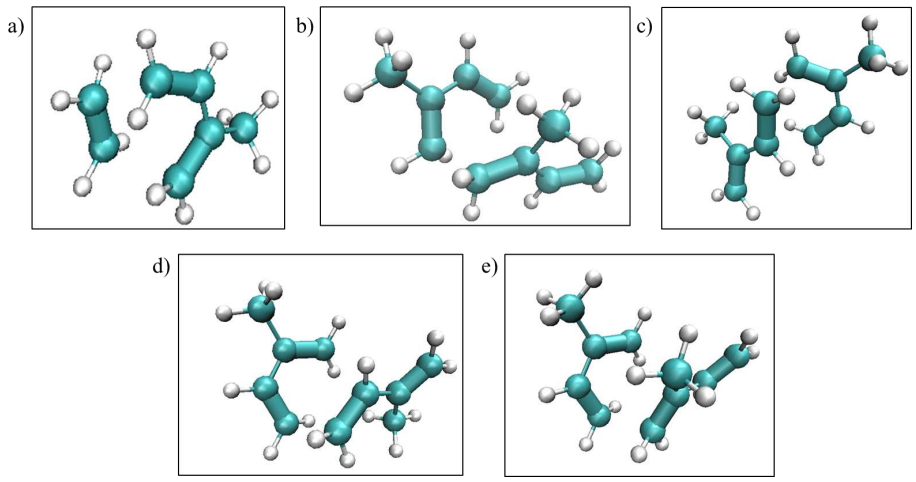


Figure A1: The most energetically stable unadsorbed configurations of our Diels-Alder transition states (TST). a) Product C7 TST, b) Product C10-para2 TST, c) Product C10-para1 TST, d) Product C10-meta1 TST, e) Product C10-meta2 TST.

| | step | $\Delta H_r^0$ | $\Delta S_r^0$ | $Ea_f$ | $k_f$ | $K_{eq}$ |
|---|---|---|---|---|---|---|
| $r_0$ | (0)-(1) | 10\10\10 | 1\1\1 | 58\58\58 | 5E+04\5E+04\5E+04 | 4E-02\4E-02\4E-02 |
| $r_1$ | (2)-(3) | -21\-19\-19 | -31\-66\-66 | 101\116\116 | 4E-02\3E-04\3E-04 | 2E+01\2E-01\2E-01 |
| $r_2$ | (5)-(4) | -55\-53\-53 | -28\-62\-62 | 64\79\79 | 6E+03\5E+01\5E+01 | 2E+06\2E+04\2E+04 |
| $r_3$ | (6)-(7) | 15\15\15 | 1\1\1 | 21\21\21 | 9E+09\9E+09\9E+09 | 7E-03\7E-03\7E-03 |
| $r_4$ | (7)-(11) | -18\-16\-16 | -20\-62\-62 | 12\12\12 | 1E+11\1E+11\1E+11 | 3E+01\1E-01\1E-01 |
| $r_5$ | (7)-(10) | 0\2\2 | -23\-65\-65 | 26\26\26 | 1E+09\1E+09\1E+09 | 7E-02\2E-04\2E-04 |
| $r_6$ | (8)-(7) | 35\35\35 | 2\2\2 | 39\39\39 | 2E+07\2E+07\2E+07 | 2E-05\2E-05\2E-05 |
| $r_7$ | (3)-(4) | -18\-18\-18 | 2\2\2 | 129\115\115 | 3E-06\4E-04\4E-04 | 4E+02\4E+02\4E+02 |
| $r_8$ | (8)-(12) | 14\14\14 | 3\3\3 | 30\30\30 | 4E+08\4E+08\4E+08 | 2E-02\2E-02\2E-02 |
| $r_9$ | (6)-(9) | 11\11\11 | 11\11\11 | 34\34\34 | 1E+08\1E+08\1E+08 | 1E-01\1E-01\1E-01 |
| $r_{10}$ | (8)-(9) | 30\30\30 | 12\12\12 | 35\35\35 | 7E+07\7E+07\7E+07 | 2E-04\2E-04\2E-04 |
| $r_{11}$ | (12+0g)-(13) | -220\-218\-220 | -207\-249\-204 | 80\79\79 | 4E+01\5E+01\5E+01 | 3E+20\9E+17\4E+20 |
| $r_{12}$ | (12+1g)-(14) | -212\-210\-212 | -248\-290\-242 | 91\89\89 | 8E-01\2E+00\2E+00 | 1E+17\4E+14\3E+17 |
| $r_{13}$ | (12+1g)-(15) | -215\-213\-215 | -246\-288\-240 | 90\88\88 | 1E+00\2E+00\2E+00 | 5E+17\2E+15\1E+18 |
| $r_{14}$ | (12+1g)-(16) | -208\-206\-208 | -248\-290\-243 | 83\81\81 | 1E+01\2E+01\2E+01 | 3E+16\1E+14\7E+16 |
| $r_{15}$ | (12+1g)-(17) | -207\-205\-207 | -246\-288\-240 | 81\79\79 | 2E+01\4E+01\4E+01 | 3E+16\1E+14\6E+16 |
| $r_{16}$ | (0g+7g)-(2g) | -197\-197\-197 | -196\-196\-196 | 88\88\88 | 9E+04\9E+04\9E+04 | 6E+17\6E+17\6E+17 |
| $r_{17}$ | (1g+7g)-(3g) | -172\-172\-172 | -216\-216\-216 | 110\110\110 | 7E+01\7E+01\7E+01 | 1E+13\1E+13\1E+13 |
| $r_{18}$ | (1g+7g)-(5g) | -164\-164\-164 | -221\-221\-221 | 115\115\115 | 1E+01\1E+01\1E+01 | 6E+11\6E+11\6E+11 |
| $r_{19}$ | (1g+7g)-(4g) | -172\-172\-172 | -215\-215\-215 | 115\115\115 | 1E+01\1E+01\1E+01 | 1E+13\1E+13\1E+13 |
| $r_{20}$ | (1g+7g)-(6g) | -167\-167\-167 | -221\-221\-221 | 115\115\115 | 1E+01\1E+01\1E+01 | 2E+12\2E+12\2E+12 |
| $r_{21}$ | (1g)-(7g) | 16\16\16 | 5\5\5 | 31\31\31 | 1E+13\1E+13\1E+13 | 9E-03\9E-03\9E-03 |
| $r_{22}$ | (0+0g)-(2) | -57\-59\-59 | -141\-106\-106 | 39\39\39 | 2E+07\2E+07\2E+07 | 5E+00\7E+02\7E+02 |
| $r_{23}$ | (1+0g)-(5) | -52\-54\-54 | -144\-109\-109 | 40\40\40 | 2E+07\2E+07\2E+07 | 7E-01\8E+01\8E+01 |
| $r_{24}$ | (0+1g)-(8) | -95\-98\-98 | -188\-147\-147 | 54\54\54 | 2E+05\2E+05\2E+05 | 5E+03\1E+06\1E+06 |
| $r_{25}$ | (1+1g)-(6) | -86\-89\-89 | -188\-147\-147 | 54\54\54 | 2E+05\2E+05\2E+05 | 3E+02\8E+04\8E+04 |
| $r_{26}$ | (0+7g)-(12) | -98\-100\-100 | -190\-148\-148 | 55\55\55 | 1E+05\1E+05\1E+05 | 9E+03\3E+06\3E+06 |
| $r_{27}$ | (0+2g)-(13) | -121\-121\-123 | -201\-201\-157 | 58\58\58 | 5E+04\5E+04\5E+04 | 4E+06\4E+06\2E+09 |
| $r_{28}$ | (0+3g)-(14) | -138\-138\-140 | -222\-222\-175 | 64\64\64 | 6E+03\6E+03\6E+03 | 8E+07\8E+07\5E+10 |
| $r_{29}$ | (0+5g)-(15) | -149\-149\-151 | -216\-216\-168 | 62\62\62 | 1E+04\1E+04\1E+04 | 7E+09\7E+09\4E+12 |
| $r_{30}$ | (0+4g)-(16) | -134\-134\-136 | -224\-224\-177 | 65\65\65 | 5E+03\5E+03\5E+03 | 2E+07\2E+07\1E+10 |
| $r_{31}$ | (0+6g)-(17) | -137\-137\-139 | -216\-216\-168 | 62\62\62 | 1E+04\1E+04\1E+04 | 2E+08\2E+08\1E+11 |

Table A1: Standard reaction enthalpy ($kJ/mol$), reaction entropy ($J/mol/K$), activation energy ($kJ/mol$), foreword rate constant ($s^{-1}$) and equilibrium constant ($dimensionless$) for each elementary reaction step outlined within Figure A2 at 368.15 ($K$). The lower and upper bounds are given by the HO and Free Translator approximations defined in section A.2, including an additional approximation where the adsorbate entropies are estimated using the Free Translator with the exception of the adsorbed cycloadducts, which are approximated by the HO. These three approximations are ordered: HO/(Free Translator & HO)/Free Translator within the table. Gas phase DA reactions were also included within the MKM; and are represented by $r_{16-21}$, which correspond to C7, C10-para1, C10-para2, C10-meta1, C10-meta2 product formation and *trans*-isoprene isomerization to *cis*-isoprene.

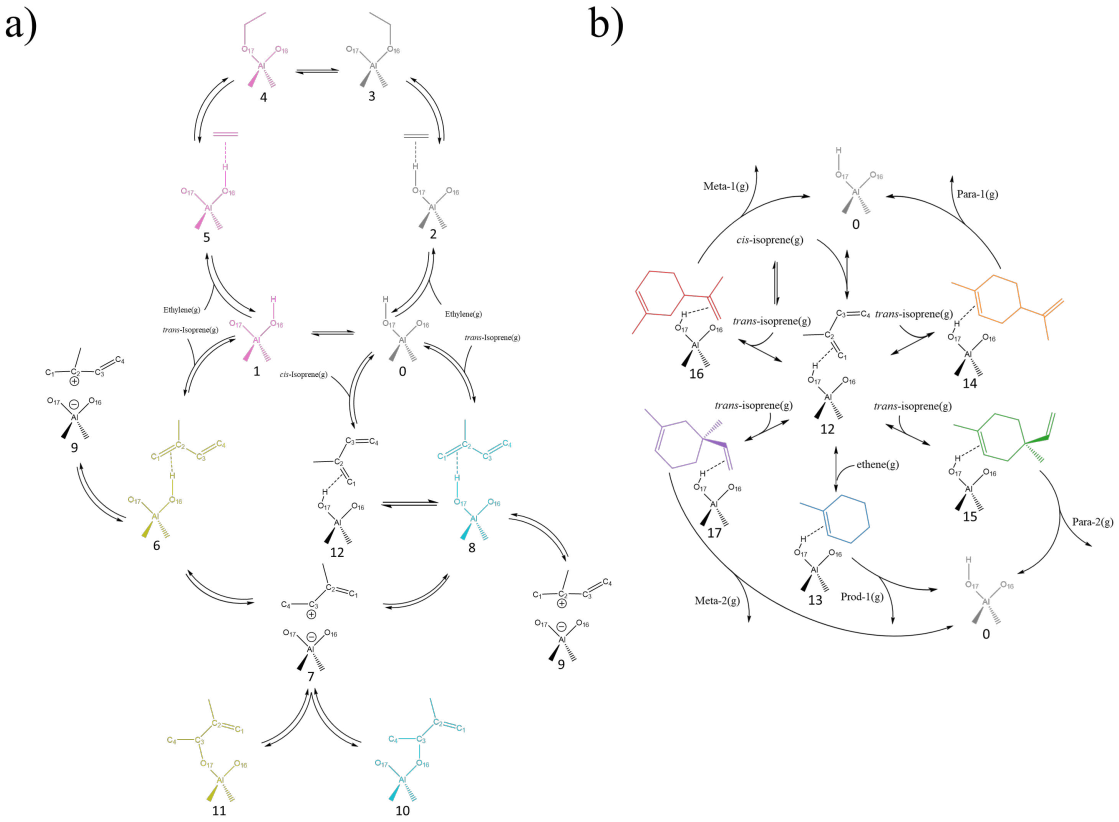


Figure A2: The reaction network, consisting of all catalytic elementary steps within the MKM with the coloring scheme and index for each intermediate having been preserved from the Gibbs free energy surfaces described in the manuscript. Each step is assumed elementary; and is categorized between: a) competitive physisorpion/chemisorption network b) DA cycloaddition network. The corresponding enthalpy, entropy, activation energy, forward rate constant, and equilibrium constants for each step are given in table A1. Gas phase DA reactions were also included within the MKM, but were not listed here.

| adsorbate | Conversion % | Yield % | Selectivity % | Fractional Coverage |
|---|---|---|---|---|
| ethene | 0.2/0.2/0.01 | - | - | 6.8E-2/3.4E-4/5.7E-7 |
| isoprene | 1.1/1.3/0.04 | - | - | 1.6E-3/1.1E-3/1.8E-6 |
| C7 | - | 0.8/0.8/0.03 | 69.9/62.3/97.7 | 2.0E-2/8.9E-3/6.2E-3 |
| C10-para1 | - | 0.01/0.01/<0.001 | 0.7/0.9/0.09 | 9.5E-3/1.2E-2/1.2E-2 |
| C10-para2 | - | 0.1/0.2/<0.001 | 10.3/12.9/0.79 | 6.0E-1/6.5E-1/6.6E-1 |
| C10-meta1 | - | 0.01/0.02/<0.001 | 1.0/1.3/0.08 | 4.1E-2/5.0E-2/5.1E-2 |
| C10-meta2 | - | 0.2/0.3/<0.001 | 18/23/1.4 | 2.6E-1/2.7E-1/2.8E-1 |

Table A2: The conversion, selectivity, and coverage per adsorbate quantified from the MKM under the entropic approximations discussed within A.2, including an additional approximation where the adsorbate entropies are estimated using the Free Translator with the exception of the adsorbed cycloadducts, which are approximated by the HO. These three approximations are ordered: HO/(Free Translator & HO)/Free Translator within the table. The isoprene conversion included both isomers (*cis*− and *trans*−). The coverage for ethene and isoprene included their physisorbed, chemisorbed, and carbenium states on O17 and O16, including both isomers of isoprene (*cis*− and *trans*−). The conversion, yield, and selectivity between 1) the Free Translator & HO and 2) the HO, were much more similar relative to the Free Translator approximation. The coverage under the Free Translator & HO was more similar to the Free Translator approximation for the cycloadducts, but more similar to the HO for the reactants.

| | step | $r_{foreward}$ | $r_{reverse}$ | $r_{net}$ |
|---|---|---|---|---|
| $r_0$ | (0)-(1) | 6E+04\3E+02\5E-01 | 6E+04\3E+02\5E-01 | 7E-01\7E-01\1E-03 |
| $r_1$ | (2)-(3) | 3E-02\2E-04\3E-07 | 3E-02\2E-04\3E-07 | 7E-10\7E-08\1E-10 |
| $r_2$ | (5)-(4) | 3E+01\1E-01\2E-04 | 3E+01\1E-01\2E-04 | -7E-10\-7E-08\-1E-10 |
| $r_3$ | (6)-(7) | 4E+09\3E+09\4E+06 | 4E+09\3E+09\4E+06 | 7E-01\7E-01\1E-03 |
| $r_4$ | (7)-(11) | 5E+08\3E+08\5E+05 | 5E+08\3E+08\5E+05 | 2E-07\-6E-08\-1E-08 |
| $r_5$ | (7)-(10) | 5E+06\3E+06\6E+03 | 5E+06\3E+06\6E+03 | 0E+00\1E-09\-7E-11 |
| $r_6$ | (8)-(7) | 4E+09\3E+09\4E+06 | 4E+09\3E+09\4E+06 | -7E-01\-7E-01\-1E-03 |
| $r_7$ | (3)-(4) | 7E-05\4E-05\7E-08 | 7E-05\4E-05\7E-08 | 7E-10\7E-08\1E-10 |
| $r_8$ | (8)-(12) | 8E+10\6E+10\1E+08 | 8E+10\6E+10\1E+08 | 2E+01\2E+01\3E-02 |
| $r_9$ | (6)-(9) | 6E+07\4E+07\7E+04 | 6E+07\4E+07\7E+04 | 2E-02\2E-02\3E-05 |
| $r_{10}$ | (8)-(9) | 1E+10\1E+10\2E+07 | 1E+10\1E+10\2E+07 | -2E-02\-2E-02\-4E-05 |
| $r_{11}$ | (12+0g)-(13) | 1E+01\1E+01\2E-02 | 3E-16\7E-14\1E-16 | 1E+01\1E+01\2E-02 |
| $r_{12}$ | (12+1g)-(14) | 7E-02\1E-01\2E-04 | 8E-15\5E-12\9E-15 | 7E-02\1E-01\2E-04 |
| $r_{13}$ | (12+1g)-(15) | 1E-01\1E-01\2E-04 | 2E-13\1E-10\2E-13 | 1E-01\1E-01\2E-04 |
| $r_{14}$ | (12+1g)-(16) | 1E+00\1E+00\2E-03 | 2E-12\1E-09\2E-12 | 1E+00\1E+00\2E-03 |
| $r_{15}$ | (12+1g)-(17) | 2E+00\3E+00\4E-03 | 2E-11\1E-08\2E-11 | 2E+00\3E+00\4E-03 |
| $r_{16}$ | (0g+7g)-(2g) | 6E-01\6E-01\6E-01 | 8E-18\8E-18\3E-19 | 6E-01\6E-01\6E-01 |
| $r_{17}$ | (1g+7g)-(3g) | 1E-04\1E-04\1E-04 | 1E-18\2E-18\5E-21 | 1E-04\1E-04\1E-04 |
| $r_{18}$ | (1g+7g)-(5g) | 2E-05\2E-05\2E-05 | 7E-18\1E-17\2E-20 | 2E-05\2E-05\2E-05 |
| $r_{19}$ | (1g+7g)-(4g) | 2E-05\2E-05\2E-05 | 3E-18\4E-18\7E-21 | 2E-05\2E-05\2E-05 |
| $r_{20}$ | (1g+7g)-(6g) | 2E-05\2E-05\2E-05 | 4E-17\6E-17\1E-19 | 2E-05\2E-05\2E-05 |
| $r_{21}$ | (1g)-(7g) | 6E+10\6E+10\7E+10 | 6E+10\6E+10\7E+10 | 5E-01\5E-01\6E-01 |
| $r_{22}$ | (0+0g)-(2) | 3E+06\2E+04\3E+01 | 3E+06\2E+04\3E+01 | 2E-09\7E-08\1E-10 |
| $r_{23}$ | (1+0g)-(5) | 1E+05\5E+02\8E-01 | 1E+05\5E+02\8E-01 | -8E-10\-7E-08\-1E-10 |
| $r_{24}$ | (0+1g)-(8) | 6E+03\3E+01\5E-02 | 6E+03\2E+01\3E-02 | 2E+01\2E+01\3E-02 |
| $r_{25}$ | (1+1g)-(6) | 3E+02\1E+00\2E-03 | 3E+02\6E-01\1E-03 | 7E-01\7E-01\1E-03 |
| $r_{26}$ | (0+7g)-(12) | 5E+01\2E-01\4E-04 | 5E+01\1E-01\2E-04 | 1E-01\1E-01\2E-04 |
| $r_{27}$ | (0+2g)-(13) | 2E+01\8E-02\6E-06 | 3E+01\1E+01\2E-02 | -1E+01\-1E+01\-2E-02 |
| $r_{28}$ | (0+3g)-(14) | 1E-02\6E-05\3E-10 | 8E-02\1E-01\2E-04 | -7E-02\-1E-01\-2E-04 |
| $r_{29}$ | (0+5g)-(15) | 3E-02\2E-04\6E-10 | 1E-01\1E-01\2E-04 | -1E-01\-1E-01\-2E-04 |
| $r_{30}$ | (0+4g)-(16) | 1E-01\7E-04\2E-09 | 1E+00\1E+00\2E-03 | -1E+00\-1E+00\-2E-03 |
| $r_{31}$ | (0+6g)-(17) | 5E-01\4E-03\1E-08 | 2E+00\3E+00\4E-03 | -2E+00\-3E+00\-4E-03 |

Table A3: The foreward, reverse, and net rate ($\mu$-mol/hr) from the MKM with respect to the numbering scheme in A2 under the entropic approximations discussed within A.2, including an additional approximation where the adsorbate entropies are estimated using the Free Translator with the exception of the adsorbed cycloadducts, which are approximated by the HO. These three approximations are ordered: HO/(Free Translator & HO)/Free Translator within the table. The net rates within the Free Translator & HO approximation more closely resemble the net rates under the HO approximation. Gas phase DA reactions were also included within the MKM; and are represented by $r_{16-21}$, which correspond to C7, C10-para1, C10-para2, C10-meta1, C10-meta2 product formation and *trans*-isoprene isomerization to *cis*-isoprene.

## A.6 List of zeolite frameworks and adsorbates

| | | | | | | |
|---|---|---|---|---|---|---|
| ABW | AWO | ESV | IWW | MOR | RHO | SSY |
| ACO | AWW | EUO | JBW | MRE | RSN | STI |
| AEI | BCT | EZT | JOZ | MSE | RWR | SZR |
| AEL | BEA | FAU | JRY | MTF | RWY | TER |
| AEN | BEC | FER | JST | MTN | SAF | THO |
| AET | BOF | GIS | JSW | MTT | SAO | TON |
| AFO | BOG | GON | KFI | MVY | SAS | TSC |
| AFR | BOZ | GOO | LOV | NAB | SAV | UEI |
| ANA | BSV | IHW | LTA | NON | SBE | UFI |
| APC | CAS | IMF | LTJ | NPT | SBN | UOS |
| APD | CFI | ISV | LTN | OBW | SFG | UOZ |
| AST | CGS | ITE | MEL | OSI | SFH | UWY |
| ASV | CON | ITH | MEP | OWE | SFV | VET |
| ATN | DFT | ITR | MER | PAU | SGT | VNI |
| ATS | DON | IWR | MFI | PHI | SIV | VSV |
| ATT | EDI | IWS | MFS | PON | SOD | WEI |
| ATV | EON | IWV | MON | PUN | SOS | ZON |

Table A4: The orthorhombic, pure silica zeolite framework codes according to The International Zeolite Database.

| Name | Category | SMILES |
|---|---|---|
| alcohols | propan-2-ol | CC(C)O |
| alcohols | propan-1-ol | CCCO |
| alcohols | pentane-1,5-diol | C(CCO)CCO |
| alcohols | octan-1-ol | CCCCCCCCO |
| alcohols | methanol | CO |
| alcohols | ethanol | CCO |
| alcohols | 1,2-ethanediol | C(CO)O |
| alcohols | pentan-1-ol | CCCCCO |
| alcohols | 2-butanol | CCC(C)O |
| alcohols | 1,3-propanediol | C(CO)CO |
| alcohols | 2-methoxyethan-1-ol | COCCO |
| alcohols | 2-methylpropan-2-ol | CC(C)(C)O |
| aldehydes | octanal | CCCCCCCC=O |
| aldehydes | pentanal | CCCCC=O |
| aldehydes | acetaldehyde | CC=O |
| alkanes | 2-methylpropane | CC(C)C |
| alkanes | 2,2-dimethylpropane | CC(C)(C)C |
| alkanes | n-decane | CCCCCCCCCC |
| alkanes | 3,4-dimethylhexane | CCC(C)C(C)CC |
| alkanes | n-butane | CCCC |
| alkanes | propane | CCC |
| alkanes | ethane | CC |
| alkanes | n-dodecane | CCCCCCCCCCCC |
| alkanes | n-octane | CCCCCCCC |
| alkanes | methane | C |
| alkanes | 2,5-dimethylhexane | CC(C)CCC(C)C |
| alkanes | 2,3-dimethylbutane | CC(C)C(C)C |
| alkanes | 2,2-dimethylhexane | CC(C)(C)CCCC |
| alkanes | n-pentane | CCCCC |
| alkenes | isoprene | C=C(C)C=C |
| alkenes | propene | C=CC |
| alkenes | 1-octene | C=CCCCCCC |
| alkenes | 1,3-butadiene | C=CC=C |
| alkenes | 1,5-hexadiene | C=CCCC=C |
| alkenes | 2-methylpropene | CC(C)=C |
| alkenes | 1-butene | C=CCC |
| alkenes | ethene | C=C |
| alkenes | cis-2-butene | C |
| alkenes | trans-2-butene | CC=CC |
| aromatics | p-xylene | CC1=CC=C(C=C1)C |
| aromatics | propylbenzene | CCCC1=CC=CC=C1 |
| aromatics | naphthalene | C1=CC=C2C=CC=CC2=C1 |
| aromatics | thiophene | C1=CSC=C1 |
| aromatics | isopropylbenzene | CC(C)C1=CC=CC=C1 |
| aromatics | ethylbenzene | CCC1=CC=CC=C1 |
| aromatics | m-xylene | CC1=CC(=CC=C1)C |
| aromatics | toluene-7 | CC1=CC=CC=C1 |
| cyclic-alkanes | cyclopentane | C1CCCC1 |
| cyclic-alkanes | cyclooctane | C1CCCCCCC1 |
| cyclic-alkanes | cyclohexane | C1CCCCC1 |
| cyclic-alkanes | cyclooctadecane | C1CCCCCCCCCCCCCCCCCC1 |
| cyclic-alkanes | cyclododecane | C1CCCCCCCCCC1 |
| cyclic-ethers | oxane | C1COCCC1 |
| cyclic-ethers | 1,4-dioxane | O1CCOCC1 |
| cyclic-ethers | tetrahydrofuran | C1COCC1 |
| cyclic-ethers | 1,3-dioxolane | O1COCC1 |
| esters | ethyl-methacrylate | CCOC(=O)C(=C)C |
| esters | ethyl-acrylate | CCOC(=O)C=C |
| esters | n-octyl-acrylate | CCCCCCCCOC(=O)C=C |
| esters | n-butyl-methacrylate | CCCCOC(=O)C(=C)C |
| esters | 2-ethylhexyl-acrylate | CCCCC(CC)COC(=O)C=C |
| esters | n-butyl-acrylate | CCCCOC(=O)C=C |
| esters | 2-hydroxyethyl-acrylate | C=CC(=O)OCCO |
| esters | methyl-methacrylate | CC(=C)C(=O)OC |
| ethers | 1,2-dimethoxyethane | COCCOC |
| ethers | ethyl-methyl-ether | CCOC |
| ethers | diethyl-ether | CCOCC |
| ethers | dipropyl-ether | CCCOCCC |
| ethers | dimethyl-ether | COC |
| ethers | methyl-tert-butyl-ether | CC(C)(C)OC |
| ethers | diisopropyl-ether | CC(C)OC(C)C |
| ketones | acetone | CC(=O)C |
| ketones | 2-pentanone | CCCC(=O)C |
| ketones | 2-octanone | CCCCCCC(=O)C |
| nitriles | acetonitrile | CC#N |
| nitriles | propionitrile | CCC#N |
| others | sarin | CC(C)OP(=O)(C)F |
| perfluoroalkanes | perfluoroethane | C(C(F)(F)F)(F)(F)F |
| perfluoroalkanes | n-perfluorodecane | C(C(C(C(C(C(F)(F)F)(F)F)(F)F)(F)F)(F)F)(C(C(C(C(F)(F)F)(F)F)(F)F)(F)F)(F)F)(F)F |
| perfluoroalkanes | n-perfluoropentane | C(C(C(F)(F)F)(F)F)(C(C(F)(F)F)(F)F)(F)F |
| perfluoroalkanes | n-perfluorohexane | C(C(C(C(F)(F)F)(F)F)(F)F)(C(C(F)(F)F)(F)F)(F)F |
| sulfides | diethyl-disulfide | CCSSCC |
| sulfides | dimethyl-sulfide | CSC |
| sulfides | diethyl-sulfide | CCSCC |
| sulfides | ethylmethyl-sulfide | CCSC |
| thiols | 2-methyl-1-propanethiol | CC(C)CS |
| thiols | 2-methyl-2-propanethiol | CC(C)(C)S |
| thiols | octanethiol | CCCCCCCCS |
| thiols | 2-butanethiol | CCC(C)S |
| thiols | methanethiol | CS |
| thiols | ethanethiol | CCS |
| thiols | pentanethiol | CCCCCS |

Table A5: Adsorbate species organized per functional category with corresponding SMILES string.

# Christopher Rzepa

Chemical Engineer · Data Scientist

*312 Iacocca Hall, 111 Research Dr., Bethlehem, PA, 18015*

(+1) 215-837-2104 | chr218@lehigh.edu | chr218 | chr218 | Christopher Rzepa

## Summary

PhD candidate in chemical engineering at Lehigh University who specializes in applying density functional theory, kinetic modeling, Monte-Carlo simulations, and data science to study catalysis within microporous frameworks. Highly motivated to solve challenging problems, develop new computational tools, and make a difference. Enjoys 3D modeling, animation, and game development in spare time.

## Work Experience

### Lehigh University

*Bethlehem, PA*

Research Assistant under Prof. Rangarajan, S.‡ and Prof. Mittal, J.† *Jan. 2018 - Aug. 2022*

- Developed micro-kinetic models of heterogeneous catalytic systems using parameters quantified by density functional theory simulations.‡
- Applied Monte-Carlo simulations to compute thermodynamic properties of adsorbate/framework systems to deduce relationships.‡ †
- Developed an artificial neural network (ANN) to identify key molecular/framework descriptors in predicting thermodynamics of adsorption.‡

### Kimberly Clark (Kleenex®)

*Neenah, WI.*

Product Development Engineer Intern *Sep. 2017 - Dec. 2017*

- Developed prototype bond patterns for "Daily Cleansing Cushion" to enhance detection of flawed product by "Cognex" vision system.
- Researched effects of fiber blends and linear densities on moisture retention for enhanced facial cleansing products.
- Promoted brand awareness by coordinating events with student interns across the nation.

### Kimberly Clark (Huggies®)

*Ogden, UT.*

Process Engineer Intern *Sep. 2016 - Dec. 2016*

- Conducted research in collaboration with quality assessment personnel in minimizing number of products to reject during machine start-up.
- Researched sources of major waste upon product change. Cause identified as erroneous inventory tracking due to glitch in scanning system.

## Publications

- **Rzepa, C.**, Siderius, W. D., Hatch, H. W., Shen, V. K., Rangarajan, S., and Mittal, J.,"*Computational Investigation of Correlations in Adsorbate Entropy for Pure-Silica Zeolite Adsorbents*", J. Physical Chemistry C 2020

- Yeh, Y., **Rzepa, C.**, Rangarajan, S., Gorte, R. J., "*Influence of brønsted-acid and cation-exchange sites on ethene adsorption in ZSM-5*", Microporous and Mesoporous Materials 2019

- Tian, H., **Rzepa, C.**, Upadhyay, R., Rangarajan, S., "*Estimating vibrational and thermodynamic properties of adsorbates with uncertainty using data driven surrogates.*", AIChE J. 2019

- **Rzepa, C.**, Rangarajan, S., "*DFT based microkinetic modeling of confinement driven [4+2] Diels-Alder reactions between ethene and isoprene in H-ZSM5*" (**RSC Cat. Sci. Tech.**)

- **Rzepa, C.**, Rangarajan, S., "*Elucidating entropy driven structure-property relations using neural networks*" (**In preparation**)

## Conference & Symposium Presentations

- **Rzepa, C.**, Rangarajan, S.,"*Developing data-driven structure-entropy relations for host-guest interactions in siliceous zeolites*", The Catalysis Club of Philadelphia's Annual Student Poster Session 2022 (Poster)
  Runner-up Best Poster Award

- **Rzepa, C.**, Hatch, H. W., Shen, V. K., Mittal, J., Siderius, W. D., Rangarajan, S. "*Developing data-driven structure-entropy relations for host-guest interactions in siliceous zeolites*", The 27th North American Catalysis Society Meeting (Talk)

- **Rzepa, C.**, Rangarajan, S.,"*An ab initio investigation of confinement driven [4+2] Diels-Alder reactions between ethene and isoprene in H-ZSM5*", The 27th North American Catalysis Society Meeting (Poster)

- **Rzepa, C.**, Rangarajan, S.,"*Inferring structure-entropy correlations of zeolite-adsorbate interactions using Monte-Carlo simulations and machine learning*", American Chemical Society Annual Meeting Spring 2022 (Talk)

- **Rzepa, C.**, Rangarajan, S.,"*Inferring structure-entropy correlations of zeolite-adsorbate interactions using Monte-Carlo simulations & machine learning*", American Chemical Society Annual Meeting Spring 2022 (Poster)
  Sci-Mix distinguished poster

- **Rzepa, C.**, Rangarajan, S.,"*A DFT and Microkinetic Modeling Study of Confinement Driven Diels-Alder Reactions in Acidic Zeolites*", AIChE Annual Meeting 2021 (Poster)

- **Rzepa, C.**, Rangarajan, S.,"*A DFT study of shape selective xylene production from H-ZSM-5 catalyzed Diels-Alder reactions*", The Catalysis Society of Metropolitan New York Annual Symposium 2021 (Virtual Flash-talk)
  Best Poster Award

- **Rzepa, C.**, Siderius, W. D., Hatch, H. W., Shen, V. K., Rangarajan, S., and Mittal, J.,"*Towards Data-Driven Structure-Property Relations for Predicting Adsorption Entropy in Siliceous Zeolites*", Catalysis Club of Philadelphia's Annual Student Poster Session 2020 (Virtual Flash-talk)
  2nd Place Best Poster Award

- **Rzepa, C.**, Siderius, Rangarajan, S.,"*Computational Quantification of Adsorption Driven Entropic Trends within Zeolites*", Northeast Corridor Zeolite Association Annual Meeting 2019 (Poster & Talk)
  Best Poster Award

- **Rzepa, C.**, Siderius, W. D., Hatch, H. W., Shen, V. K., Rangarajan, S., and Mittal, J.,"*Towards Data-Driven Structure-Property Relations for Predicting Adsorption Entropy in Siliceous Zeolites*", Virtual AIChE Annual Meeting 2020 (Virtual Flash-talk)

- **Rzepa, C.**, Siderius, W. D., Hatch, H. W., Shen, V. K., Rangarajan, S., and Mittal, J.,"*Towards Data-Driven Structure-Property Relations for Predicting Adsorption Entropy in Siliceous Zeolites*", Northeast Corridor Zeolite Association Annual Meeting 2020 (Virtual Flash-talk)

- **Rzepa, C.**, Siderius, W. D., Hatch, H. W., Shen, V. K., Rangarajan, S., and Mittal, J.,"*Quantification of the Thermodynamic Effects Regarding Alkane Adsorption within ZSM-5 Using Advanced Sampling Grand*

*Canonical Monte Carlo Simulations*", AIChE Annual Meeting 2019 (Talk)

## Education

**Lehigh University** — *Bethlehem, PA*

*Ph.D.* in Chemical Engineering, GPA: 3.93/4.00 — *Jan. 2018 - Aug. 2022*

- John C. Chen Endowed Fellow in Chemical Engineering

**Pennsylvania State University** — *University Park, PA*

*B.S.* Chemical Engineering, GPA: 3.49/4.00 — *Aug. 2012 - May. 2017*

- Minor in Mathematics

## Honors & Awards

| | | |
|---|---|---|
| 2022 | **Runner-up Best Poster Award,** The Catalysis Club of Philadelphia | *Villanova* |
| 2022 | **Richard J. Kokes Travel Award,** The 27th North American Catalysis Society Meeting | *New York* |
| 2022 | **CATL Graduate Student Travel Award,** American Chemical Society | *San Diego* |
| 2022 | **John C. Chen Endowed Fellow in Chemical Engineering,** for excellence in research | *Lehigh University* |
| 2021 | **Mohamed El-Aasser Travel Award,** for excellence in research | *Lehigh University* |
| 2021 | **Best Poster Award,** The Catalysis Society of Metropolitan New York | *Virtual Meeting* |
| 2020 | **Poster Award 2nd Place,** The Catalysis Club of Philadelphia | *Virtual Meeting* |
| 2019 | **Best Poster Award,** Northeast Corridor Zeolite Association | *UPenn* |
| 2018 | **Leonard A. Wenzel Award,** for best performance in Ph.D. qualifying exam | *Lehigh University* |

## Leadership

**Undergraduate Thermodynamics** — *Lehigh University*

Volunteer Recitation Instructor — *Jan. 2022 - May. 2022*

- Assisted Prof. Reichmanis, E. in leading recitation, holding office hours, and substituting lectures.

**Methods of Analysis in Chemical Engineering** — *Lehigh University*

Teacher's Assistant — *Aug. 2021 - Dec. 2021*

- Assisted Prof. Rangarajan, S. by helping students during the python based programming labs.

**Chemical Engineering Graduate Student Association** — *Lehigh University*

President — *Jun. 2020 - Jun. 2021*

- Coordinated with student officers to plan and host the first fully virtual Graduate Student Symposium.

**Graduate Kinetics and Reactor Design** — *Lehigh University*

Teacher's Assistant — *Aug. 2019 - Dec. 2019*

- Assisted Prof. Rangarajan, S. in grading student homework, holding office hours, and substituting lectures.

## Skills

| | |
|---|---|
| **Software** | Vienna Ab-Initio Simulation Package (VASP), Gaussian, The Free Energy and Advanced Sampling Simulation Toolkit (FEASST), Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS), Atomic Simulation Environment (ASE), TAMkin, CasADi, Openbabel, Avogadro, Vesta, VMD, ChemDraw, LaTeX, Microsoft Office |
| **Data Science** | Artificial Neural Networks (ANN), Principal Component Analysis (PCA), Regression Analysis(LASSO/Ridge/Kernel), Pytorch, scikit-learn, numpy, scipy, pandas, Matplotlib |
| **Programming** | Python, BASH, Cython, MATLAB, C++, High Performance Computing (HPC), SLURM, Parallel programming, Git |
| **Other** | Blender, Unity game engine, Web Scraping (Beautiful Soup), Fluent in Polish |