

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Ab Initio Methods for Modeling the Thermodynamics of Molecules Adsorbed in Zeolites

Permalink

<https://escholarship.org/uc/item/2k30v4kq>

Author

Bettinson, Lance

Publication Date

2022

Peer reviewed|Thesis/dissertation

Ab Initio Methods for Modeling the Thermodynamics of Molecules Adsorbed in Zeolites

By

Lance Bettinson

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemical Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Alexis T. Bell, Chair
Professor Martin Head-Gordon
Professor Kranthi K. Mandadapu

Spring 2022

**Ab Initio Methods for Modeling the Thermodynamics of Molecules
Adsorbed in Zeolites**

Thesis by
Lance Bettinson

University of California, Berkeley
Berkeley, California, USA

©2022

Abstract

Ab Initio Methods for Modeling the Thermodynamics of Molecules Adsorbed in Zeolites

by

Lance Bettinson

Professor Alexis T. Bell, Chair

Zeolites are crystalline microporous solids composed of corner-sharing, tetrahedrally-coordinated silicate (SiO_4) units. The isomorphic substitution of a framework Si atom by an Al atom is charge-compensated by a proton, introducing Brønsted-acidic bridging-hydroxy groups. These proton-exchanged zeolites are used in a large number of processes, including hydrocarbon cracking, isomerization, and alkylation, and conversion of petroleum to transportation fuel. In addition to catalytic processes, zeolites can be used as adsorbents for carbon capture, molecular sieving, and pollution control technologies. It is therefore of great interest to predict the impact of zeolite structure and composition on its functional properties to screen new catalysts and improve existing catalysts. Quantum chemical calculations can provide molecular-scale information on zeolite-adsorbate interactions, as well as model the energetic changes and dynamics of important reactions that occur within the channels and pores of zeolite catalysts. However, the application of quantum chemical calculations for the study of chemical reactions occurring in zeolites is made difficult by the lack of reliable methods to generalize the theory beyond zero Kelvin. Modeling adsorption and desorption free energies is particularly troublesome, relying on a subtle balance between enthalpic and entropic terms. While the enthalpic term is becoming ever more accurate through density functional development, the much more temperature-sensitive entropic term remains generally underquantified by frequently-assumed harmonic approximations. The consequence is an inability to computationally replicate experimental observations, such as rate coefficients or equilibrium constants, with chemical accuracy. This work is concerned with a re-examination of harmonic approximations, including quantifying its failures in reaction kinetics of aldol condensation on isolated metal sites, exploring alternative approximation methods for anharmonic intramolecular motions, and developing new methods for approximating anharmonic external molecular motions for species adsorbed in zeolites.

Contents

I.	Introduction	1
	Bibliographical Notes	6
II.	Computational Studies of Ketonization and Aldol Condensation Reactions over Site-Isolated Zirconium Catalysts	10
	1. Introduction	11
	2. Computational Methods	12
	3. Computational Results	13
	3.1 Analysis of Propanoic Acid Ketonization	13
	3.2 Analysis of Propan-2-one Condensation	20
	4. Conclusions	25
	5. Acknowledgements	26
	Bibliographical Notes	26
III.	Pitzer-Gwinn Approximations for Gas Phase Molecules with Significant Torsional Anharmonicity and Mode Coupling	29
	1. Introduction	30
	2. Theory and Methods	32
	3. Computational Details	40
	4. Results and Discussion	41
	5. Conclusions	48
	6. Acknowledgements	49
	Bibliographical Notes	49
IV.	Effects of Confinement on the Estimation of Enthalpies and Entropies of Small Molecule Adsorption in Zeolites	53
	1. Introduction	54
	2. Theory and Methods	55
	2.1 Molecular Model and Energy Calculations	57
	2.2 Adsorbate Rotations	57
	Parametrization of a Molecular Top	57
	Potential Energy Surface	60
	Partition Functions	64
	2.3 Translational Partition Functions	66
	2.4 Computational Procedure and Additional Details	67
	3. Results and Discussion	67

3.1 Adsorption Structures and ab Initio Electronic Energies	67
3.2 Thermal Corrections to the Electronic Energy	70
4. Conclusion	75
5. Future Directions	75
6. Acknowledgements	76
Bibliographical Notes	76
Appendices	82
A. Supplementary Information for Chapter II	83
1. Experimental Methods	83
1.1 Catalyst Synthesis	83
1.2 Catalyst Activity	83
1.3 Catalyst Characterization	84
2. Experimental Results	84
2.1 Catalyst Characterization	84
2.2 Catalytic Activity of Silica-Supported Zr Catalysts	85
3. Supporting Mathematics	87
3.1 The Energetic Span Model	87
3.2 Generalized Degrees of Rate Control and Apparent Activation Energy	89
3.3 Gibbs Activation Energy	90
3.4 Energetic Span Model, Generalized Degrees of Rate Control, and Apparent Activation Energy: Aldol Condensation	90
4. Supporting Figures	91
Bibliographical Notes	100
B. Supplementary Information for Chapter III	102
1. Thermodynamic Derivations	102
1.1 Quantum Harmonic Oscillator	103
1.2 Classical Harmonic Oscillator	103
1.3 Classical Uncoupled Mode Approximation	103
1.4 Classical Coupled Mode Approximation	104
2. No U-Turn Sampling	105
3. Data Visualization	107
Bibliographical Notes	111
C. Supplementary Information for Chapter IV	112
1. Molecular Rotations	112
1.1 Euler Angle Addendum	112
1.2 Variational Method for Non-linear Rotors	113
Kinetic Energy Matrix Elements	114
Potential Energy Matrix Elements	115
1.3 Thermodynamic Derivations	116
Quantum Harmonic Oscillator	117
Classical Harmonic Oscillator	117

Classical Rigid Rotor	117
Quantum Rigid Rotor	117
Pitzer-Gwinn Approximations	118
2. Sampling Grid Size Effects	118
3. Model Parameters	118
Bibliographical Notes	118

Chapter I

Introduction

Zeolites are crystalline microporous solids composed of corner-sharing, tetrahedrally coordinated silicate (SiO_4) units. The isomorphic substitution of a framework Si atom by an Al atom is charge-compensated by a proton, introducing Brønsted-acidic bridging-hydroxy groups. These materials are solid acid catalysts that are used in a large number of processes, including hydrocarbon cracking, isomerization, alkylation, and conversion of petroleum to transportation fuel [1, 2, 3, 4, 5]. In addition to catalytic processes, zeolites can be used as adsorbents for carbon capture, molecular sieving, and pollution control technologies [6, 7, 8, 9, 10, 11]. It is therefore of great interest to predict the impact of zeolite structure and composition on their functional properties to screen new catalysts and improve existing catalysts. Understanding the thermodynamics through a rigorous quantitative theory is increasingly realizable, with continuing advances in *ab initio* electronic structure methods [12] such as density functional theory (DFT) [13, 14, 15, 16, 17] permitting a closer examination of challenging chemical problems in heterogeneous catalysis. In the context of zeolites, quantum chemical calculations can provide molecular-scale information on zeolite-adsorbate interactions, as well as model the energetic changes and dynamics of important reactions that occur within the channels and pores of zeolite catalysts [18, 19, 20, 21, 22].

The application of these methods for the study of chemical reactions occurring on any catalyst, however, is made difficult by the lack of reliable and computationally feasible methods to generalize the theory beyond zero Kelvin. To do so involves computing the energy levels E_n of all relevant nuclear motions from the Schrödinger equation,

$$\hat{H} |\Phi_n\rangle = [\hat{T} + \hat{V}] |\Phi_n\rangle = E_n |\Phi_n\rangle \quad , \quad (1.1)$$

where \hat{H} is the Hamiltonian operator, $|\Phi_n\rangle$ is the n th eigenstate, \hat{T} is the kinetic energy operator of the relevant motion, and \hat{V} is the potential energy operator. The nuclear partition function is solved by summation over the Boltzmann-distributed energy levels,

$$Q^{qu} = \sum_n g_n e^{-\beta E_n} \quad , \quad (1.2)$$

where g_n is the degeneracy of the n th energy level, and the superscript *qu* denotes that the partition function is derived from a discrete sum. This is, in principle, a challenging problem because it requires modeling the full-dimensional potential energy surface (PES),

which is often computationally prohibitive except for extremely small systems. Path integral Monte Carlo (PIMC) and ab initio molecular dynamics (AIMD), for instance, incorporate the full-dimensional PES into the evaluation of thermodynamics, but can require millions of energy calculations to reach convergence [20, 23], and therefore quickly become unattractive for larger systems or more rigorous levels of electronic structure theory. Instead, for simple molecular systems, the problem is usually solved by assuming that vibrational modes behave like uncoupled harmonic oscillators (HOs). Thus all the energy levels as well as thermodynamic functions can be derived by a normal mode analysis, in which each vibrational mode’s potential energy is approximated as a quadratic potential whose curvature is determined by frequencies (ν_i) obtained from the second order derivatives of the electronic PES. The simplicity of this approximation is evident in the closed-form solution to the energy levels and partition function, which is shown in the following equations for a given mode i .

$$E_{i,n} = h\nu_i \left(n + \frac{1}{2} \right) \quad \forall n \in \mathbb{Z}^{\geq} \quad (1.3)$$

$$\begin{aligned} Q^{qu,HO} &= \sum_{n=0}^{\infty} e^{-\beta h\nu_i(n+\frac{1}{2})} \\ &= \frac{e^{-\beta h\nu_i/2}}{1 - e^{-\beta h\nu_i}} \end{aligned} \quad (1.4)$$

Here, h is the Planck constant, ν_i is the harmonic frequency of mode i , n is the quantum number corresponding to the HO energy levels.

Modeling adsorption and desorption free energies using the HO approximation poses its own set of problems, particularly when it comes to approximating adsorption entropies. The HO frequencies corresponding to these modes tend to be small in comparison to the frequencies of intramolecular modes, leading to nonphysical divergences in entropies derived from the HO partition function [24]. Furthermore, studies comparing experimental adsorption entropies for a wide range of molecules have consistently found as much as two-thirds of the gas phase entropy is retained from the gas phase, indicating significant retention of both translational and rotational motions [25, 26]. This is in contrast to harmonic approximations, which approximate translational and rotational degrees of freedom as effectively immobile relative to their gas phase counterparts. The result of using the HO approximation for all modes is a systematic underestimation of adsorption entropies.

Attempts have been made to correct errors inherent to the HO approximation through interpolation approaches [27], frequency down-scaling [28, 29], and anharmonic sampling along normal mode coordinates [19, 20, 30, 31, 32]. Perhaps the most widely used alternative to the HO approximation is the quasi-rigid rotor harmonic oscillator (q-RRHO) approximation proposed by Grimme [27], in which the entropies of low frequency normal modes are replaced by a semi-empirical interpolation between rigid rotor (RR) and harmonic oscillator (HO) entropies using the Head-Gordon damping function ϖ [33]:

$$S^{q-RRHO} = \sum_i [\varpi(\nu_i)S^{HO} + (1 - \varpi(\nu_i))S^{RR}] \quad (1.5)$$

$$\varpi(\nu_i) \equiv \frac{1}{1 + (100 \text{ cm}^{-1}/\nu_i)^4} \quad (1.6)$$

Investigation of the quasi-harmonic approach is examined in detail in Chapter II. This was done in the context of ketonization and aldol condensation reactions over site-isolated metal catalysts on a silica support. This effort involved modeling the stationary states of reaction intermediates on a silsesquioxane cluster model [34, 35], the reaction mechanism of which was informed by experiments. While the q-RRHO method accounts for rotational entropy of the adsorbate, we did not find this ad hoc approach to be particularly accurate. The degree to which the quasi-harmonic approximation cannot describe entropic contributions was quantified by a sensitivity analysis of adsorption entropies. Specifically, we computed the apparent activation energies, turnover frequencies, and partial pressure dependences for ketonization and aldol condensation reactions as a function of the adsorption entropy contribution to the free energy. Comparison of these quantities against experimental results affirmed a systematic underestimation of the adsorption entropy by quasi-harmonic approaches, with significant gas phase motion needing to be retained in order to obtain agreement with experiment.

An alternate approach for computing the adsorbed phase thermodynamics, pursued by Sauer and colleagues [19, 20, 31, 32], is to approximate an anharmonic (AnH) representation of the potential energy V by sampling electronic energies along linear displacements of the normal mode coordinates. The vibrational energy levels of the system can then be solved using basis functions through the variational method [36]. By this method, the wavefunction is approximated as a linear combination of orthonormal basis functions, chosen to be the eigenfunctions of the HO,

$$|\Phi\rangle = \sum_n c_n |n\rangle, \quad (1.7)$$

$$\psi_n(q) \equiv \langle q|n\rangle = \frac{1}{\sqrt{2^n n!}} \left(\frac{\mu\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{\mu\omega q^2}{2\hbar}} \mathcal{H}_n\left(\sqrt{\frac{\mu\omega}{\hbar}}q\right) \quad (1.8)$$

where \mathcal{H}_n is the n th Hermite polynomial, μ is the mass of the oscillator, and ω is its frequency. A matrix representation of the Hamiltonian is obtained by computing the matrix elements,

$$H_{mn} = \langle m|\hat{T} + \hat{V}_{\text{AnH}}|n\rangle. \quad (1.9)$$

Diagonalization of this matrix yields the energy levels, through which the quantum partition function can be computed by Equation 1.2.

The above approach accounts for local anharmonicity of the potential energy, but because it still requires sampling along local normal mode coordinates, it does not account for multiple conformers on the extended domain. This was demonstrated by previous work studying internal rotations (i.e. torsions) of gas phase molecules [30], a comparable problem to that of hindered rotation and translation of adsorbed species in that the potential energy surface is highly anharmonic and possibly contains multiple energetically-favorable configurations. In this work, the total anharmonic PES was assumed to be the sum of each torsional contribution (the uncoupled mode approximation for torsions, labeled UM-T). Each torsional PES was obtained by sampling along the dihedral angular coordinate (ϕ), and the energy

levels were obtained by the variational method using Fourier basis functions:

$$\psi_n(\phi) \equiv \langle \phi | n \rangle = \begin{cases} \frac{1}{\sqrt{2\pi}} & n = 0 \\ \frac{1}{\sqrt{\pi}} \cos\left(\frac{n+1}{2}\phi\right) & n \text{ odd} \\ \frac{1}{\sqrt{\pi}} \sin\left(\frac{n}{2}\phi\right) & n \text{ even} \end{cases}, \quad \forall n \in \mathbb{Z}^{\geq} \quad (1.10)$$

The UM-T method yielded superior results against the HO and anharmonic normal mode alternatives when compared against experimental values, particularly the gas phase entropy. This suggests that the usual normal mode analysis, even under a locally anharmonic potential, could benefit from incorporating non-local information obtained by sampling the entire range of the relevant motion.

The drawback to non-local uncoupled mode methods is that some modes, particularly hindered translations and hindered rotations, are realistically coupled to each other. Moreover, obtaining the desired mode-coupled partition function by the variational method is computationally prohibitive because requires not only sampling electronic energies of the full-dimensional PES, but also diagonalization of a sizeable Hamiltonian matrix [37]. Mode coupling and an approximation to the quantum mode-coupled partition function can be described by the Pitzer-Gwinn (PG) approximation, through which the classical partition function of an anharmonic, non-local motion is scaled by a prefactor given by the quotient of the quantum partition function of a known local reference potential and its classical counterpart [37, 38, 39, 40]. We represent this approximation as $\text{PG}_{\text{AnH}}^{\text{ref.}}$, where the superscript “ref.” indicates the chosen reference potential (e.g. the HO reference) for which the quantum and classical partition functions can be computed, and the subscript “AnH” indicates the target anharmonic, non-local potential that the reference potential approximates (e.g. the torsional or hindered rotational potential).

$$Q^{\text{PG}_{\text{AnH}}^{\text{ref.}}} = \frac{Q^{qu,\text{ref.}}}{Q^{cl,\text{ref.}}} Q^{cl,\text{AnH}} \quad (1.11)$$

As before, the superscript *qu* denotes that the partition function is derived from a sum of discrete elements, and *cl* denotes that the partition function is derived classically, i.e. by integration over canonical position and momentum variables.

In Chapter III, we demonstrate how PG approximations can be used to capture anharmonic mode-coupling effects of internal rotations, a high-dimensional problem for which solving the energy levels using a variational solution to the Schrödinger equation is computationally prohibitive. We present the $\text{PG}_{\text{CM-T}}^{\text{UM-T}}$ method, which uses the uncoupled torsional (UM-T) potential energy as the reference potential to approximate the torsional mode-coupled (CM-T) thermodynamics. The quantum UM-T partition function is computed using the variational method in the Fourier basis [30], and the classical mode-coupled partition function is solved by Monte Carlo integration. The Monte Carlo sampler we implemented in this work was a hybrid Monte Carlo algorithm called the No U-Turn Sampler (NUTS), which combines principles of AIMD and Markov chain Monte Carlo [41, 42, 43]. This involved drawing random dihedral angular momenta and solving their trajectories on the multidimensional PES by time-integration of Hamilton’s equations of motion. The thermodynamics derived from the partition functions of the $\text{PG}_{\text{CM-T}}^{\text{UM-T}}$ approximation are found to

outperform HO approximations, HO-reference Pitzer-Gwinn approximations ($\text{PG}_{\text{CM-T}}^{\text{HO}}$ and $\text{PG}_{\text{UM-T}}^{\text{HO}}$), and UM-T approximations when compared against experimental gas phase enthalpies, entropies, and heat capacities. Notably, we find the underapproximation of the entropy to be greatly improved by the anharmonic, non-local description of molecular torsions.

The results of the work of Chapter III introduce the possibility of using PG approximations for more challenging anharmonic multidimensional problems, including modes such as the hindered translation and rotation for adsorbed species. This motivated us to pursue a description of the rotational PES for adsorbed species under a rigid rotor (RR) approximation, which is the rotational analog to the UM-T method. The undertaking of this work is presented in Chapter IV, where we present a discrete sampling scheme for adsorbate orientations of non-linear molecules adsorbed in the zeolites H-MFI and H-CHA. The RR PES is obtained via a discrete, spherical Fourier transform using Wigner D-Matrix element basis functions:

$$V(\phi, \theta, \chi) = \sum_{\ell=0}^{L_{\max}} \sum_{m=-\ell}^{\ell} \sum_{k=-\ell}^{\ell} \hat{v}_{mk}^{\ell} D_{mk}^{\ell}(\phi, \theta, \chi) \quad (1.12)$$

$$D_{mk}^{\ell}(\phi, \theta, \chi) = e^{-im\phi} d_{mk}^{\ell}(\theta) e^{-ik\chi} \quad (1.13)$$

$$d_{mk}^{\ell}(\theta) = \sqrt{\frac{(\ell+k)!(\ell-k)!}{(\ell+m)!(\ell-m)!}} \left(\sin \frac{\theta}{2}\right)^{k-m} \left(\cos \frac{\theta}{2}\right)^{k+m} P_{\ell-k}^{(k-m, k+m)}(\cos \theta) \quad (1.14)$$

The result is a description of the rotational PES that can be used to (1) solve the discrete RR energy levels using a variational solution to the Schrödinger equation in the same basis,

$$\psi_{mk}^{\ell}(\phi, \theta, \chi) \equiv \langle \phi, \theta, \chi | \ell m k \rangle = \sqrt{\frac{2\ell+1}{8\pi^2}} D_{mk}^{\ell*}(\phi, \theta, \chi) \quad \forall \ell \in \mathbb{Z}^{\geq}, -\ell \leq m \leq \ell, -\ell \leq k \leq \ell \quad (1.15)$$

(2) obtain the HO-reference PG partition function for a RR (the $\text{PG}_{\text{RR}}^{\text{HO}}$ method), and (3) obtain the classical RR partition function. These results were subsequently paired with harmonic and non-local descriptions of translation, and their accuracies were then compared against the strictly local HO approximation.

The accuracy of the methods in Chapter IV is shown to depend on the strength of the adsorption, with lightly bound alkanes better represented under non-local and RR approximations when compared against experimental adsorption data, particularly adsorption entropies. Strong adsorptions, such as the adsorption of methanol and ethanol in H-MFI, are found to strongly favor harmonic translational approximations. RR and $\text{PG}_{\text{RR}}^{\text{HO}}$ methods are nevertheless found to be comparable to local harmonic approximations in the extreme of strong adsorption, speaking to their general applicability in describing hindered rotation. More strongly bound adsorbates are also found to potentially benefit from direct consideration of rotational degrees of freedom, with RR and $\text{PG}_{\text{RR}}^{\text{HO}}$ methods accounting for multiple stable adsorption conformers. Moreover, the externality of rotations and translations render these methods exceptionally scalable, with the number of samples needed to construct the PES independent of molecule size. This opens the possibility of applying higher levels of electronic structure theory for describing the rotational PES.

These results and the RR results obtained in this work are the foundation of future development to include the full-dimensional PES of translations and rotations. At the end of Chapter IV, we propose the next steps to realize this objective. Foremost, considerable improvements should be made to the anharmonic description of adsorbate translations, which yet lack the same rigor as the RR method. A general method for handling translational anharmonicity and non-locality in zeolites is proposed, which involves sampling adsorbate positions in three dimensions to obtain the uncoupled translational PES. This and the RR PES will serve as a computationally feasible reference potentials to apply in a PG scheme that interrogates translational and rotational mode coupling of adsorbed species in zeolites.

References

- [1] Kunhao Li, Julia Valla, and Javier Garcia-Martinez. “Realizing the Commercial Potential of Hierarchical Zeolites: New Opportunities in Catalytic Cracking”. *ChemCatChem* **6.1** (2014), 46–66
- [2] Vincent Blay et al. “Engineering Zeolites for Catalytic Cracking to Light Olefins”. *ACS Catal.* **7.10** (2017), 6542–6566
- [3] Avelino Corma. “Inorganic Solid Acids and Their Use in Acid-catalyzed Hydrocarbon Reactions”. *Chemical Reviews* **95.3** (1995), 559–614
- [4] W Vermeiren and J-P Gilson. “Impact of Zeolites on the Petroleum and Petrochemical Industry”. *Top. Catal.* **52.9** (2009), 1131–1161
- [5] Jihong Yu and Ruren Xu. “Rational Approaches Toward the Design and Synthesis of Zeolitic Inorganic Open-framework Materials”. *Acc. Chem. Res.* **43.9** (2010), 1195–1204
- [6] Ahmed I Osman et al. “Recent Advances in Carbon Capture Storage and Utilisation Technologies: a Review”. *Environ. Chem. Lett.* **19.2** (2021), 797–849
- [7] Cong Chao et al. “Post-Combustion Carbon Capture”. *Renewable and Sustainable Energy Reviews* **138** (2021), 110490. ISSN: 1364-0321. DOI: <https://doi.org/10.1016/j.rser.2020.110490>. URL: <https://www.sciencedirect.com/science/article/pii/S1364032120307760>
- [8] Helge Bux et al. “Zeolitic Imidazolate Framework Membrane with Molecular Sieving Properties by Microwave-Assisted Solvothermal Synthesis”. *J. Am. Chem. Soc.* **131.44** (2009), 16000–16001
- [9] Chen Zhang et al. “Unexpected Molecular Sieving Properties of Zeolitic Imidazolate Framework-8”. *J. Phys. Chem. Lett.* **3.16** (2012), 2130–2134
- [10] Sabéha Kesraoui-Ouki, Christopher R. Cheeseman, and Roger Perry. “Natural Zeolite Utilisation in Pollution Control: A Review of Applications to Metals’ Effluents”. *J. Chem. Technol. Biotechnol.* **59.2** (1994), 121–126. DOI: <https://doi.org/10.1002/jctb.280590202>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/jctb.280590202>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jctb.280590202>

- [11] Rolando Roque-Malherbe. “Applications of Natural Zeolites in Pollution Abatement and Industry”. Dec. 2001, pp. 495–522. DOI: [10.1016/B978-012513910-6/50069-4](https://doi.org/10.1016/B978-012513910-6/50069-4)
- [12] Evgeny Epifanovsky et al. “Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package”. *J. Chem. Phys.* **155.8** (2021), 084801
- [13] Stefan Grimme. “Semiempirical Hybrid Density Functional with Perturbative Second-Order Correlation”. *J. Chem. Phys.* **124.3** (2006), 034108
- [14] Urs Zimmerli, Michele Parrinello, and Petros Koumoutsakos. “Dispersion Corrections to Density Functionals for Water Aromatic Interactions”. *J. Chem. Phys.* **120.6** (2004), 2693–2699
- [15] János G Angyán et al. “Van der Waals Forces in Density Functional Theory: Perturbational Long-Range Electron-Interaction Corrections”. *Phys. Rev. A* **72.1** (2005), 012510
- [16] O Anatole Von Lilienfeld et al. “Optimization of Effective Atom Centered Potentials for London Dispersion Forces in Density Functional Theory”. *Phys. Rev. Lett.* **93.15** (2004), 153004
- [17] Alexandre Tkatchenko and Matthias Scheffler. “Accurate Molecular van der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data”. *Phys. Rev. Lett.* **102.7** (2009), 073005
- [18] Fabian Berger, Marcin Rybicki, and Joachim Sauer. “Adsorption and Cracking of Propane by Zeolites of Different Pore Size”. *J. Catal.* **395** (2021), 117–128
- [19] GiovanniMaria Piccini, Maristella Alessio, and Joachim Sauer. “Ab Initio Study of Methanol and Ethanol Adsorption on Brønsted Sites in Zeolite H-MFI”. *Phys. Chem. Chem. Phys.* **20.30** (2018), 19964–19970
- [20] GiovanniMaria Piccini et al. “Accurate Adsorption Thermodynamics of Small Alkanes in Zeolites. Ab Initio Theory and Experiment for H-Chabazite”. *J. Phys. Chem. C* **119.11** (2015), 6128–6137
- [21] Florian Göltl et al. “Van der Waals interactions between hydrocarbon molecules and zeolites: Periodic calculations at different levels of theory, from density functional theory to the random phase approximation and Møller-Plesset perturbation theory”. *J. Chem. Phys.* **137.11** (2012), 114111
- [22] Yi-Pei Li, Martin Head-Gordon, and Alexis T Bell. “Computational Study of p-Xylene Synthesis from Ethylene and 2, 5-Dimethylfuran Catalyzed by H-BEA”. *J. Phys. Chem. C* **118.38** (2014), 22090–22095
- [23] Berend Smit and Theo LM Maesen. “Molecular Simulations of Zeolites: Adsorption, Diffusion, and Shape Selectivity”. *Chemical Reviews* **108.10** (2008), 4125–4184
- [24] Christopher J Cramer. *Essentials of Computational Chemistry: Theories and Models*. John Wiley & Sons, 2013
- [25] Charles T Campbell and Jason RV Sellers. “The Entropies of Adsorbed Molecules”. *J. Am. Chem. Soc.* **134.43** (2012), 18109–18115

- [26] Paul J Dauenhauer and Omar A Abdelrahman. “A Universal Descriptor for the Entropy of Adsorbed Molecules in Confined Spaces”. *ACS Central Science* **4.9** (2018), 1235–1243
- [27] Stefan Grimme. “Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory”. *Chem. Eur. J.* **18.32** (2012), 9955–9964
- [28] Anthony P Scott and Leo Radom. “Harmonic vibrational frequencies: an evaluation of Hartree-Fock, Møller-Plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors”. *J. Phys. Chem.* **100.41** (1996), 16502–16513
- [29] Mattias P Andersson and Per Uvdal. “New scale factors for harmonic vibrational frequencies using the B3LYP density functional method with the triple- ζ basis set 6-311+G (d, p)”. *J. Phys. Chem. A* **109.12** (2005), 2937–2941
- [30] Yi-Pei Li, Alexis T Bell, and Martin Head-Gordon. “Thermodynamics of Anharmonic Systems: Uncoupled Mode Approximations for Molecules”. *J. Chem. Theory Comput.* **12.6** (2016), 2861–2870
- [31] GiovanniMaria Piccini and Joachim Sauer. “Quantum Chemical Free Energies: Structure Optimization and Vibrational Frequencies in Normal Modes”. *J. Chem. Theory Comput.* **9.11** (2013), 5038–5045
- [32] GiovanniMaria Piccini and Joachim Sauer. “Effect of Anharmonicity on Adsorption Thermodynamics”. *J. Chem. Theory Comput.* **10.6** (2014), 2479–2487
- [33] Jeng-Da Chai and Martin Head-Gordon. “Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections”. *Phys. Chem. Chem. Phys.* **10.44** (2008), 6615–6620
- [34] Elsje Alessandra Quadrelli and Jean-Marie Basset. “On Silsesquioxanes’ Accuracy as Molecular Models for Silica-grafted Complexes in Heterogeneous Catalysis”. *Coordination Chemistry Reviews* **254.5-6** (2010), 707–728
- [35] David G Hanna et al. “Experimental and Theoretical Study of n-Butanal Self-condensation over Ti Species Supported on Silica”. *ACS catalysis* **4.9** (2014), 2908–2916
- [36] Attila Szabo and Neil S. Ostlund. *Modern Quantum Chemistry : Introduction to Advanced Electronic Structure Theory*. Mineola, N.Y: Dover Publications, 1996. ISBN: 0486691861
- [37] Luis Simón-Carballido et al. “Anharmonicity of Coupled Torsions: the Extended Two-dimensional Torsion Method and its Use to Assess More Approximate Methods”. *J. Chem. Theory Comput.* **13.8** (2017), 3478–3492
- [38] Kenneth S Pitzer and William D Gwinn. “Thermodynamic Functions for Molecules with Internal Rotation”. *J. Chem. Phys.* **9.6** (June 1941), 485–486. ISSN: 0021-9606. DOI: [10.1063/1.1750941](https://doi.org/10.1063/1.1750941). URL: <https://doi.org/10.1063/1.1750941>
- [39] Benjamin A Ellingson et al. “Statistical Thermodynamics of Bond Torsional Modes: Tests of Separable, Almost-Separable, and Improved Pitzer–Gwinn Approximations”. *J. Chem. Phys.* **125.8** (Aug. 2006), 84305. ISSN: 0021-9606. DOI: [10.1063/1.2219441](https://doi.org/10.1063/1.2219441). URL: <https://doi.org/10.1063/1.2219441>

- [40] Vanessa Audette Lynch, Steven L Mielke, and Donald G Truhlar. “High-Precision Quantum Thermochemistry on Nonquasiharmonic Potentials: Converged Path-Integral Free Energies and a Systematically Convergent Family of Generalized PitzerGwinn Approximations”. *J. Phys. Chem. A* **109**.44 (Nov. 2005), 10092–10099. ISSN: 1089-5639. DOI: [10.1021/jp051742n](https://doi.org/10.1021/jp051742n). URL: <https://doi.org/10.1021/jp051742n>
- [41] Matthew D Hoffman, Andrew Gelman, et al. “The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo.” *J. Mach. Learn. Res.* **15**.1 (2014), 1593–1623
- [42] John Salvatier, Thomas V Wiecki, and Christopher Fonnesbeck. “Probabilistic Programming in Python using PyMC3”. *PeerJ Computer Science* **2** (2016), e55
- [43] Andrew Gelman, Daniel Lee, and Jiqiang Guo. “Stan: A probabilistic programming language for Bayesian inference and optimization”. *J. Educ. Behav. Stat.* **40**.5 (2015), 530–543

Chapter II

Computational Studies of Ketonization and Aldol Condensation Reactions over Site-Isolated Zirconium Catalysts

THE REACTION KINETICS of aldol condensation and ketonic decarboxylation (ketonization) are investigated for reactions over isolated Zr centers supported on a high surface area silica. These reactions were modeled by a Zr cluster model using density functional theory, after which the thermodynamics, computed via harmonic approximations, were compared against experimental data. Catalysts were synthesized by the grafting of Cp_2ZrCl_2 on the surface of amorphous silica. The connectivity of Zr was characterized by XRD, UV-vis, and Raman spectroscopy. For the lowest Zr loading, Zr is present predominantly as isolated monomeric species. As the Zr loading is increased, a progressively larger fraction of Zr forms oligomeric species and ZrO_2 nanoparticles. Measurements of catalytic activity show that the turnover frequency for carboxylic acid ketonic decarboxylation reaction and aldol condensation of ketones decreases monotonically with increasing Zr loading. An H/D kinetic isotope effect was not observed over isolated Zr catalysts, suggesting that α -H abstraction is not the rate-determining step, rather C–C bond forming may be rate limiting for both reactions. This conclusion is supported by computational modeling of the reaction mechanism. The proposed catalytic cycle for ketonization proceeds via a β -keto acid intermediate on isolated Zr sites that are always coordinatively saturated with C–C bond formation as the rate-limiting step. C–C bond formation is also rate-determining for aldol condensation, with an apparent activation energy that is in good agreement with the experiment if the resting state is a saturated $\equiv\text{ZrOH}$ site with two adsorbed ketone molecules. The extent to which the resting state is malapproximated by the quasi-rigid rotor-harmonic oscillator (qRRHO) approximation are probed by a sensitivity analysis of turnover frequencies computed by the energetic span model. The results of this study affirm the failure of harmonic approximations to describe the entropies of species adsorbed from the gas phase.

1. Introduction

Biomass and biomass-derived products are an attractive source of renewable carbon for producing chemicals and fuels [1, 2]. A significant drawback to using these feedstocks is their high content of oxygen. For example, bio-oil produced by fast pyrolysis of biomass contains a high fraction of acids, alcohols, aldehydes, esters, phenolics, and other oxygenates [3, 4]. Of these compounds, the largest fraction comprises low molecular weight carboxylic acids, aldehydes, and ketones (R-COOH, R-CHO, R-CO-R). Carboxylic acids can also be produced by fermentation of sugars [5]. Because of their low molecular weight and high oxygen content, these compounds cannot be used as fuels without being upgraded to products containing more carbon atoms and fewer oxygen atoms [6, 7, 8, 9, 10].

An attractive approach for removing oxygen from biomass-derived carboxylic acids is ketonic decarboxylation (ketonization) [11, 12]. This reaction condenses two carboxylic acid molecules, eliminating 75% of the oxygen in the reactants to produce a linear ketone with $2n-1$ carbon atoms, CO_2 , and H_2O [11]. The alkanone produced by ketonic decarboxylation can undergo subsequent aldol condensation, thereby further increasing the chain length of the product and its energy density via the removal of oxygen as water [13]. While ketonic decarboxylation is conventionally promoted by homogeneous base catalysts, such as NaOH or KOH, catalyst separation from the reaction mixture is difficult and catalyst disposal is expensive [14].

Previous studies have shown that bulk and dispersed metal oxides are promising heterogeneous catalysts for ketonic decarboxylation and aldol condensation because they contain Lewis acid-Brønsted base pairs on their surface [11]. Oxides with a high lattice energy, such as TiO_2 or ZrO_2 , are particularly active and selective for C-C bond formation because carboxylic acids or aldehydes can adsorb at Lewis-acidic Ti or Zr sites. The resulting adsorbate can then undergo abstraction of the α -proton at an adjacent Brønsted-basic oxygen sites [9, 15, 16, 17, 18, 19, 20].

The purpose of this study was to elucidate the role of local coordination and connectivity of supported Zr sites on the gas-phase, ketonic decarboxylation of carboxylic acids and aldol condensation of ketones. Catalysts were prepared by impregnation of the support with a metal alkoxide precursor and then characterized by XRD, Raman and UV-Vis spectroscopy. Since we were particularly interested in investigating the properties of isolated Zr structures, well-defined active sites were obtained either by the grafting an organometallic Zr precursor onto a silica support or by incorporating Zr into the framework of a mesoporous silica. We observed that isolated, tetrahedrally coordinated Zr species ($\equiv\text{Zr-OH}$) are more active for ketonic decarboxylation of carboxylic acids and aldol condensation of ketones than tetrahedrally coordinated Zr dimer species and octahedrally coordinated ZrO_2 nanoparticles dispersed on silica. Based on this finding, we focused attention on the kinetics of ketonic decarboxylation and aldol condensation occurring on isolated $\equiv\text{Zr-OH}$ species. The mechanism by which these reactions proceed was probed both experimentally and by computational quantum chemistry calculations through density functional theory (DFT).

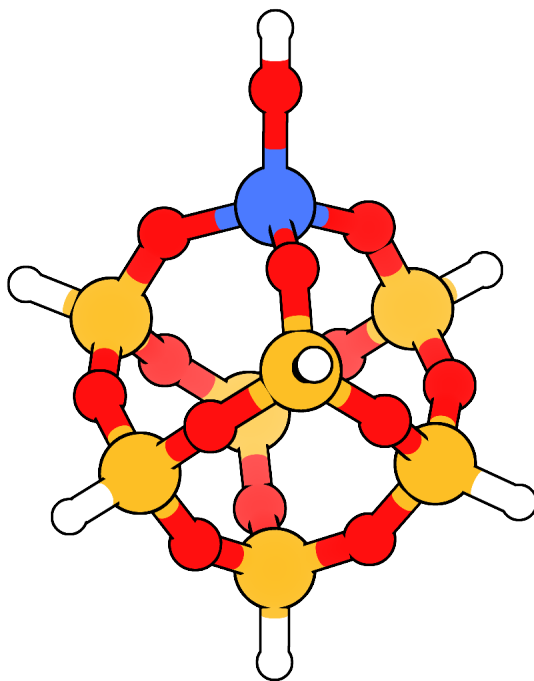


Figure 1: Cluster model of isolated M-OH site on silica support. White, red, and yellow spheres represent H, O, and Si atoms, respectively. The blue sphere represents the Zr metal site.

2. Computational Methods

The active site for ketonization and aldol condensation was represented by a cluster comprised of an isolated ZrOH group incorporated into the corner of a silsesquioxane, as shown in Figure 1. This model has previously been used to investigate aldol condensation on isolated TiOH groups supported on silica [20], and more generally silsesquioxanes have been used as molecular models of silica-grafted metal hydroxo and oxo species [21]. Geometry optimization of the initial structure preceded all calculations of adsorbed species. The structure of each adsorbed species on the optimized active site was hypothesized a priori, and then further optimized by relaxing the Zr metal center, all atoms in its first coordination sphere of Zr, the hydrogen of the terminal hydroxyl group, and all atoms of the adsorbate. The vibrations of the converged species were then determined by vibrational analysis. These calculations were done using the ω B97X-D functional [22] and the def2-SVP basis set [23], including the def2- effective core potential (ECP) for the Zr atom [24]. The resulting Hessian matrix output with zero negative eigenvalues (i.e. zero imaginary frequencies) confirmed that a given structure was at an energetic minimum. Structural convergence was followed by higher accuracy single-point energy calculations at the ω B97M-V/def2-TZVP level of theory [23, 25].

Estimates of the transition state structures occurring between reactant and product states were obtained using the frozen string method [26]. These estimates were then refined to transition structures at the ω B97X-D/def2-SVP level of theory. Similar to geometry optimization of reaction intermediates, transition structures were confirmed by the presence of one negative eigenvalue in the Hessian matrix corresponding to motion in a single direction (the reaction coordinate). Visualization of the corresponding imaginary frequency was used to identify a vibrational mode in the direction of bond formation. All calculations were performed using the Q-Chem software package [27].

Enthalpies and entropies at the reaction temperature were determined by calculating the zero-point vibrational energy and the temperature corrections to the enthalpy using the quasi-rigid rotor harmonic oscillator (q-RRHO) approach proposed by Grimme [28]. This approach replaces the vibrational entropy and enthalpy for low frequency modes ($< 100\text{cm}^{-1}$) by an interpolation between the rigid rotor harmonic oscillator (RRHO) vibrational values and free-rotor rotational values. The turnover frequencies (TOFs) of the catalytic cycles were predicted using Kozuch’s model of a catalytic cycle, which uses transition state theory to predict the TOF from estimates of the Gibbs free energy and enthalpy for all reaction intermediates and transition states, weighted by appropriate reactant and product partial pressure contributions [29, 30].

We used this model to determine the rate-limiting transition structures and most abundant surface intermediates for both ketonization and aldol condensation. Where applicable, we performed a degree of rate control analysis [31, 32] to determine dominant intermediate states in the reaction mechanism. These were then used to predict experimentally-observed activation energies from the free energies of reaction intermediates following a procedure described by Mao and Campbell [33].

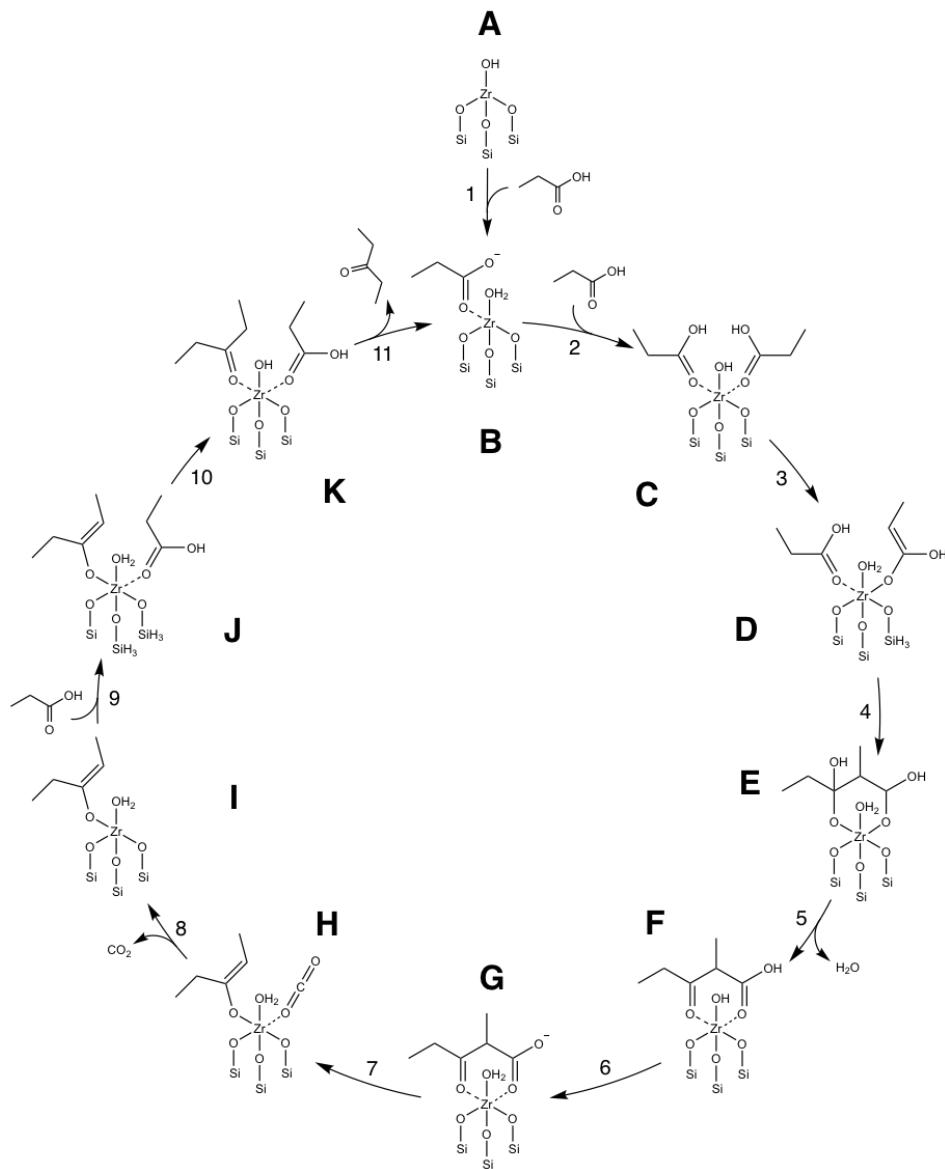
Experiments were performed by Shylesh [34], with the methods for catalyst synthesis, measurement of activity, and characterization included in Appendix A. The results of these experiments, also contained in Appendix A, are referenced against the computational predictions under our approximations.

3. Computational Results

3.1. Analysis of Propanoic Acid Ketonization

The minimum energy pathway for ketonization of propanoic acid over isolated Zr sites is shown in Scheme II.1, for which reaction intermediates were hypothesized based on the mechanism supported by experimental evidence. The reaction begins with the adsorption of two propanoic acid molecules, each one interacting with the nucleophilic oxygen of the carbonyl groups and the electrophilic Zr, as depicted in reactions 1 and 2. These elementary steps result in Zr becoming a saturated hexacoordinated species. Polarization of the carbonyl groups makes the carbon more electrophilic and the α -H more acidic, thereby facilitating abstraction of the α -H by the Lewis basic hydroxyl oxygen to form an adsorbed enolized carboxylic acid, as shown in reaction 3. The formation of this enolate intermediate mirrors that of aldol condensation previously described over isolated TiOH sites [20]. Polarization of the carbonyl group encourages nucleophilic attack of the enolate on the electrophilic carbon

on the second adsorbed propanoic acid molecule, leading to the formation of a C-C bond (reaction 4) in the hydroxy- β -keto acid intermediate (species E). Water leaves after the hydroxyl group on species E abstracts the proton on the Brønsted site, forming the β -keto acid. Deprotonation of the β -keto acid and cleavage of the C-C bond leads to decarboxylation and formation of an enolized ketone. Desorption of the CO₂ is followed by the energetically favorable repopulation of the site with another reactant molecule to maintain Zr saturation. Subsequent proton transfer from the Brønsted site to the enol π orbital forms the final product, pentan-3-one. Desorption of the product closes the cycle, reforming intermediate B.



Scheme II.1: Proposed reaction mechanism of ketonic decarboxylation of propanoic acid over isolated ZrOH to form 3-pentanone.

The predicted Gibb's free energy pathway at the reaction temperature (573 K) for the sequence described in Scheme II.1 is shown in Figure 2. The dissociative adsorption of propanoic acid on the empty site A forms intermediate B. The adsorption free energy for this step, reaction 1, is -2 kJ mol^{-1} . Because intermediate B represents the beginning of the cycle under the mechanism of Scheme II.1, we designate this state as the reference state. Our thermodynamic approximations predict adsorption of a second propanoic acid molecule to form intermediate C is approximately thermodynamically neutral at -2 kJ mol^{-1} . These two energetically similar states represent the lowest energy, most abundant intermediates. The largest free energy barrier corresponds to C-C bond formation (D^\ddagger) at 189 kJ mol^{-1} relative to the reference state, suggesting that C-C bond formation is the rate-limiting step. This conclusion is consistent with the experimental absence of an H/D kinetic isotope effect. The overall reaction is energetically favorable, since $\Delta G_{\text{rxn}} = -55 \text{ kJ mol}^{-1}$.

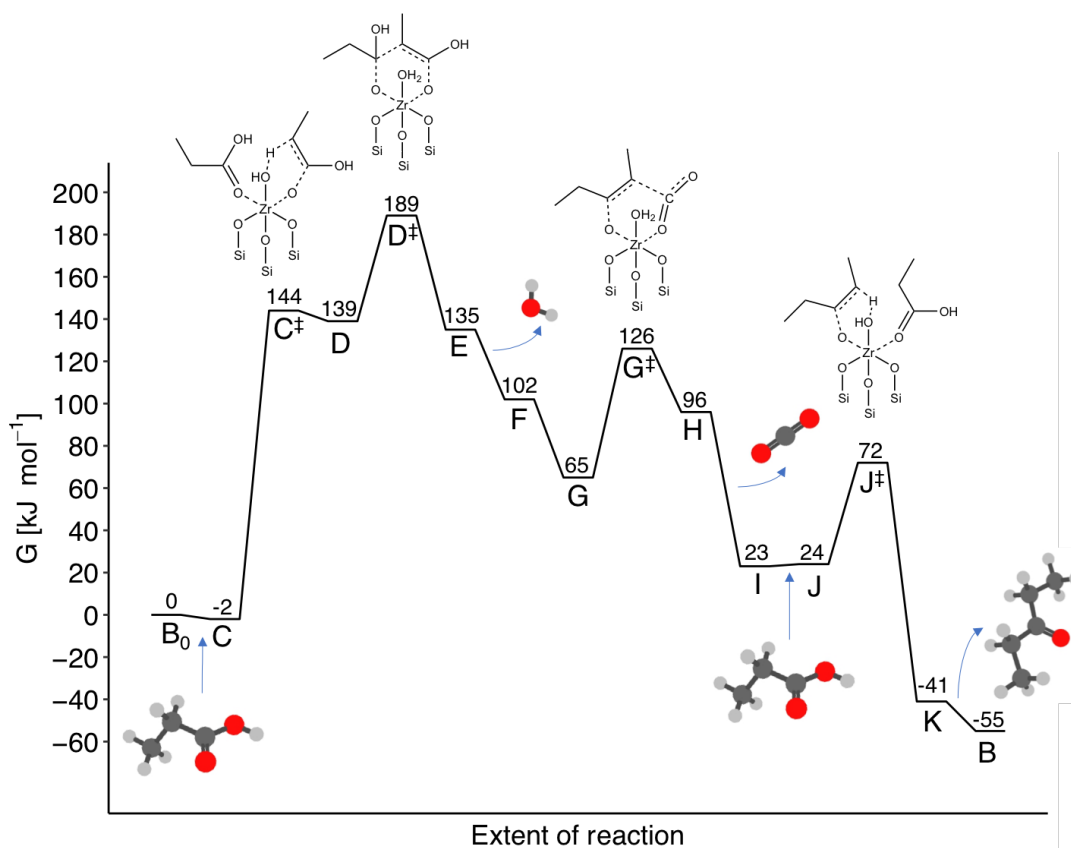


Figure 2: Free energy diagram for ketonic decarboxylation of propanoic acid over isolated $\equiv\text{ZrOH}$ species. The elementary steps are labeled according to the reaction sequence shown in Scheme II.1. Relevant transition state structures are depicted, as are the points where gas phase molecules enter and exit the cycle. Values were calculated at reaction conditions: $T = 573 \text{ K}$, $P_{\text{Tot.}} = 1 \text{ atm}$.

A rate expression for the kinetics of propanoic acid ketonization was derived using the reaction sequence shown in Scheme II.1 and the free energy profile presented in Figure 2 following the procedure described by Kozuch [29, 30]. Details of this derivation are given

in the Appendix (Beginning with Equation A.1). Figure 3 shows the dependence of the turnover frequency (TOF) on the partial pressure of propanoic acid (P_{acid}) at 573 K. The model of the TOF predicts less than first order kinetics in the range of experimental reactant partial pressures (0.1–0.4 kPa). At higher partial pressures (greater than approximately 2 kPa), it predicts approximately zero order in P_{acid} .

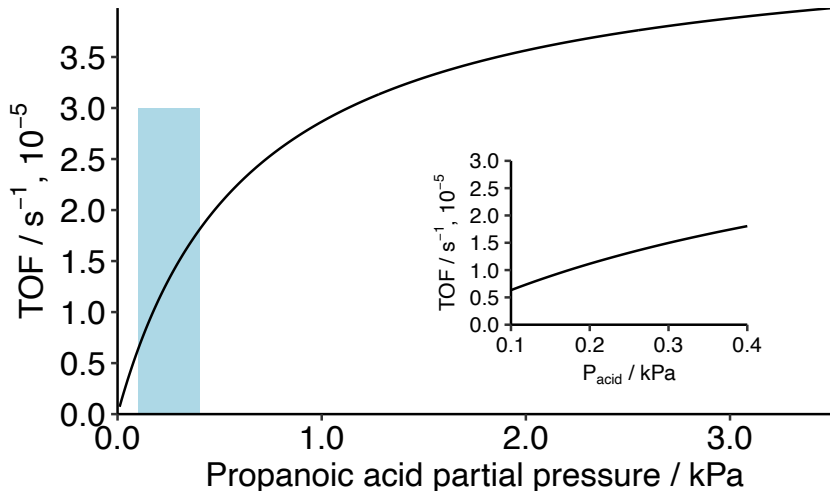


Figure 3: Predicted turnover frequency for ketonization of propanoic acid based on the mechanism proposed in Scheme II.1 and the free energy landscape presented in Figure 2. The reaction range (0.1–0.4 kPa) is shaded and expanded for clarity in the inset. Reaction conditions: $T = 573$ K, $P_{\text{Tot.}} = 1$ atm.

Under reaction conditions used in this study, the conversion of propanoic acid was $< 2\%$ and hence the concentrations of products was very small. Consequently, all terms in the rate expression that depend on product concentration are set to zero and the TOF is well approximated by Equation 2.1. The percent error between this expression and the complete equation for TOF is on the order of 10^{-3} s^{-1} for the range of P_{acid} considered. The result is an expression for the TOF as a function of reactant partial pressure that effectively captures the character of the reaction mechanism with the free energies of states B, C, and D^\ddagger .

$$\text{TOF} = \frac{k_B T}{h} \frac{P_{\text{acid}} e^{-\beta \Delta G_{\text{rxn}}}}{P_{\text{acid}} e^{\beta(G_{D^\ddagger} - G_C - \Delta G_{\text{rxn}})} + e^{\beta G_{D^\ddagger} - G_B}} \quad (2.1)$$

The TOF is dictated by both kinetic and thermodynamic quantities. While the kinetics are determined by the transition state barrier of the rate-limiting step, G_{D^\ddagger} , the thermodynamic quantities depend on the free energy minima of the reaction (the free energies of intermediates B and C). Because $G_C - \Delta G_{\text{rxn}}$ and G_B are roughly equal, our model predicts the order of the TOF with respect to P_{acid} is dependent on P_{acid} itself and is always < 1 . At vanishingly small P_{acid} , the second term in the denominator dominates the expression and the model predicts approximately first order kinetics. This signifies reaction equilibrium favors a singly adsorbed reactant (intermediate B) as the resting state. At high P_{acid} , the TOF is zero order in reactant partial pressure, signifying intermediate C is thermodynamically

avored. At reaction partial pressures (0.1–0.4 kPa), our model predicts kinetics between zero and first order, meaning intermediates B and C are similarly favorable and contribute significantly to the overall TOF. The resting state is therefore represented by a combination of singly and doubly adsorbed reactant species.

Campbell and coworkers defined the generalized degree of rate control (DRC) as a means for quantifying the thermodynamic importance of each intermediate and the kinetic importance of each transition state for the rate of a multi-step reaction mechanism [31, 32]. We performed this analysis (Equations A.5-A.8) in order to define the relative significance of each adsorbed species as a function of P_{acid} (see Figure A.14). As anticipated, we found that for the reaction conditions used in this study, both intermediates (B and C) have a significant DRC. Because both terms in the denominator of Equation 2.1 are significant, the apparent activation energy (E_{app}) cannot be calculated straightforwardly. This is evident on inspection of the enthalpy profile given in Figure 8. Therefore, E_{app} was approximated using an expression derived by Campbell [33], which expresses E_{app} using a sum of reactive intermediate and transition state enthalpies weighted by their appropriate DRCs (Equation A.3). The result is a predicted E_{app} reflecting a weighted average of enthalpy differences between states D^\ddagger and B and states D^\ddagger and C. The predicted E_{app} versus P_{acid} is depicted in Figure 5.

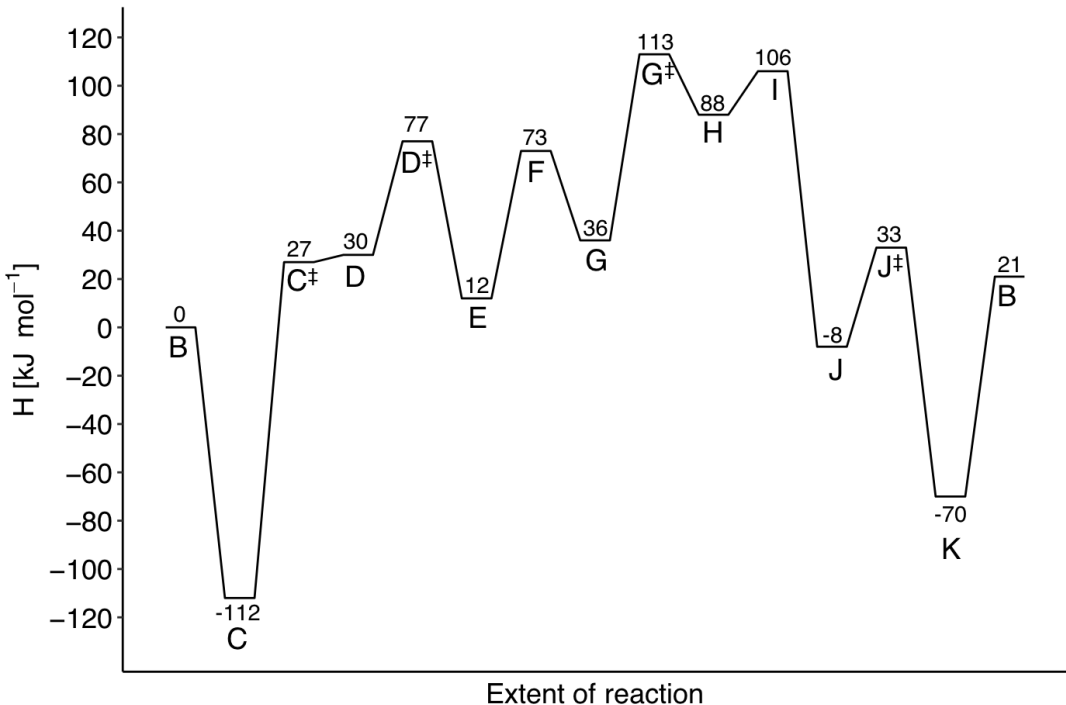


Figure 4: Enthalpy diagram of ketonic decarboxylation over isolated ZrOH. The elementary steps are labeled according to the reaction mechanism shown in Scheme II.1. Values were calculated at reaction conditions: $T = 573$ K, $P_{\text{Tot.}} = 1$ atm.

Under reaction conditions (0.1-0.4 kPa), we predict E_{app} to be between 96 and 124 kJ mol⁻¹. An Arrhenius plot of $\log(\text{TOF})$ vs T^{-1} using our model of this free energy

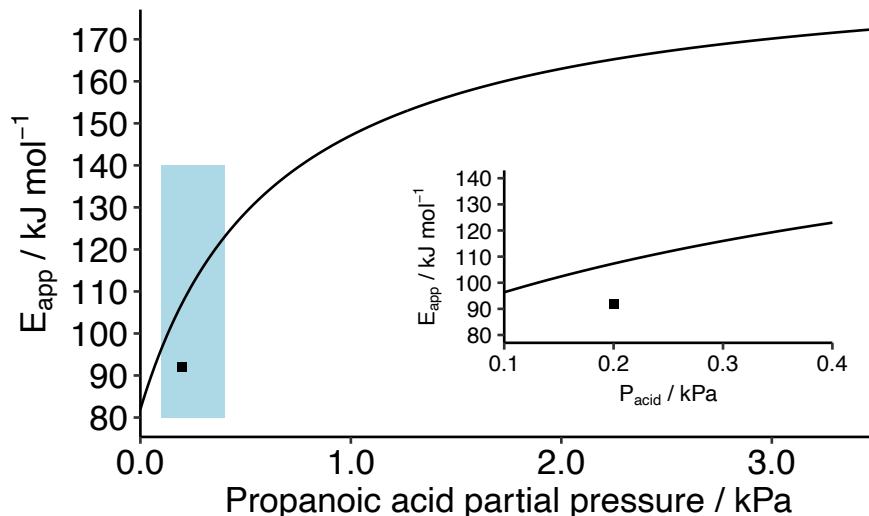


Figure 5: Predicted apparent activation energy of ketonic decarboxylation over isolated ZrOH as a function of reactant partial pressure. The reaction range (0.1-0.4 kPa) is shaded and expanded for clarity. The experimentally-observed activation barrier is depicted by a square data point. The reaction range (0.1-0.4 kPa) is shaded and expanded for clarity in the inset. Reaction conditions: $T = 573$ K, $P_{\text{Tot.}} = 1$ atm.

pathway under reaction conditions is in accord with this finding, predicting an E_{app} of 106 kJ mol^{-1} (Figure A.15) at $P_{\text{acid}} = 0.2$ kPa and $T = 563$ -593 K. Both analyses yield values in apparent agreement with the experimentally observed E_{app} of 92 kJ mol^{-1} . In further support of this mechanism, the predicted Gibbs activation energy of the reaction (ΔG^\ddagger) calculated from Equation A.11 is 198 kJ mol^{-1} , in reasonable agreement with 182 kJ mol^{-1} calculated from the experimental TOF.

The model predicts a reaction order in P_{acid} of 0.7, which differs from that observed experimentally, which is 0.1 between 0.1 and 0.4 kPa. We believe this discrepancy may be because our approximation of an immobile adsorbate with only quasi-harmonic vibrational modes fails to account for internal and surface rotations of adsorbed species, leading to an over-estimation of entropy loss upon adsorption ($-T\Delta S_{\text{ads}}$). This is a well-known problem of harmonic approximation techniques [35], and implies the free energy of intermediate C is more favorable than initially predicted. We therefore conducted an analysis to examine the extent to which changes to the adsorption entropy simultaneously affect the predicted ΔG^\ddagger , E_{app} , and P_{acid} dependence. The results of this analysis are shown in Figure 6, which shows that relatively small changes in $-T\Delta S_{\text{ads}}$ cause significant changes in ΔG^\ddagger , E_{app} , and the order in P_{acid} . Over the range of values of $-T\Delta S_{\text{ads}}$ shown, it is possible to achieve reasonable agreement with the values of ΔG^\ddagger and E_{app} observed experimentally but not the order in P_{acid} . Alternatively, it is possible to match the observed or in P_{acid} with a 16 kJ mol^{-1} adjustment to the adsorption entropy, but not the observed values of ΔG^\ddagger and E_{app} .

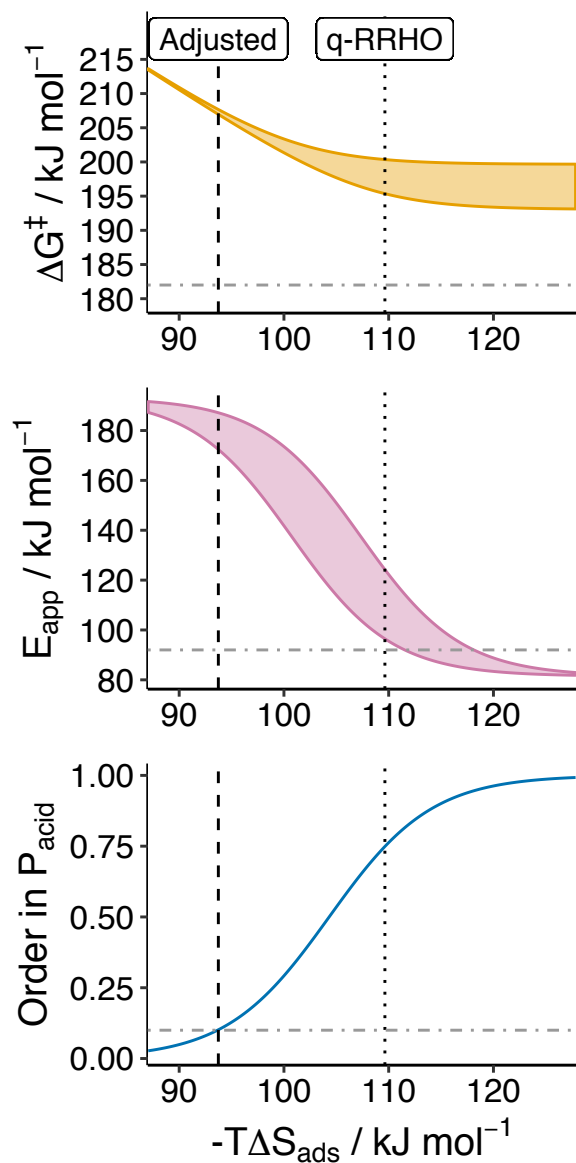


Figure 6: Effect of propanoic acid adsorption entropy on ketonization reaction order in partial pressure of reactant (blue, bottom), apparent activation energy (red, middle), and Gibbs activation energy (yellow, top). Each plot reflects the reaction region of $P_{\text{acid}} = 0.1\text{-}0.4$ kPa. The vertical dotted line corresponds to the q-RRHO approximation. The horizontal dashed lines on each plot correspond to experimentally observed quantities. Reaction conditions: $T = 573$ K, $P_{\text{Tot.}} = 1$ atm.

3.2. Analysis of Propan-2-one Condensation

The pathway for aldol condensation of propan-2-one over $\equiv\text{ZrOH}$ sites analyzed by DFT is shown in Scheme II.2, and mirrors that proposed for aldol condensation on $\equiv\text{TiOH}$ species [20]. The associated free energy profile is given in Figure 7. The reaction begins with adsorption of the carbonyl oxygen of propan-2-one to the Lewis-acidic Zr center. Adsorption of a second propan-2-one molecule prior to enolization of the adsorbed propan-2-one causes the α -hydrogen to become more acidic, thereby facilitating abstraction of the α -proton. We note that at reaction temperature ($T = 473$ K), our thermal analysis predicts each adsorption to be uphill in free energy by about 11 kJ mol^{-1} . Following the formation of the enolate, C-C coupling of the nucleophilic carbon of the π -system to the carbonyl carbon of the non-enolized co-adsorbate forms the diacetone alcohol, with a free energy activation barrier of 114 kJ mol^{-1} relative to the empty site. This is the free energy maximum of the cycle. The adsorbed diacetone alcohol readily dehydrates to form mesityl oxide.

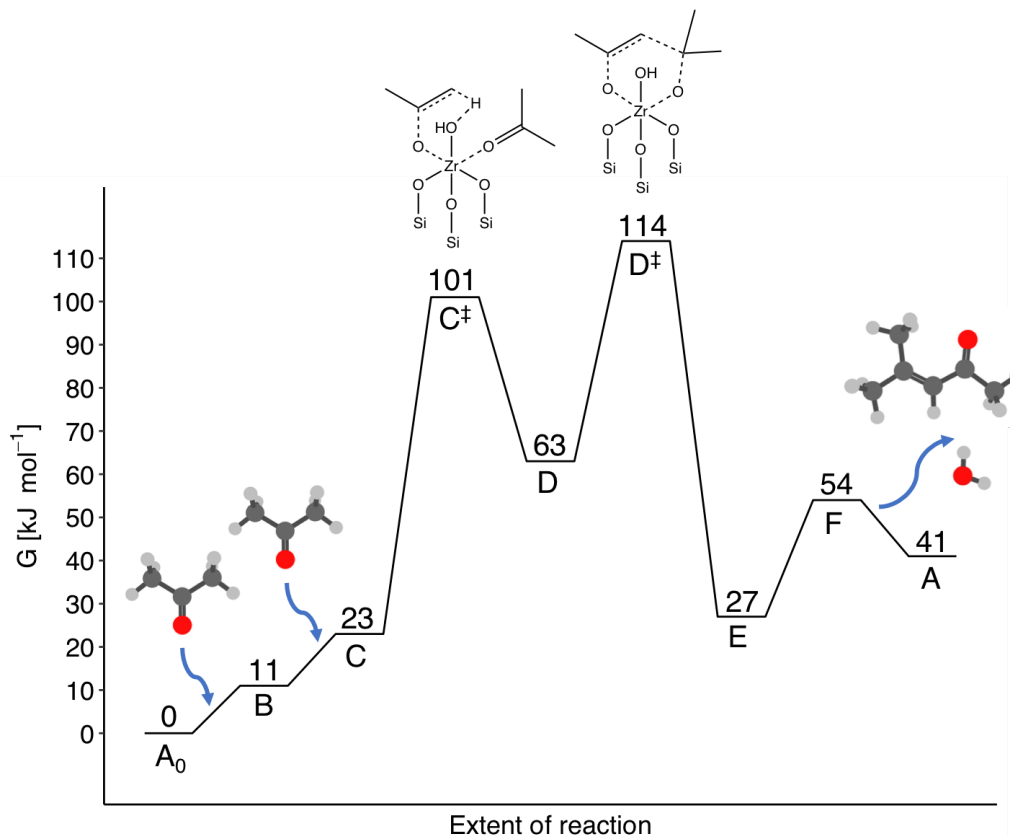
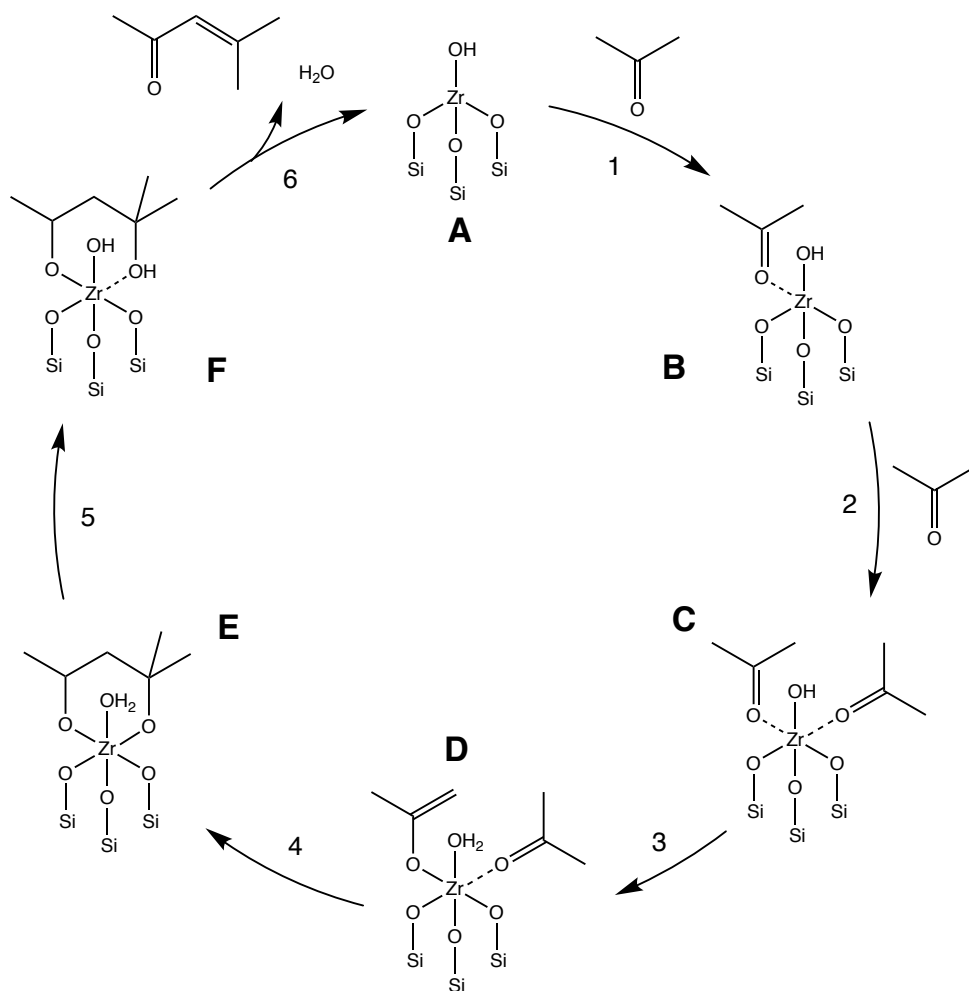


Figure 7: Free energy diagram for aldol condensation on isolated $\equiv\text{ZrOH}$ species. The elementary steps are labeled according to the reaction mechanism shown in Scheme II.2. Relevant transition state structures are depicted alongside, as are the points where gas-phase species enter and exit the cycle. Values were calculated at reaction conditions: $T = 573$ K, $P_{\text{Tot.}} = 1$ atm.

As with ketonization, the free energy profile suggests that the rate-limiting step is C-C



Scheme II.2: Minimum energy reaction mechanism of aldol condensation of propan-2-one over isolated $\equiv\text{ZrOH}$ species to form mesityl oxide.

bond formation (D^\ddagger), the free energy maximum in this cycle. This conclusion is consistent with an experimentally observed H/D kinetic isotope effect of unity. It is evident from the enthalpy diagram for this process (Figure 8) that the theoretically determined E_{app} is strongly influenced by which adsorbed state is identified as the resting state. If either A or B is the resting state, for example, E_{app} is negative, but if intermediate C is the resting state, E_{app} is 70 . With the experimentally observed E_{app} of 76 kJ mol^{-1} , intermediate C can be reasonably inferred as the resting state. The approximately zero order dependence of the reaction rate with P_{ketone} observed experimentally supports this hypothesis. However, the free energy profile predicting uphill adsorption of propan-2-one predicts a second order P_{ketone} dependence. As with ketonization, we postulate our quasi-harmonic approximations to the free energy overestimate the adsorption free energy of propan-2-one to the active site.

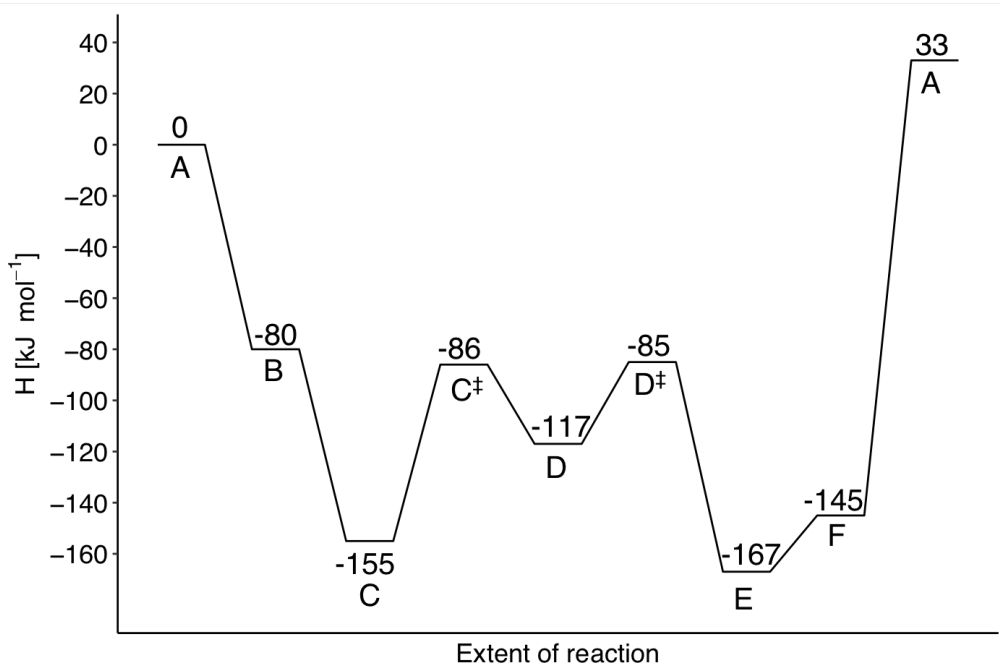


Figure 8: Enthalpy diagram of aldol condensation on isolated ZrOH. The elementary steps are labeled according to the reaction mechanism shown in Scheme II.2. Values were calculated at reaction conditions: $T = 573 \text{ K}$, $P_{\text{Tot.}} = 1 \text{ atm}$.

In light of these apparent limitations, we applied the same analysis used to investigate the effects of adjusting $-T\Delta S_{\text{ads}}$ in the ketonization reaction to aldol condensation. The computational details of the underlying TOF and degree of rate control calculations for this reaction are reported in Appendix A. As depicted in Figure 9, our current free energy approximation techniques under q-RRHO predict a reaction second order in P_{ketone} with a negative E_{app} . Applying a $-27.5 \text{ kJ mol}^{-1}$ adjustment to $-T\Delta S_{\text{ads}}$ for each propan-2-one adsorption leads to a predicted 0.1 order dependence in propan-2-one partial pressure and an E_{app} in the reaction range between 57 and 70 kJ mol^{-1} , in good agreement with the experimentally observed E_{app} of 76 kJ mol^{-1} . The predicted ΔG^\ddagger , after the adjustment to $-T\Delta S_{\text{ads}}$ is applied, is between 145 and 146 kJ mol^{-1} in the reaction range, in excellent

agreement with 145 kJ mol^{-1} calculated from experimental TOF data.

In contrast to ketonization, the adjustment to $-T\Delta S_{\text{ads}}$ required to correct the q-RRHO approximation to 0.1 order reactant partial pressure dependence is notably larger for aldol condensation ($-27.5 \text{ kJ mol}^{-1}$ for propan-2-one adsorption compared to -16 kJ mol^{-1} for propanoic acid adsorption). Whereas both propan-2-one and propanoic acid are C_3 oxygenates, we might expect similar error in our model of adsorption thermodynamics. We postulate this is because propanoic acid behaves more like an immobile adsorbate than propan-2-one due to the capacity of propanoic acid to hydrogen bond to the surface oxygen atoms near the adsorption site. Since propan-2-one lacks such a stabilizing force, it can retain more of its gas-phase rotational entropy upon adsorption. Indeed, factoring in the adjustment to adsorption entropy, adsorbed propan-2-one species are expected to retain approximately 17.5% of their total gas phase entropy, compared to 7% for adsorbed propanoic acid species. By comparison, rotational entropy for both propanoic acid and propan-2-one makes up about 30% of the total gas phase entropy, supporting the hypothesis that adsorbed propan-2-one may retain significantly more rotational entropy.

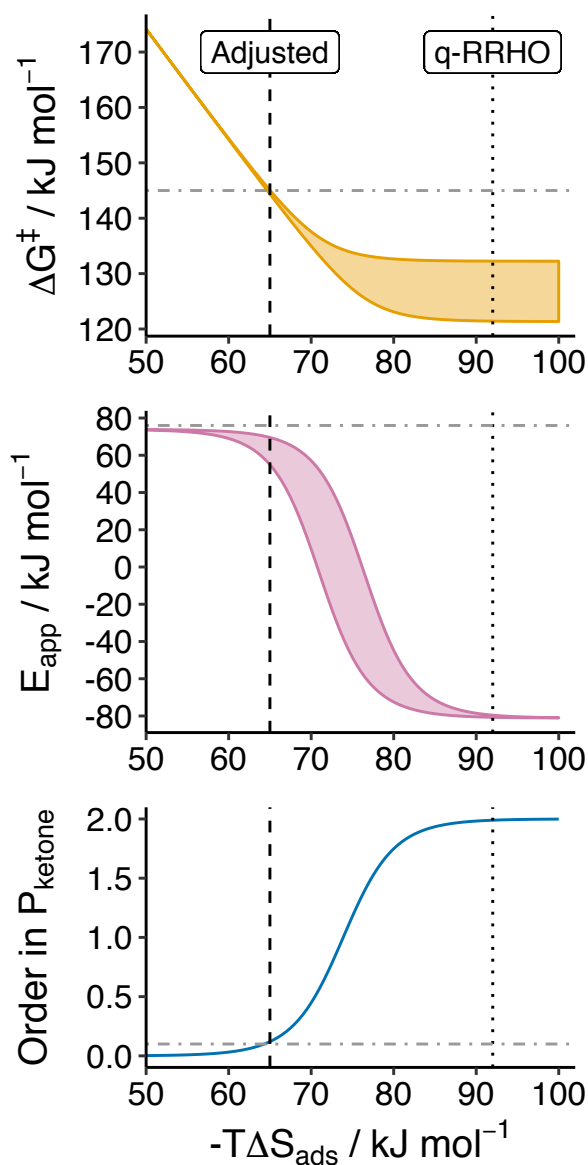


Figure 9: Effect of propan-2-one acid adsorption entropy on aldol condensation reaction order in partial pressure of reactant (blue, bottom), apparent activation energy (red, middle), and Gibbs activation energy (yellow, top). Each plot reflects the reaction region of $P_{\text{ketone}} = 0.1\text{-}0.4$ kPa. The vertical dotted line corresponds to the q-RRHO approximation. The vertical dashed line reflects the required adjustment for agreement with experimental partial pressure dependence. The horizontal dashed lines on each plot correspond to experimentally observed quantities. Reaction conditions: $T = 573$ K, $P_{\text{Tot.}} = 1$ atm.

4. Conclusions

We have examined the effects of Zr coordination environment and connectivity on the rate of ketonic decarboxylation of carboxylic acids and aldol condensation of propan-2-one over silica-supported zirconia prepared by grafting Zr onto the surface of amorphous silica. The turnover frequency for both reactions decreases monotonically with increasing Zr loading, leading to the conclusion that isolated $\equiv\text{ZrOH}$ species are more active than Zr oligomers or ZrO_2 nanoparticles. For both reactions, H/D isotope measurements indicate that cleavage of the α -C-H bond is not rate limiting. Instead, carbon-carbon bond formation is the rate-limiting step. Both reactions depend on the presence of a strong Lewis acid center for adsorption of either the carboxylic acid or the ketone.

Theoretical analysis of ketonic decarboxylation on isolated $\equiv\text{ZrOH}$ species suggests that the reaction mechanism for this process occurs via a β -keto acid intermediate and that isolated Zr sites remain coordinatively saturated throughout the cycle. Our analysis also suggests that C-C bond formation is the free energy maximum of the process, and is therefore the rate-limiting step. A thorough analysis of generalized degrees of rate control (DRCs) and associated partial pressure dependences suggests that both singly and doubly adsorbed propanoic acid species are significant thermodynamic intermediates under our thermodynamic approximations at reaction conditions. Varying the adsorption entropy of reactant adsorption above and below that predicted by the q-RRHO approximation, we demonstrated the sensitivity of reactant partial pressure dependence, apparent activation energy, and Gibbs activation energy to the predicted free energy of adsorption. This allowed us to quantify the degree to which our thermodynamic approximations might underestimate the entropy of adsorbed propanoic acid species, and to explain otherwise irreconcilable experimental quantities. While we found justifiable agreement between experimental and computed values among each observed value, the ensemble suggests combined limitations in computational and experimental methods.

Theoretical analysis supports the conclusion that C-C bond formation is the rate-limiting step for aldol condensation of propan-2-one. Experimental evidence suggests the resting state for this reaction is a saturated $\equiv\text{ZrOH}$ site with two adsorbed ketone molecules. Assuming that resting state, calculation of the apparent activation energy by inspection of the associated enthalpy diagram yields good agreement with experiment. The same sensitivity analysis performed for ketonization showed that we underestimate the entropy of adsorbed propan-2-one to a greater degree as adsorbed propanoic acid. Correcting for this over-approximation results in good agreement with the experimentally observed partial pressure dependence, apparent activation energy, and Gibbs activation energy.

More broadly, we have also obtained insight into the applicability of our computational methodology. Notwithstanding the pervasiveness of error in harmonic approximation techniques, the combined use of Kozuch’s model of a catalytic cycle and Campbell’s generalized degree of rate control allowed for both quantification of the extent of this error and the promise of reconciliation of computational studies with experimental results notwithstanding this error.

5. Acknowledgements

This work was funded by Director, Office of Science, Office of Basic Energy Sciences of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. DFT Calculations were performed on a computing cluster sponsored by the National Institutes of Health (NIH S10OD023532). We also gratefully acknowledge the contributions of Andrew Keeton, Keyang Sun, James P. Dombrowski, and Christopher Ho to the experimental section of Appendix A.

References

- [1] George W Huber, Sara Iborra, and Avelino Corma. "Synthesis of Transportation Fuels from Biomass: Chemistry, Catalysts, and Engineering". *Chemical Reviews* **106.9** (2006), 4044–4098
- [2] Steven Chu and Arun Majumdar. "Opportunities and Challenges for a Sustainable Energy Future". *Nature* **488.7411** (2012), 294–303
- [3] David Martin Alonso, Stephanie G Wettstein, and James A Dumesic. "Bimetallic Catalysts for Upgrading of Biomass to Fuels and Chemicals". *Chemical Society Reviews* **41.24** (2012), 8075–8098
- [4] Maria J Climent, Avelino Corma, and Sara Iborra. "Conversion of Biomass Platform Molecules into Fuel Additives and Liquid Hydrocarbon Fuels". *Green Chemistry* **16.2** (2014), 516–547
- [5] Cateryna Aiello-Mazzarri, Frank K Agbogbo, and Mark T Holtzapple. "Conversion of Municipal Solid Waste to Carboxylic Acids using a Mixed Culture of Mesophilic Microorganisms". *Bioresource Technology* **97.1** (2006), 47–56
- [6] Avelino Corma, Michael Renz, and Colin Schaverien. "Coupling Fatty Acids by Ketonic Decarboxylation using Solid Catalysts for the Direct Production of Diesel, Lubricants, and Chemicals". *ChemSusChem: Chemistry & Sustainability Energy & Materials* **1.8-9** (2008), 739–741
- [7] Sankaranarayananpillai Shylesh et al. "Integrated Catalytic Sequences for Catalytic Upgrading of Bio-derived Carboxylic Acids to Fuels, Lubricants and Chemical Feedstocks". *Sustainable Energy & Fuels* **1.8** (2017), 1805–1809
- [8] Juan Carlos Serrano-Ruiz and James A Dumesic. "Catalytic Routes for the Conversion of Biomass into Liquid Hydrocarbon Transportation Fuels". *Energy & Environmental Science* **4.1** (2011), 83–99
- [9] Shuai Wang and Enrique Iglesia. "Experimental and Theoretical Assessment of the Mechanism and Site Requirements for Ketonization of Carboxylic Acids on Oxides". *J. Catal.* **345** (2017), 183–206
- [10] Sankaranarayananpillai Shylesh et al. "Novel Strategies for the Production of Fuels, Lubricants, and Chemicals from Biomass". *Acc. Chem. Res.* **50.10** (2017), 2589–2597

- [11] Tu N Pham et al. “Ketonization of Carboxylic Acids: Mechanisms, Catalysts, and Implications for Biomass Conversion”. *ACS Catal.* **3.11** (2013), 2456–2473
- [12] James A Bennett et al. “Acetic Acid Ketonization over Fe₃O₄/SiO₂ for pyrolysis bio-oil upgrading”. *ChemCatChem* **9.9** (2017), 1648
- [13] Madhesan Balakrishnan et al. “Novel Pathways for Fuels and Lubricants from Biomass Optimized using Life-cycle Greenhouse Gas Assessment”. *Proc. Natl. Acad. Sci.* **112.25** (2015), 7645–7649
- [14] Gerald Jödicke et al. “Developing Environmentally-sound Processes in the Chemical Industry: a Case Study on Pharmaceutical Intermediates”. *J. Clean. Prod.* **7.2** (1999), 159–166
- [15] Gianfranco Pacchioni. “Ketonization of Carboxylic Acids in Biomass Conversion over TiO₂ and ZrO₂ Surfaces: a DFT Perspective”. *ACS Catal.* **4.9** (2014), 2874–2888
- [16] Angeles Pulido et al. “Ketonic Decarboxylation Reaction Mechanism: a Combined Experimental and DFT Study”. *ChemSusChem* **6.1** (2013), 141–151
- [17] S Damyanova, Paul Grange, and Bernard Delmon. “Surface Characterization of Zirconia-coated Alumina and Silica Carriers”. *J. Catal.* **168.2** (1997), 421–430
- [18] K Okumura and Y Iwasawa. “Zirconium Oxides Dispersed on Silica Derived from Cp₂ZrCl₂, [(i-PrCp)₂ZrH(μ-H)]₂, and Zr(OEt)₄ Characterized by X-Ray Absorption Fine Structure and Catalytic Ketonization of Acetic Acid”. *J. Catal.* **164.2** (1996), 440–448
- [19] Thomas Maschmeyer et al. “Heterogeneous Catalysts Obtained by Grafting Metallocene Complexes onto Mesoporous Silica”. *Nature* **378.6553** (1995), 159–162
- [20] David G Hanna et al. “Experimental and Theoretical Study of n-Butanal Self-Condensation over Ti Species Supported on Silica”. *ACS Catal.* **4.9** (2014), 2908–2916
- [21] Elsje Alessandra Quadrelli and Jean-Marie Basset. “On Silsesquioxanes’ Accuracy as Molecular Models for Silica-grafted Complexes in Heterogeneous Catalysis”. *Coordination Chemistry Reviews* **254.5-6** (2010), 707–728
- [22] Jeng-Da Chai and Martin Head-Gordon. “Long-range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections”. *Phys. Chem. Chem. Phys.* **10.44** (2008), 6615–6620
- [23] Florian Weigend and Reinhart Ahlrichs. “Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy”. *Phys. Chem. Chem. Phys.* **7.18** (2005), 3297–3305
- [24] Florian Weigend and Reinhart Ahlrichs. “Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy”. *Phys. Chem. Chem. Phys.* **7.18** (2005), 3297–3305
- [25] Narbe Mardirossian and Martin Head-Gordon. “ω B97M-V: A Combinatorially Optimized, Range-separated Hybrid, Meta-GGA Density Functional with VV10 Nonlocal Correlation”. *J. Chem. Phys.* **144.21** (2016), 214110

- [26] Shaama Mallikarjun Sharada et al. “Automated Transition State Searches without Evaluating the Hessian”. *J. Chem. Theory Comput.* **8.12** (2012), 5166–5174
- [27] Yihan Shao et al. “Advances in Molecular Quantum Chemistry Contained in the Q-Chem 4 Program Package”. *Molecular Physics* **113.2** (2015), 184–215
- [28] Stefan Grimme. “Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory”. *Chem. Eur. J.* **18.32** (2012), 9955–9964
- [29] Sebastian Kozuch and Sason Shaik. “How to Conceptualize Catalytic Cycles? The Energetic Span Model”. *Acc. Chem. Res.* **44.2** (2011), 101–110
- [30] Sebastian Kozuch. “A Refinement of Everyday Thinking: the Energetic Span Model for Kinetic Assessment of Catalytic Cycles”. *Wiley Interdiscip. Rev. Comp. Mol. Sci.* **2.5** (2012), 795–815
- [31] Charles T Campbell. *The Degree of Rate Control: a Powerful Tool for Catalysis Research*. 2017
- [32] Carsten Stegelmann, Anders Andreasen, and Charles T Campbell. “Degree of Rate Control: How Much the Energies of Intermediates and Transition States Control Rates”. *J. Am. Chem. Soc.* **131.23** (2009), 8077–8082
- [33] Zhongtian Mao and Charles T Campbell. “Apparent Activation Energies in Complex Reaction Mechanisms: a Simple Relationship via Degrees of Rate Control”. *ACS Catal.* **9.10** (2019), 9465–9473
- [34] Sankaranarayananpillai Shylesh et al. “Experimental and Computational Studies of Carbon–Carbon Bond Formation via Ketonization and Aldol Condensation over Site-isolated Zirconium Catalysts”. *ACS Catal.* **10.8** (2020), 4566–4579
- [35] Charles T Campbell and Jason RV Sellers. “The Entropies of Adsorbed Molecules”. *J. Am. Chem. Soc.* **134.43** (2012), 18109–18115

Chapter III

Pitzer-Gwinn Approximations for Gas Phase Molecules with Significant Torsional Anharmonicity and Mode Coupling

PARTITION FUNCTIONS of gas phase molecular species with significant anharmonic internal rotations were calculated using Pitzer-Gwinn (PG) approximations that account for internal rotation mode coupling. The idea behind PG methods is that classical partition functions can be corrected for quantum effects by applying a quantum correction given by the ratio of the quantum to classical partition functions under a known, simpler reference potential. Or, equivalently, mode-coupling effects neglected in uncoupled quantum treatments can be included as a correction at the classical level. In its original formulation, the PG approximation was applied to the classical partition function of internal rotation (i.e. torsion) using the harmonic oscillator (HO) potential energy as the reference potential. Recently-developed computational protocols for solving the energy levels of one-dimensional, torsional uncoupled mode (UM-T) potential energy surfaces are explored in this work as improved references. The difficulty in evaluating the UM-T-reference PG approximation stems from needing to solve the multidimensional integral in the classical mode-coupled partition function with the inclusion of torsional rotations. This is overcome through Markov Chain Monte Carlo, namely the No U-Turn Sampler (NUTS), a Hamiltonian Monte Carlo algorithm that operates as an efficient *ab initio* molecular dynamics simulation. The calculated partition functions were compared against simpler approximations, including typical harmonic oscillator (HO) approximations and PG approximations that include HO reference potentials. The heat capacities, enthalpies, and entropies were thus computed and compared against accepted values for a set of twelve organic molecules containing multiple torsional degrees of freedom. It was observed that PG methods that use UM-T reference potentials accounting for torsional mode-coupling are more accurate relative to the HO, UM-T, and harmonic reference alternatives.

1. Introduction

Theoretical determination of the thermodynamic properties of molecules and their interactions with their environment (e. g., a solvent or the walls of a porous adsorbent) are the subject of ongoing research having impact on many practical problems, including the thermodynamics of vapor liquid condensation and adsorption. Advances in ab initio electronic structure methods, such as density functional theory (DFT) [1, 2, 3, 4, 5], have enabled calculation of the electronic energy of molecules in their ground state with chemical accuracy (± 1 kcal mol⁻¹). However, determination of thermodynamic properties at finite temperatures remains a challenge because it requires the determination of the molecular partition function, Q , which necessitates knowledge of potential energy surface (PES) governing molecular motion. The challenge is that the motion involves vibrations, rotations, and torsions, which are in general coupled to one another. Therefore, standard variational approaches to obtain the quantum nuclear partition function from the Schrödinger equation are computationally prohibitive except for extremely small systems because it requires both sampling of the full-dimensional PES and diagonalization of a sizeable Hamiltonian matrix [6]. Methods such as path integral Monte Carlo (PIMC) and path integral molecular dynamics (PIMD) account for these mode-mode coupling effects, but these methods are also burdened with high computational costs because they can require millions of samples to reach convergence [7]. For this reason, more approximate representations of the PES have been assessed. The problem can be greatly simplified, for example, by assuming molecular motions are separable. These uncoupled partition functions are relatively easy to evaluate because they can be divided into a set of one-dimensional systems and solved independently. Evaluation of the partition function therefore only requires description of the PES along the axes of each internal coordinate.

The simplest of these approximations is the harmonic oscillator (HO) approximation, in which all internal modes are assumed to have a quadratic potential. This model has widespread use in vibrational analysis of molecular models, since the frequencies required to attain each oscillator’s partition function are easily obtained by diagonalization of the mass-weighted Hessian. However, for internal modes that significantly deviate from a harmonic potential, this approximation leads to errors in spectroscopic and thermodynamic observables. Molecules and molecular complexes with large-amplitude internal rotations (e.g. torsions around single bonds) are notoriously anharmonic and non-local, which can lead to difficulties in accurately describing the entropy [8]. Neglect of mode coupling may typically lead to underestimates of the entropy, while the well-known zero frequency divergence of the HO partition function can lead to overestimates [9, 10, 11].

While ad hoc techniques such as the quasi-rigid rotor/harmonic oscillator approximation [12] or corrections to low-frequency normal modes [13, 14] - which tend to correlate to hindered rotation or translation - attempt to correct this anharmonicity, a more rigorous approach is to model the internal degrees of freedom assuming a realistic anharmonic potential. Perhaps the simplest way this can be done is by sampling electronic energies for each normal mode by finite differences along each uncoupled coordinate, as detailed by previous works [6, 8, 15, 16, 17, 18]. This most basic form, the uncoupled normal mode (UM-N) model, uses finite displacement along each normal mode coordinate, leading to an anhar-

monic model whose computational cost scales with molecular size in a manner similar to that for harmonic analysis. However, the application of this approach has not led to significant improvement in the calculated thermodynamic properties compared to harmonic analysis [8]. Rather, significant improvements can be obtained by using torsional (T) coordinates to capture large-amplitude motions and harmonic vibrational modes for the remaining degrees of freedom (UM-T model). For torsions, this requires obtaining each one-dimensional torsional potential by independently sampling dihedral angles along the torsional internal coordinate. In the UM-T approximation, the associated PES potentials at each sampling point are used to obtain a continuous representation of the PES as a function of the dihedral angles of internal rotations [6, 8, 18], and the partition function is obtained by direct eigenvalue summation after variationally solving the Schrödinger equation for this potential using Fourier basis functions.

Despite the improved accuracy of the UM-T approximation versus the HO approximation, a method that includes torsional mode coupling effects is still desired because these non-local modes are expected to exhibit the greatest degree of interaction. One such method is the Pitzer-Gwinn (PG) approximation, which has specifically been used to investigate torsional mode coupling for low-dimensional systems, notably by Truhlar and coworkers [6, 19, 20, 21]. Under this class of approximation, the torsional partition function is represented as the product of the classical torsional partition function, which includes mode coupling terms, and a quantum correction factor, given by the ratio of the quantum to classical partition functions for a known reference potential. Much of the research into these methods has been devoted to improving the reference potential beyond the HO reference originally proposed by Pitzer and Gwinn [22], and including additional mode coupling terms in the classical partition function [6, 19, 20]. Among these improvements to the PG method is the improved-reference PG approximation introduced by Ellingson, Lynch, et al. in their comprehensive studies of PG methods applied to hydrogen peroxide and its isotopologs [20]. In their approach, the reference potential was taken to be the uncoupled torsional potential (UM-T), with the corrective factor applied to the classical partition function of the single torsion coupled to anharmonic vibrations and rotations. Simón-Carbadillo et al. used a similar approach for two-dimensional systems in their E2DT method [6], in which the reference potential was described by the two-dimensional torsional potential – whose quantum partition function was simple enough to be obtained by direct eigenvalue summation. However, these anharmonic-reference, mode-coupled PG methods have not yet been generalized to molecules with more than two torsions. This is undoubtedly because evaluation of the torsional partition function for a coupled system requires solving a challenging multidimensional integral, the evaluation of which suffers from the same curse of dimensionality that uncoupled approximations are intended to avoid.

Fortunately, substantial improvements to Markov chain Monte Carlo (MCMC) algorithms have been made to enable computationally tractable evaluation of challenging multidimensional integrals [23, 24, 25]. Of particular note is a class of MCMC algorithms called Hamiltonian Monte Carlo (HMC), which has been demonstrated to accurately model joint distributions of high-dimensional problems with fewer iterations needed for convergence versus conventional random walk algorithms [23]. Using this method, sampling of variables is done by random simulations of Hamiltonian dynamics [23, 26, 27, 28]. More recently, an extension of HMC was developed that disallows "U-turns" in the simulated trajectory and automati-

cally optimizes sampling parameters that would otherwise require costly hand-tuning [24].

With these promising developments at our disposal, we revisit and extend PG-corrective schemes to treat mode coupling in high-dimensional problems. The broad purpose of this undertaking is to investigate the accuracy of using uncoupled modes as reference potentials for more complex potential energy surfaces of interest, and to compare the results to those of more approximate methods, including uncoupled mode approximations. We focus our attention on the internal rotational modes in gas phase molecules, whose large-amplitude motions are expected to contribute to significant mode coupling. We first examine the theory and implementation of uncoupled and torsional PG approximations, ultimately presenting the UM-T-reference coupled PG approximation, whose non-separable, classical partition function is calculated by the No U-Turn Sampling (NUTS) HMC algorithm. We call this method the $\text{PG}_{\text{CM-T}}^{\text{UM-T}}$ method, where the superscript indicates that the reference potential is the UM-T potential, and the subscript indicates that the corrective factor is applied to the classical partition function of the mode-coupled energy for torsions (CM-T). This method is compared against the similarly-named $\text{PG}_{\text{UM-T}}^{\text{HO}}$ and $\text{PG}_{\text{CM-T}}^{\text{HO}}$ methods, as well as the UM-T and HO methods. Across all methods, non-torsional vibrations and whole-molecule rotations are assumed to be separable from molecular torsions, and are simply calculated using the HO and rigid rotor approximations. This allows us to then assess the extent to which torsional uncoupled mode approximations might be in error by isolating the mode coupling to internal rotations. We report significant improvement using the $\text{PG}_{\text{CM-T}}^{\text{UM-T}}$ method relative to harmonic-reference uncoupled PG approximations ($\text{PG}_{\text{UM-T}}^{\text{HO}}$), harmonic-reference coupled PG approximations ($\text{PG}_{\text{CM-T}}^{\text{HO}}$), UM-T, and HO approximations when comparing the calculated enthalpies, entropies, and heat capacities of each method against experimental references for a series of molecules. We also find reduction in the number of calculations required to achieve these thermodynamic quantities versus the direct solution, demonstrating the possibility of achieving accurate thermodynamics for significantly larger systems of interest than would otherwise be possible.

2. Theory and Methods

As detailed in previous studies [15, 16, 17, 29], uncoupled mode approximations generally begin with evaluation of the mass-weighted Hessian matrix at the equilibrium geometry. The diagonalization of this matrix decouples the problem into a set of harmonic normal modes. The associated eigenvalues of the normal modes are the squared frequencies of oscillation along the corresponding normal mode axes. The eigenvectors of the Schrödinger equation with a harmonic oscillator potential are obtained analytically as a family of Hermite polynomial functions with equally-spaced eigenvalues given by

$$E_{i,v} = \hbar\omega_i \left(v + \frac{1}{2} \right), \quad \forall v \in \mathbb{Z}^{\geq} \quad (3.1)$$

where i is the index of the vibrational mode and ω_i is the harmonic frequency of the same mode. The partition function and all thermodynamic quantities derived therefrom are functions only of the harmonic frequencies ($\omega_i, i = 1, \dots, N_{\text{vib}}$), as shown in Equations 3.2-3.3,

where the superscript *qu* denotes that energies of the partition function are quantized.

$$Q_{\text{vib},1\text{-D},i}^{qu,\text{HO}} = \sum_{\nu=0}^{\infty} e^{-\beta E_{i,\nu}} = \frac{e^{-\beta\hbar\omega_i/2}}{1 - e^{-\beta\hbar\omega_i}} \quad (3.2)$$

$$Q_{\text{vib}}^{qu,\text{HO}} = \prod_{i=1}^{N_{\text{vib}}} Q_{\text{vib},1\text{-D},i}^{qu} \quad (3.3)$$

Inclusion of anharmonic torsions under the same uncoupled framework (i.e. the UM-T method) is done by distorting the geometry of a given molecule along its torsional internal coordinate, ϕ_τ , with all other modes unrelaxed and performing energy calculations at each geometry sampled. The potential energies, $V_\tau(\phi_\tau)$, can be easily fitted to a Fourier series [6, 18] or interpolated [8], resulting in one-dimensional descriptions of the potential energy as a function of the dihedral angles of each torsion. For N torsions, this leads to the following estimation of the total torsional potential:

$$V(\phi_1, \dots, \phi_N) \approx \sum_{\tau=1}^N V_\tau(\phi_\tau) \quad (3.4)$$

The discrete energy levels of each one-dimensional potential can then be obtained as a numerical solution to the Schrödinger equation using the variational method, with the Hamiltonian matrix composed of the integrals

$$H_{mn} = \langle \Psi_m | \frac{-\hbar^2}{2I_\tau} \frac{\partial^2}{\partial \phi_\tau^2} + V_\tau(\phi_\tau) | \Psi_n \rangle \quad (3.5)$$

The matrix elements can be conveniently solved using Fourier basis functions,

$$\Psi_m(\phi) = \begin{cases} \frac{1}{\sqrt{2\pi}} & m = 0 \\ \frac{1}{\sqrt{\pi}} \cos\left(\frac{m+1}{2}\phi\right) & m \text{ odd} \\ \frac{1}{\sqrt{\pi}} \sin\left(\frac{m}{2}\phi\right) & m \text{ even} \end{cases}, \quad \forall m \in \mathbb{Z}^{\geq} \quad (3.6)$$

after which the one-dimensional torsional partition function is obtained by summing over the energy eigenvalues from diagonalizing the Hamiltonian matrix. Accounting for the torsional symmetry of the rotor, σ_τ , this leads to

$$Q_{\text{tors},1\text{-D},\tau}^{qu} = \frac{1}{\sigma_\tau} \sum_{m=0}^{M_\tau} e^{-\beta E_{\tau,m}} \quad (3.7)$$

where M_τ is the number of Fourier basis functions for torsion τ required for the sum to converge. Because each torsion is assumed to be uncoupled, the total torsional partition function is represented by the product of each individual torsional partition function,

$$Q_{\text{tors}}^{qu,\text{UM-T}} = \prod_{\tau=1}^N Q_{\text{tors},1\text{-D},\tau}^{qu} \quad (3.8)$$

This result is multiplied by the one-dimensional partition functions for each non-torsional internal degree of freedom – for which torsion-free normal mode coordinates are used – to obtain the total internal vibrational partition function.

Pitzer, Gwinn, and Kilpatrick also addressed the thermodynamics of internal rotations in gas phase alkanes [22, 30], representing the torsional partition function as the product of the classical anharmonic partition function of a torsion ($Q_{\text{tors}}^{cl, \text{CM-T}}$), multiplied by a quantum correction factor given by the ratio of the quantum to classical partition functions for a known reference potential. In the original Pitzer-Gwinn (PG) formulation, the reference potential was taken to be the harmonic potential: the second order Taylor series approximation of the anharmonic potential about the equilibrium position.

$$V(\phi_1, \dots, \phi_N) = \underbrace{\frac{1}{2} \sum_{\tau\tau'} \frac{\partial^2 V}{\partial \phi_\tau \partial \phi_{\tau'}} \Big|_{\text{eq.}} \phi_\tau \phi_{\tau'}}_{\text{HO reference potential}} + \frac{1}{3!} \sum_{\tau\tau'\tau''} \frac{\partial^3 V}{\partial \phi_\tau \partial \phi_{\tau'} \partial \phi_{\tau''}} \Big|_{\text{eq.}} \phi_\tau \phi_{\tau'} \phi_{\tau''} + \dots \quad (3.9)$$

The HO-reference coupled Pitzer-Gwinn partition function, labeled $\text{PG}_{\text{CM-T}}^{\text{HO}}$, is given mathematically by

$$Q_{\text{tors}}^{\text{PG}_{\text{CM-T}}^{\text{HO}}} = \frac{Q_{\text{tors}}^{qu, \text{HO}}}{Q_{\text{tors}}^{cl, \text{HO}}} Q_{\text{tors}}^{cl} \quad (3.10)$$

and the HO-reference prefactor is given by the expression

$$\frac{Q_{\text{tors}}^{qu, \text{HO}}}{Q_{\text{tors}}^{cl, \text{HO}}} = \prod_{\tau=1}^N \frac{e^{-\beta \hbar \omega_\tau / 2} (1 - e^{-\beta \hbar \omega_\tau})^{-1}}{(\beta \hbar \omega_\tau)^{-1}}. \quad (3.11)$$

The numerator of the HO prefactor is derived from Equations 3.2-3.3, with harmonic potential curvature determined by the harmonic frequencies corresponding to torsions instead of non-torsional vibrations (ω_τ , $\tau = 1, \dots, N$). The classical HO partition function counterpart is given by the expression in the denominator. The torsional frequencies used in both expressions are determined by the usual normal mode analysis using the torsional coordinate Hessian. The second term in the $\text{PG}_{\text{CM-T}}^{\text{HO}}$ approximation, Q_{tors}^{cl} , is the classical CM-T partition function, which is given by the multidimensional integral over all torsional phase-space:

$$Q_{\text{tors}}^{cl} = \frac{1}{h^N} \int_{-\infty}^{\infty} dp_1 \cdots \int_{-\infty}^{\infty} dp_N \int_0^{2\pi/\sigma_1} d\phi_1 \cdots \int_0^{2\pi/\sigma_N} d\phi_N e^{-\beta \left(\frac{p_1^2}{2I_1} + \cdots + \frac{p_N^2}{2I_N} + V(\phi_1, \dots, \phi_N) \right)}. \quad (3.12)$$

As mentioned previously, evaluation of this multidimensional integral for a given potential is the computational bottleneck. The simplest approximation to make this integral tractable is, again, the UM-T approximation. In contrast to Equation 3.7, the classical partition function for a single torsion is given by

$$Q_{\text{tors}, 1\text{-D}, \tau}^{cl} = \left(\frac{I_\tau}{2\pi\beta\hbar^2} \right)^{1/2} \int_0^{2\pi/\sigma_\tau} d\phi_\tau e^{-\beta V_\tau(\phi_\tau)} \quad (3.13)$$

and for N anharmonic torsional potentials, the total classical UM-T partition function is, similar to Equation 3.8, given by the product of the individual one-dimensional contributions.

$$Q_{\text{tors}}^{cl, \text{UM-T}} = \prod_{\tau=1}^N Q_{\text{tors}, 1\text{-D}, \tau}^{cl} \quad (3.14)$$

This simplification is the $\text{PG}_{\text{UM-T}}^{\text{HO}}$ approximation, which can be thought of as an approximation to the UM-T method solved by the variational method. This is because it involves the classical UM-T partition function “corrected” to the quantum level by the HO-reference prefactor. We will examine $\text{PG}_{\text{UM-T}}^{\text{HO}}$ alongside its more rigorous $\text{PG}_{\text{CM-T}}^{\text{HO}}$ counterpart after the numerical evaluation of the multidimensional integral is addressed.

Pitzer and Gwinn reasoned that in the high-temperature limit, the HO-reference prefactor approaches unity and the classical torsional partition function is recovered. Similarly, they reasoned that in the low-temperature limit, if the lowest energy levels of the torsional potential energy are reasonably approximated by the HO approximation, $Q_{\text{tors}}^{cl, \text{CM-T}}$ and $Q_{\text{tors}}^{cl, \text{HO}}$ become equal, recovering the quantum HO approximation. The qualification in their second assertion is a recognition that for many systems, particularly systems with multiple low-energy conformers, the HO approximation is suboptimal. For this reason, more sophisticated reference potentials for which energy levels can be calculated have been introduced to improve the original PG approximation. One such improvement can be found in multistructural models, in which torsions with multiple low energy conformers are represented by HO potentials centered at their local minima [19, 20, 21]. We posit that an even closer reference for the anharmonic potential can be found in the UM-T model, whose quantum and classical partition functions are readily computable by Equations 3.7-3.8 and 3.13-3.14. A pictorial representation of these reference methods is compared against a one-dimensional isopentane torsion in Figure 1, which illustrates the advantage of referencing the UM-T potential for multidimensional problems. A more quantitative reasoning based on the original literature can be found by expanding the N -dimensional potential in a Fourier series:

$$V(\phi_1, \dots, \phi_N) = \underbrace{\sum_{\tau=1}^N V_{\tau}(\phi_{\tau})}_{\text{UM-T reference}} + \sum_{m_1} \dots \sum_{m_N} c_{m_1 \dots m_N} e^{im_1 \phi_1} \dots e^{im_N \phi_N} \quad (3.15)$$

The higher order terms are corrections specifically accounting for mode coupling. These terms are expected to be much less important than the higher order terms in the Taylor series expansion given in Equation 3.9. This is in large part because the UM-T ansatz accounts for the non-locality (i.e. high probability of low-energy conformers) of molecular torsions, whereas the terms in the Taylor expansion are derived from the PES near the equilibrium position. The $\text{PG}_{\text{CM-T}}^{\text{UM-T}}$ approximation is then given by

$$Q_{\text{tors}}^{\text{PG}_{\text{CM-T}}^{\text{UM-T}}} = \frac{Q_{\text{tors}}^{qu, \text{UM-T}}}{Q_{\text{tors}}^{cl, \text{UM-T}}} Q_{\text{tors}}^{cl} \quad (3.16)$$

A pictorial representation of the components needed for this method is given in Figure 2.

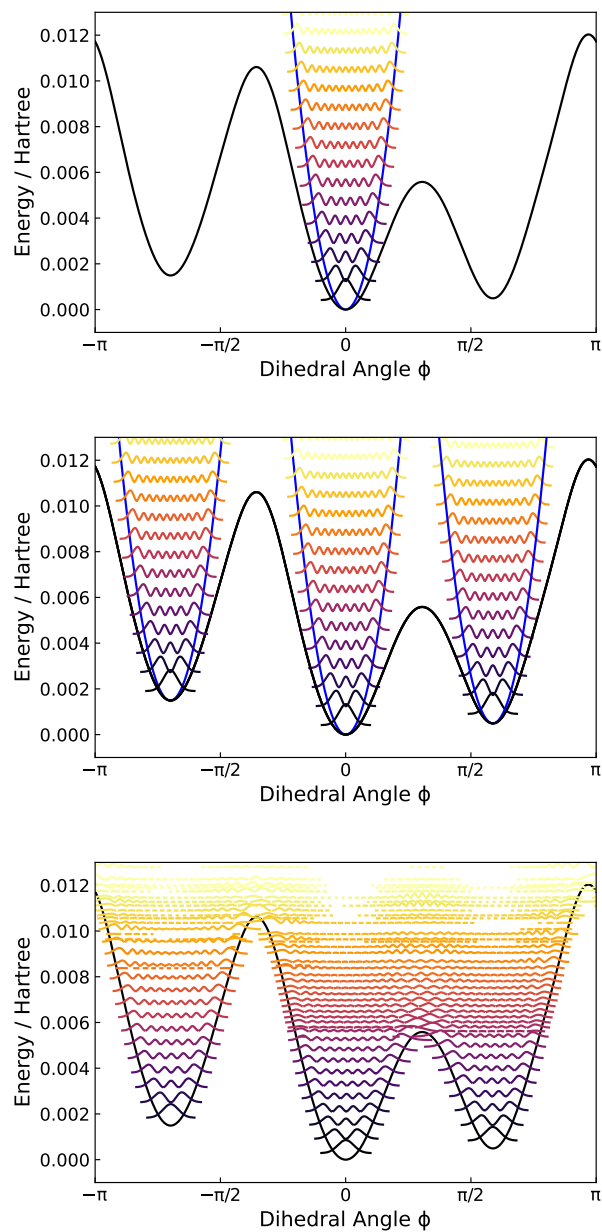


Figure 1: Graphical representations of HO-reference, multi-conformer HO-reference, and UM-T-reference approximations on a single torsion. The anharmonic potential of a single isobutane torsion is represented in black, the classical harmonic approximations to the same torsion in blue, and the discrete energy levels and squared probability amplitudes of the quantum results in multicolor.

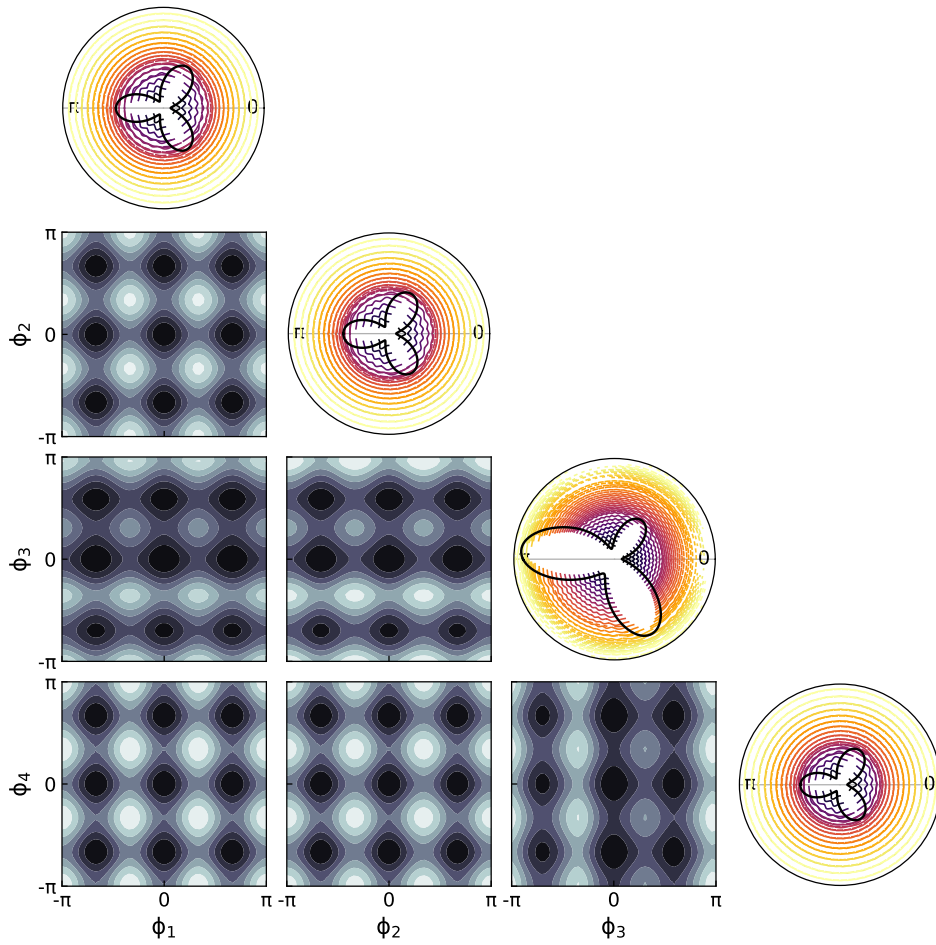


Figure 2: Schematic representation of the $\text{PG}_{\text{CM-T}}^{\text{UM-T}}$ method for an isobutane molecule. On the diagonal, classical UM-T potential (black) and the UM-T discrete energy levels and squared probability amplitudes (multicolor) associated with internal rotations are depicted with polar plots. The lower triangular plots depict the coupled, classical potential.

The approach for solving the classical CM-T partition functions necessary for the PG_{CM-T}^{UM-T} and PG_{CM-T}^{HO} methods is now addressed. Kilpatrick and Pitzer proposed the following solution to Equation 3.12 following studies by Eidinoff and Aston on nonrigid molecules [30, 31]:

$$Q_{\text{tors}}^{\text{cl}} = \left(\frac{1}{2\pi\beta\hbar^2} \right)^{N/2} \int_0^{2\pi/\sigma_1} d\phi_1 \cdots \int_0^{2\pi/\sigma_N} d\phi_N |D|^{1/2} e^{-\beta V(\phi_1, \dots, \phi_N)} \quad (3.17)$$

The term $|D|$ is the determinant of the torsional moment of inertia matrix [30, 32], which remains in the integral because it is not strictly independent of the molecular configuration [6]. The matrix is given by

$$D = \begin{bmatrix} I_1 & -\Lambda_{12} & \cdots & -\Lambda_{1N} \\ -\Lambda_{21} & I_2 & \cdots & -\Lambda_{2N} \\ & \vdots & \ddots & \vdots \\ -\Lambda_{N1} & -\Lambda_{N2} & \cdots & I_N \end{bmatrix} \quad (3.18)$$

The diagonal elements, I_τ , correspond to the reduced moments of inertia of the τ th torsion, and off-diagonal elements, $\Lambda_{\tau\tau'}$, are the coupling of torsional moments of inertia I_τ and $I_{\tau'}$. Any given element of the moment of inertia matrix is dependent on A_τ , the moment of the τ th top itself, I_x, I_y, I_z , the rigid rotational moments of inertia of the entire molecule, and $\lambda_{\tau i}$, the direction cosine between the axis of the τ th top and the i th principal axis of rotation [22, 30].

$$I_\tau = A_\tau - \sum_{i=x,y,z} \lambda_{\tau i}^2 A_\tau^2 / I_i \quad (3.19)$$

$$\Lambda_{\tau\tau'} = A_\tau A_{\tau'} \sum_{i=x,y,z} \lambda_{\tau i} \lambda_{\tau' i} / I_i \quad (3.20)$$

The non-separable integral in Equation 3.17 is solved by a Markov chain Monte Carlo approach, through which we can find thermodynamic quantities from samples drawn from the classical coupled distribution by statistical inference. This is done by Hamiltonian Monte Carlo (HMC), which has a direct equivalence to the physics of the problem. By this method, a Markov chain is simulated in which each iteration samples a random momentum and then performs a Metropolis update with a proposal determined after time-integrating the Hamiltonian equations of motion [23, 26, 27, 28]. The sampler thus explores multidimensional phase-space through solving dynamic trajectories. In the context of modeling the coupled molecular torsions, this is equivalent to stochastically simulating torsional AIMD in the canonical ensemble. The corresponding governing equations are

$$H(\Phi, \mathbf{p}) = V(\Phi) + \frac{1}{2} \mathbf{p}^T M^{-1} \mathbf{p} \quad (3.21)$$

$$\frac{d\Phi}{dt} = M^{-1} \mathbf{p} \quad (3.22)$$

$$\frac{d\mathbf{p}}{dt} = -\frac{\partial V}{\partial \Phi} \quad (3.23)$$

where H is the Hamiltonian, $\Phi = [\phi_1, \dots, \phi_N]^T$ is the vector of dihedral angles encoding a configuration of N torsions, V is the torsional potential, $\mathbf{p} = [p_1, \dots, p_N]^T$ is the vector of

dihedral angular momenta, and M is the mass matrix. In HMC, the mass matrix informs how momenta, with the chain indexed by k , are randomly drawn from the multivariate Gaussian distribution:

$$\mathbf{p}_k(t = 0) \sim \mathcal{N}(\mathbf{0}, M) . \quad (3.24)$$

The mass matrix has a direct physical analog in the torsional moment of inertia matrix D previously defined. In practice, however, M is optimally tuned or chosen to increase the efficiency of the sampler, and is not necessarily equal to D .

It is apparent from Equation 3.23 that in order to use HMC for sampling torsional conformations of a molecule, we require not only the energy at a given sampling point, but also its gradient. This can be readily obtained using DFT calculations, where the torsional energy gradients can be isolated from the internal coordinate energy gradient, which is related to the Cartesian gradient by a coordinate transform

$$\nabla_q V = B^{-1} \nabla_x V . \quad (3.25)$$

B^{-1} is the generalized inverse of the Wilson B-Matrix, the rectangular Jacobian mapping the Cartesian coordinates x to the internal coordinates q [8, 33]. We thus have the components to implement a Hamiltonian Monte Carlo method for sampling torsional potential energies from the classical mode coupled distribution.

We employed an extension of HMC called the No U-Turn Sampler (NUTS) [24]. This algorithm explores parameter space by disallowing "U-Turns", which can expend a significant amount of computational time solving the molecular dynamics for small gains in displacement. The U-Turn condition is determined by constructing a binary tree to trace the path of the trajectory in forward and backward directions using a leap frog time integrator. This method has the added benefit of automatically tuning sampling parameters such as time integration steps and the mass matrix M from which random momenta are drawn [34, 35]. Initialization of tuning parameters for this algorithm is discussed in the Computational Details section.

The torsional contributions to the energy and heat capacity can be expressed in terms of the mean and variance of the torsional potential energy V , with

$$E_{\text{tors}}^{\text{cl,CM-T}} = \frac{N}{2} RT + \langle V \rangle \quad (3.26)$$

$$C_{V, \text{tors}}^{\text{cl,CM-T}} = \frac{N}{2} R + \frac{\langle V^2 \rangle - \langle V \rangle^2}{RT^2} \quad (3.27)$$

where the angled brackets denote the mean of the quantity. Similarly, the entropy can be expressed in terms of expected values, with the following expression given in terms of the mean square root of $|D|$ and the mean negative logarithm of the probability density function (pdf) associated with Φ , $f(\Phi)$.

$$S_{\text{tors}}^{\text{cl,CM-T}} = R \left[\frac{N}{2} - \frac{N}{2} \ln(2\pi\beta\hbar^2) + \ln\left(\langle |D|^{1/2} \rangle\right) - \langle \ln f(\Phi) \rangle \right] \quad (3.28)$$

Derivations of these terms are provided in Appendix B.

Provided enough samples have been taken to accurately reflect the underlying distribution, most of the angle-bracketed quantities were found by their sample means. The exception is

the negative logarithm of the pdf, which is not directly measurable. Instead, this quantity was obtained through a kernel density estimator (KDE) [36, 37, 38], with the pdf estimated by a sum of local Gaussian functions centered at the sampling points.

$$f(\Phi) \approx \frac{1}{n} \sum_{i=1}^n K_{\Sigma}(\Phi - \Phi_i) \quad (3.29)$$

$$K_{\Sigma}(\Phi) = (2\pi)^{-N/2} |\Sigma|^{-1/2} e^{-\frac{1}{2}\Phi^T \Sigma^{-1} \Phi} \quad (3.30)$$

The bandwidth of the KDE, Σ , was chosen by Scott’s rule with $\sqrt{\Sigma_{\tau\tau}} = n^{-1/(N+4)} \varsigma_{\tau}$ [39], where n is the number of samples, N is the number of torsions, and ς_{τ} is the standard deviation of the τ th variable. This choice of KDE parameters was verified to accurately represent the pdfs of uncoupled potential energies sampled by the NUTS protocol. The expected value of the negative logarithm of the pdf was then solved as

$$\langle \ln f(\Phi) \rangle = \int d^N \Phi f(\Phi) \ln f(\Phi) \quad (3.31)$$

which was computed by Monte Carlo integration resampling from the distribution approximated in Equation 3.29.

Because thermodynamic state functions are derived from the logarithm of the partition function, the thermodynamics of the product (or quotient) of multiple partition functions are linearly additive. For example, the heat capacity, energy, and entropy of Equation 3.16 are calculated by

$$C_{V,\text{tors}}^{\text{PG}_{\text{CM-T}}^{\text{UM-T}}} = C_{V,\text{tors}}^{\text{qu,UM-T}} - C_{V,\text{tors}}^{\text{cl,UM-T}} + C_{V,\text{tors}}^{\text{cl,CM-T}} \quad (3.32)$$

$$E_{\text{tors}}^{\text{PG}_{\text{CM-T}}^{\text{UM-T}}} = E_{\text{tors}}^{\text{qu,UM-T}} - E_{\text{tors}}^{\text{cl,UM-T}} + E_{\text{tors}}^{\text{cl,CM-T}} \quad (3.33)$$

$$S_{\text{tors}}^{\text{PG}_{\text{CM-T}}^{\text{UM-T}}} = S_{\text{tors}}^{\text{qu,UM-T}} - S_{\text{tors}}^{\text{cl,UM-T}} + S_{\text{tors}}^{\text{cl,CM-T}} \quad (3.34)$$

Classical and quantum UM-T thermodynamic quantities derived from the partition functions of Equations 3.7 and 3.14 are given in Appendix B. These quantities were evaluated using numerical methods following sampling the uncoupled potentials by the protocol of Li [8, 40]. The remaining internal thermodynamics (i.e. the stretch and bend contributions) were solved under the HO approximation, with external thermodynamics, namely translational and rotational contributions, solved under the ideal gas and rigid rotor approximations. The thermodynamic variables of these modes are also given in Appendix B.

3. Computational Details

All electronic structure calculations were performed using a developmental version of the Q-Chem software package [41]. In accordance with previous work [8], all DFT calculations were carried out at the B97-D/6-31G* level of theory [42]. This level of theory demonstrated no statistically significant degradation in representing the one-dimensional torsional PES relative to ω B97X-D/6-311+G-(2df,2pd) [43], which is a far more computationally demanding level of theory. Potential energy functions for the uncoupled mode approximation (UM-T) were obtained by cubic spline interpolation following finite-difference sampling along

torsional coordinates at a 10° resolution [8]. All internal coordinate transformations and other associated matrix algebra was performed using the open-access Automated Property Estimator (APE) [40]. Thermodynamic parameters such as the symmetry numbers and moments of inertia, both for external and internal rotations, were calculated using the Reaction Mechanism Generator (RMG-Py) [44].

The No U-Turn Sampling algorithm was provided by the PyMC3 software package [34]. To enforce continuity at the periodic boundaries, sampling was done in two dimensional ($\cos \Phi$ and $\sin \Phi$) vector space. Approximate step size was fixed at approximately 10° , scaled down by $(\frac{1}{N})^{1/4}$ [39], where N is the number of dimensions (i.e. torsions). Because such a resolution was determined by previous studies to sufficiently capture the features of the torsional PES, adaptive step tuning was forgone in favor of using computational efforts of training the model into achieving a good estimate of the mass matrix. An initial equilibration run included a tuning schedule, wherein the mass matrix was iteratively adapted to a time-integration target acceptance rate of 70%. The tuning schedule comprised 500 iterations of the algorithm, which was determined sufficient in replicating distributions of uncoupled mode systems. Additional Hamiltonian Monte Carlo parameters assumed the default/recommended values, for which additional information can be found in the reference documentation [34, 35, 45]. Production runs comprised 10 parallel Markov chains of 1000 samples. These chains were pooled maintaining parallelism prior to computing standard thermodynamic quantities. These measures were sufficient to achieve convergence of the energy to a sampling standard deviation of less than $0.1 RT$ across all molecules tested. The convergence criterion was verified by computing thermodynamic quantities as a function of the number of samples. Statistical computations, namely KDEs and KDE integration, were performed using modified functions from the SciPy open-source library [46].

4. Results and Discussion

We compared the calculated gas phase heat capacities, enthalpies, and entropies for a set of twelve molecules for which these data are readily available. In order to specifically probe the performance of this method relative to uncoupled methods, the molecules in this study were chosen based on the degree to which significant torsional mode coupling was suspected, such as mode coupling through hydrogen bonding or steric hindrance. Heat capacity data, and the less common enthalpic and entropic data, were obtained from the National Institute of Standards and Technology database [47].

The calculated constant pressure heat capacities at 298 K are given in Table 1. The HO approximation tends to underestimate the expected values. This is attributed to the overestimation of the energy level spacing using HO energy levels, leading to an underestimation of the fluctuation between energy levels. Figure 1 illustrates the difference in energy spacing between the HO approximation and anharmonic UM-T approximation, where we observe closer spacing in energy levels using an anharmonic potential, particularly at excitation energies above potential barriers separating torsional conformers. The $\text{PG}_{\text{UM-T}}^{\text{HO}}$ method can be interpreted [6] to correct the errant quantum HO heat capacity toward the UM-T result by subtraction of the classical HO heat capacity and addition of a classical UM-T heat capacity (see Equation 3.32). Nevertheless, both UM-T and $\text{PG}_{\text{UM-T}}^{\text{HO}}$ tend to overestimate the heat

capacity, attributable to an overestimation of the energy variance when generally restrictive mode coupling effects are ignored. The root-mean-squared errors (RMSEs) of the HO, UM-T, and $\text{PG}_{\text{CM-T}}^{\text{HO}}$ methods are thus comparably around $2.5 \text{ cal mol}^{-1} \text{ K}^{-1}$. Superior results are obtained with mode-coupled methods, with the $\text{PG}_{\text{CM-T}}^{\text{HO}}$ and $\text{PG}_{\text{CM-T}}^{\text{UM-T}}$ methods generally coming within $1 \text{ cal mol}^{-1} \text{ K}^{-1}$ of the accepted value, as shown by the parity plot in Figure 3. The $\text{PG}_{\text{CM-T}}^{\text{HO}}$ method performs marginally better, with an RMSE of $0.85 \text{ cal mol}^{-1} \text{ K}^{-1}$, compared to $0.9 \text{ cal mol}^{-1} \text{ K}^{-1}$ for $\text{PG}_{\text{CM-T}}^{\text{UM-T}}$. The similarity in computed heat capacities between $\text{PG}_{\text{CM-T}}^{\text{HO}}$ and $\text{PG}_{\text{CM-T}}^{\text{UM-T}}$ methods, and the difference between the $\text{PG}_{\text{CM-T}}^{\text{HO}}$ and $\text{PG}_{\text{UM-T}}^{\text{HO}}$ methods, suggests these quantities depend much more on the assumption on mode coupling than they depend on the choice of reference. This hypothesis is supported by the closeness of the $\text{PG}_{\text{UM-T}}^{\text{HO}}$ and UM-T results, indicating the UM-T energy variances can be well-approximated by an appropriate adjustment of the quantum HO reference. Mathematically, these observations suggest approximate equality of quantum and classical heat capacities at a given level of theory.

The experimental results for the enthalpy for ten of the twelve molecules studied are listed in Table 2 and depicted in Figure 4, which examines the enthalpy difference between 298 K and 0 K. All methods tend to fall within 1 kcal mol^{-1} of the accepted values. This is consistent with what was found previously [8], where it was concluded that the enthalpy is not sensitive to the choice of model due to the elimination of degenerate conformers represented in the partition function by differentiation. As with the heat capacities, the HO approximation generally underestimates the experimentally determined enthalpies, while the UM-T approximation tends to overestimate the experimental enthalpies. Unlike the heat capacities, however, the accuracy of the enthalpy for PG approximations is much more sensitive to the choice of reference potential. The $\text{PG}_{\text{UM-T}}^{\text{HO}}$ and $\text{PG}_{\text{CM-T}}^{\text{HO}}$ methods, for example, lead to similar results, but can have limited accuracy where the harmonic approximation is a poor prior. Using the UM-T approximation reference potential in conjunction with the coupled torsional partition function yields more accurate results, with $\text{PG}_{\text{CM-T}}^{\text{UM-T}}$ consistently falling within less than $0.5 \text{ kcal mol}^{-1}$ of experimental values and having an RMSE of $0.22 \text{ kcal mol}^{-1}$. This result is strikingly good, and may well be controlled by errors in the DFT energies as much as by limitations of $\text{PG}_{\text{CM-T}}^{\text{UM-T}}$.

By contrast, the entropy is much more sensitive to the choice of model because of the direct dependence on the logarithm of the partition function [8]. Therefore, accurate results require explicit enumeration of degenerate conformers. As seen in Table 3 and Figure 5, this leads to an underestimation of entropies by the HO approximation, with values across most molecules falling outside the range of chemical accuracy. $\text{PG}_{\text{UM-T}}^{\text{HO}}$ and UM-T methods account for conformational effects, performing similarly well. Improvements are obtained by accounting for mode coupling, with the $\text{PG}_{\text{CM-T}}^{\text{HO}}$ and $\text{PG}_{\text{CM-T}}^{\text{UM-T}}$ method generally falling within the range of chemical accuracy, defined for the standard entropy as $\pm 3.35 \text{ cal mol}^{-1} \text{ K}^{-1}$, which is $\pm 1 \text{ kcal mol}^{-1}$ of TS at 298 K. $\text{PG}_{\text{CM-T}}^{\text{UM-T}}$ proves a better method, with a RMSE of $2.24 \text{ cal mol}^{-1} \text{ K}^{-1}$ compared to $2.53 \text{ cal mol}^{-1} \text{ K}^{-1}$ for $\text{PG}_{\text{CM-T}}^{\text{HO}}$. As with the heat capacity, the accuracy of the computed entropies is depends strongly on the assumption of mode coupling, but these results suggest improvements can also be obtained by improving the reference potential. This can be rationalized by the additional dependence of the entropy on the energy, which, as shown for the computed enthalpies, relies significantly on the choice of the reference potential.

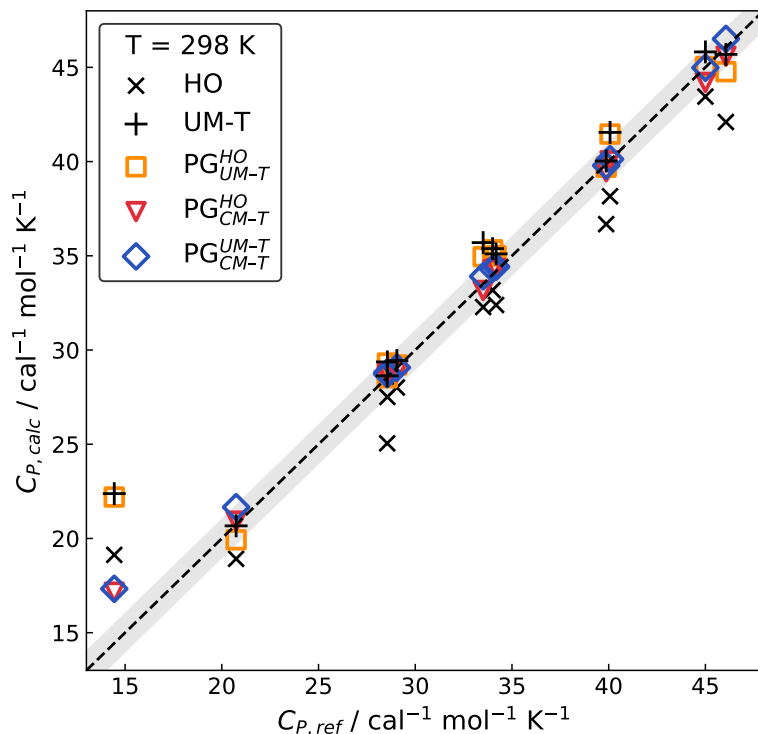


Figure 3: Parity plot for constant pressure heat capacity of select molecules in the gas-phase. The shaded region represents the interval of $\pm 1 \text{ cal mol}^{-1} \text{ K}^{-1}$ of the experimental value.

Table 1: Constant Pressure Heat Capacities for Selected Molecules at 298.15 K ($\text{cal mol}^{-1} \text{ K}^{-1}$)

Molecule	ref. ^a	HO	UM-T	$\text{PG}_{\text{UM-T}}^{\text{HO}}$	$\text{PG}_{\text{CM-T}}^{\text{HO}}$	$\text{PG}_{\text{CM-T}}^{\text{UM-T}}$
Ethylene glycol	20.74	18.92	20.67	19.93	20.92	21.66
Ethyl ether	28.55	25.05	28.63	28.53	28.62	28.72
Ethanedial	14.44	19.13	22.38	22.20	17.15	17.33
2,2-Dimethylbutane	34.00	33.18	35.38	35.32	34.26	34.32
2,3-Dimethylbutane	33.51	32.28	35.70	34.95	33.16	33.91
2,2,3,3-Tetramethylbutane	45.00	43.45	45.82	45.04	44.20	44.98
Isopentane	28.56	27.51	29.37	29.32	28.75	28.81
Neopentane	29.05	28.01	29.44	29.21	28.84	29.07
2-Methylpentane	34.18	32.39	35.10	35.00	34.32	34.42
3-Ethylpentane	39.87	36.68	40.03	39.70	39.46	39.79
2,2-Dimethylpentane	40.07	38.16	41.55	41.46	40.04	40.13
2,3,4-Trimethylpentane	46.06	42.10	45.69	44.77	45.59	46.51
RMSE		2.53	2.49	2.41	0.85	0.90

^a Ref [47], experimental values calculated at 298.15 K.

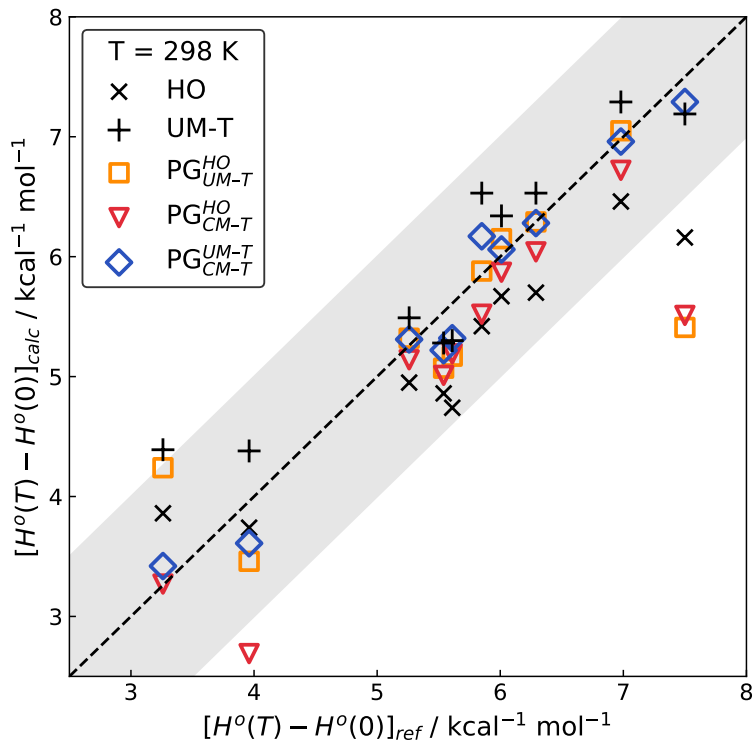


Figure 4: Parity plot for $H^\circ(T = 298\text{K}) - H^\circ(T = 0)$. The shaded region represents the interval of $\pm 1 \text{ kcal mol}^{-1}$ of the experimental value.

Table 2: Enthalpies $H^\circ(298 \text{ K}) - H^\circ(0 \text{ K})$ for Selected Molecules at 298.15 K (kcal mol^{-1})

Molecule	ref. ^a	HO	UM-T	$\text{PG}_{\text{UM-T}}^{\text{HO}}$	$\text{PG}_{\text{CM-T}}^{\text{HO}}$	$\text{PG}_{\text{CM-T}}^{\text{UM-T}}$
Ethylene glycol	3.96	3.74	4.38	3.46	2.69	3.61
Ethyl ether	5.61	4.74	5.30	5.17	5.19	5.32
Ethanedial	3.26	3.86	4.39	4.24	3.27	3.42
2,2-Dimethylbutane	6.01	5.67	6.34	6.15	5.87	6.06
2,3-Dimethylbutane	5.85	5.42	6.53	5.88	5.52	6.17
Isopentane	5.26	4.95	5.49	5.32	5.14	5.31
Neopentane	5.54	4.86	5.28	5.07	5.01	5.22
2-Methylpentane	6.29	5.70	6.53	6.29	6.04	6.28
3-Ethylpentane	7.50	6.16	7.19	5.41	5.51	7.29
2,2-Dimethylpentane	6.98	6.46	7.29	7.05	6.72	6.96
RMSE		0.67	0.50	0.78	0.79	0.22

^a Ref [47], experimental values calculated at 298.15 K. Results for 2,2,3,3-tetramethylbutane and 2,3,4-trimethylpentane are omitted for lack of reference data.

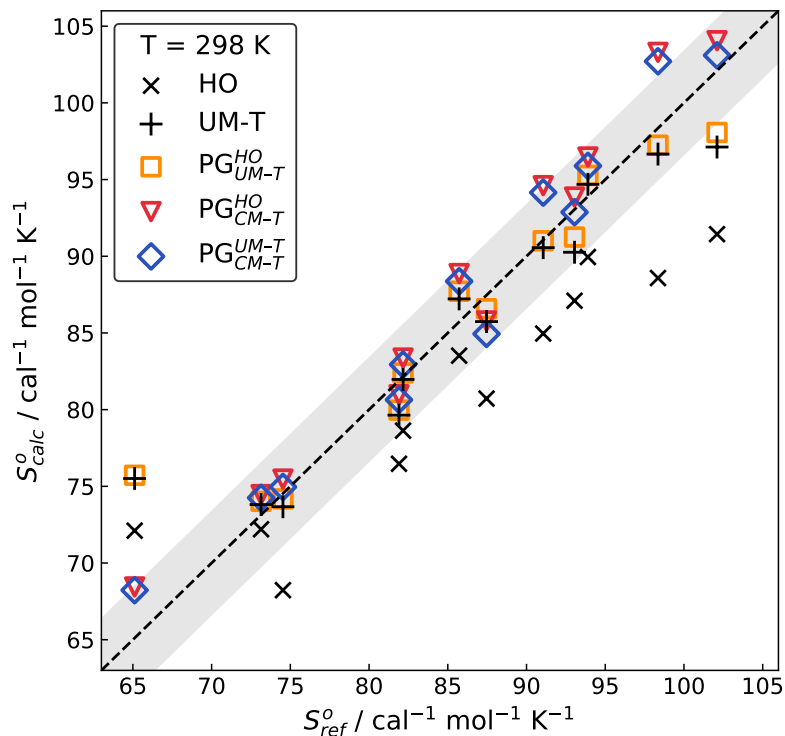


Figure 5: Parity plot for standard entropy. The shaded region represents the interval of $\pm 3.35 \text{ cal mol}^{-1} \text{ K}^{-1}$ ($\pm 1 \text{ kcal mol}^{-1}$ of TS at 298 K).

Table 3: Standard Entropies for Selected Molecules at 298.15 K ($\text{cal mol}^{-1} \text{ K}^{-1}$)

Molecule	ref. ^a	HO	UM-T	PG _{UM-T} ^{HO}	PG _{CM-T} ^{HO}	PG _{CM-T} ^{UM-T}
Ethylene glycol	74.53	68.23	73.67	74.18	75.47	74.96
Ethyl ether	81.90	76.47	79.64	79.97	80.97	80.64
Ethanedial	65.11	72.11	75.51	75.72	68.44	68.23
2,2-Dimethylbutane	85.72	83.52	87.22	87.72	88.87	88.37
2,3-Dimethylbutane	87.46	80.72	85.74	86.58	85.78	84.94
2,2,3,3-Tetramethylbutane	93.05	87.10	90.26	91.25	93.86	92.87
Isopentane	82.16	78.63	81.97	82.39	83.36	82.94
Neopentane	73.14	72.2	73.81	74.02	74.46	74.25
2-Methylpentane	91.06	84.96	90.56	91.01	94.60	94.15
3-Ethylpentane	98.35	88.58	96.66	97.24	103.3	102.7
2,2-Dimethylpentane	93.90	89.95	94.69	95.27	96.47	95.89
2,3,4-Trimethylpentane	102.1	91.44	97.12	98.07	104.1	103.1
RMSE		6.32	3.61	3.47	2.53	2.24

^a Ref [47], experimental values calculated at 298.15 K.

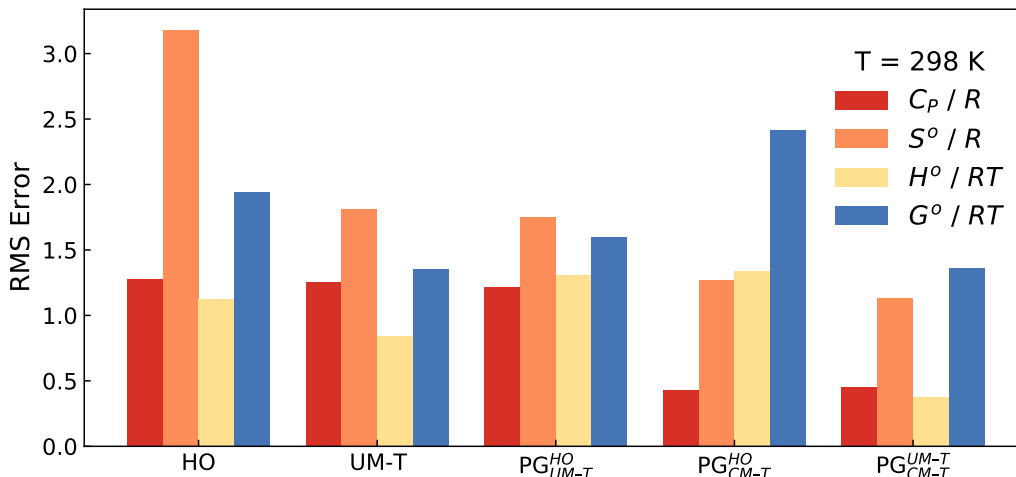


Figure 6: Root mean squared error across methods tested.

A summary of the RMSEs of the calculated heat capacity, entropy, enthalpy, and free energy is given in Figure 6 for all methods. For nearly all thermodynamic quantities calculated, the PG_{CM-T}^{UM-T} method outperforms harmonic, UM-T, PG_{UM-T}^{HO} , PG_{CM-T}^{HO} alternatives. The Gibbs free energy RMSEs of the PG_{CM-T}^{UM-T} method and the UM-T method are comparable, nevertheless, due to the compensation of generally overestimating the enthalpy with underestimating the entropy. For accuracy of enthalpy and entropy, PG_{CM-T}^{UM-T} remains the preferable method.

To analyze the temperature scaling of the PG_{CM-T}^{UM-T} method and Hamiltonian Monte Carlo relative to HO, UM-T, PG_{UM-T}^{HO} , and PG_{CM-T}^{HO} , we performed computations of the heat capacity using 2,2,3,3-tetramethylbutane as an example. We use the constant pressure heat capacity for these studies because, as opposed to other computable thermodynamic quantities, temperature dependence studies of heat capacities are widely available. The computed values for an array of temperatures were thus benchmarked against these reference data, which were obtained from the NIST database [47]. The results are given in Figure 7. We find that at intermediate temperatures, PG_{CM-T}^{UM-T} generally outperforms uncoupled mode approximations, and notably outperforms the HO approximation. As temperature increases, uncoupled methods and the HO method steer from the expected value, while coupled methods PG_{CM-T}^{HO} and PG_{CM-T}^{UM-T} remain relatively stable. At higher temperatures, however, the HMC sampling could not give convergent results in the amount of samples specified by our sampling protocol. This is attributed to the Hamiltonian Monte Carlo algorithm relying on the gradient scaled by β . When temperature increases, the energy distribution flattens, which is known to cause gradient-based methods to fail [48]. Because the NUTS algorithm relies on exploration of the PES through time integration of the potential gradient and chooses samples based on the criterion of a turn in trajectory, these low-gradient distributions cannot be reliably explored. Consequently, our implementation of the PG_{CM-T}^{HO} and PG_{CM-T}^{UM-T} methods, which use NUTS exclusively, is most useful at temperatures where the PES still has significant curvature. This could possibly be remedied by using non-gradient-based MCMC algorithms, increasing the number of samples, or adapting HMC for these edge cases, as has

been proposed to explore multimodal distributions with significant flat features [48].

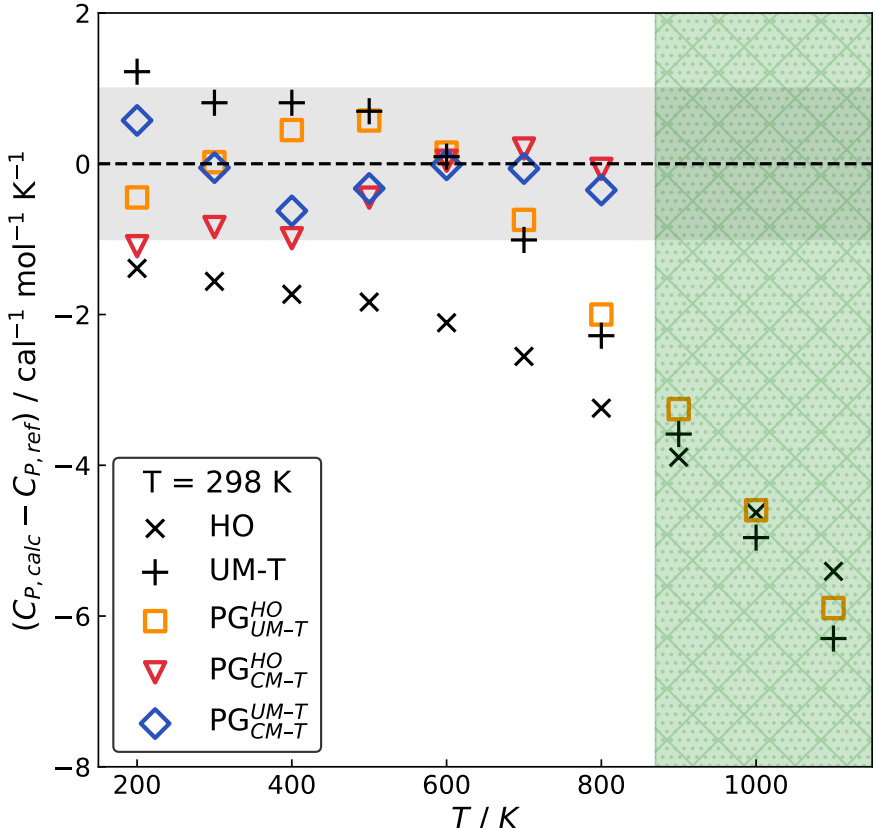


Figure 7: Temperature scaling of constant pressure heat capacity of 2,2,3,3-tetramethylbutane across HO, UM-T, PG_{UM-T}^{HO}, PG_{CM-T}^{HO}, and PG_{CM-T}^{UM-T} methods. The dashed line represents the expected value, with the gray shaded region representing the range of ± 1 cal mol⁻¹ K⁻¹. The hatched green region denotes the range where Hamiltonian Monte Carlo begins to become computational intractable.

We also examined the computational scaling of each method against each other. For a molecule with N torsions, a brute force finite-difference approach to sampling the potential energy surface at a 10° resolution scales as 36^N . This is in contrast to the uncoupled mode scaling of the HO, UM-T, and PG_{UM-T}^{HO} methods, which scales simply with N . For mode-coupled approximations (PG_{CM-T}^{HO} and PG_{CM-T}^{UM-T}) solved using HMC, we generally found the energies of all molecules to converge to values within $0.1 RT$ in fewer than 10,000 samples even in higher dimensional cases. The tradeoff, however, is that higher-dimensional systems can require more gradient calculations per sample. As a typical example, we examine the case of isobutane, which on average required 15 gradient calculations per sample using the NUTS algorithm. The convergence plot for isobutane as a function of the total number of computations, with $N = 4$ torsions, is given in Figure 8. From this analysis, we note that the

convergence of the energy is achieved between $36^{2.5}$ calculations, a significant improvement from 36^4 given by the finite-difference solution. In even the highest-dimensional systems, we experienced similar improvement, requiring on the order of 10^5 computations at most for convergence.

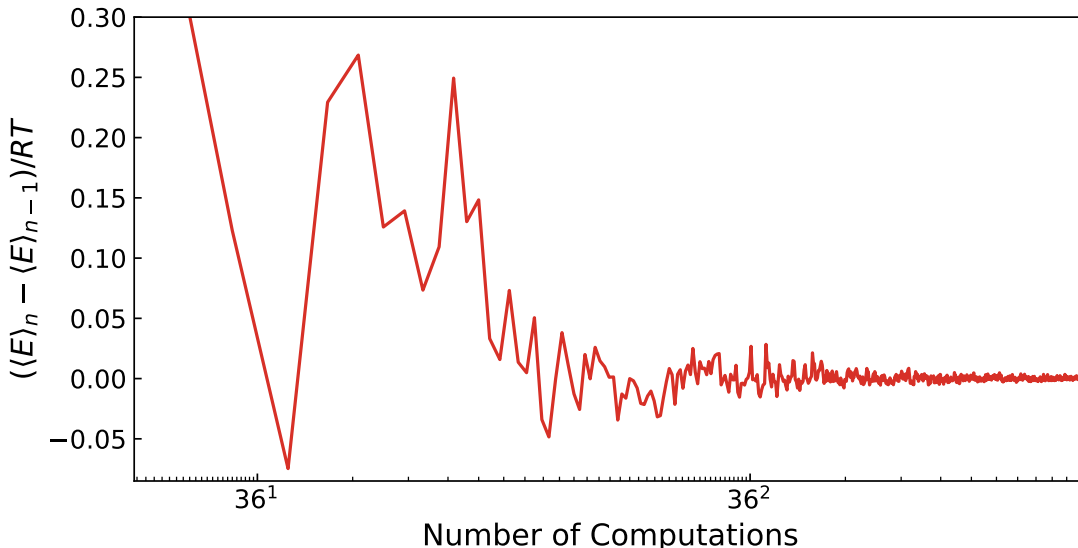


Figure 8: Convergence of the energy of isobutane as a function of number of computations. The number of computations includes both the single point energy samples at desired sampling configurations and the gradient calculations required in the HMC algorithm. The y -axis represents the difference in expected value for the n th iteration minus the previous iteration. The convergence criterion of $0.1 RT$ is given by the shaded region.

5. Conclusions

We have explored generalized Pitzer-Gwinn approximations for describing the thermodynamics of molecules with torsional mode coupling, for which a direct rigorous description is difficult. In its original formulation, the $\text{PG}_{\text{CM-T}}^{\text{HO}}$ was applied to torsional internal modes, for which it demonstrated remarkable improvement against the standard normal mode HO approximation. Continuing this work, Truhlar and coworkers have found similar improvement in obtaining improved PG corrections by using more informative harmonic reference potentials, such as that of the multiconformer HO model [19, 20, 21]. The method presented here uses an even more informative reference potential in the UM-T approximation, which describes the discrete mode uncoupled energy levels for explicitly anharmonic internal rotations. We have demonstrated that using this reference potential in the classical, torsional mode-coupled PES of molecules with significant torsional anharmonicity provides an accurate, computationally feasible way of assessing the coupling of torsions in gas phase molecules.

Benchmarked against accepted thermodynamic quantities, the $\text{PG}_{\text{CM-T}}^{\text{UM-T}}$ method generally

outperforms the HO, UM-T, $\text{PG}_{\text{UM-T}}^{\text{HO}}$, and $\text{PG}_{\text{CM-T}}^{\text{HO}}$ approximations for molecules with significant internal rotations. Harmonic approximations tend to underestimate the expected thermodynamic observables in large part because they fail to represent multiple degenerate conformers. While the UM-T and $\text{PG}_{\text{UM-T}}^{\text{HO}}$ methods account for the non-locality and anharmonicity of the PES, they too can yield substantial errors if internal rotations are treated as uncoupled. The $\text{PG}_{\text{CM-T}}^{\text{HO}}$ method includes mode coupling, but is still generally outperformed by $\text{PG}_{\text{CM-T}}^{\text{UM-T}}$, attributable to shortcomings in using the HO reference potential. The $\text{PG}_{\text{CM-T}}^{\text{UM-T}}$ method thus proves superior in predicting experimental observables because it accounts for the anharmonicity in the reference potential and coupling of internal rotations in its classical description of torsions. The superiority of $\text{PG}_{\text{CM-T}}^{\text{UM-T}}$ extends to the thermodynamics of gas-phase molecules at intermediate temperatures. Another key aspect of coupled mode PG methods is how efficiently the multidimensional configurational integral can be evaluated. For this purpose we employed the NUTS algorithm within Markov chain Monte Carlo. We found at standard conditions, the NUTS sampler requires substantially less computational time against the finite-difference solution. Beyond 800 K, however, the same effect could not be realized because of the unsuitability of the Hamiltonian Monte Carlo sampling method for flat distributions.

The generalizability of PG methods suggests that similar schemes can be employed for external degrees of freedom, such as hindered rotations and translations for which harmonic approximations are known to yield substantial underestimations of thermodynamic quantities, particularly the entropy. In this context, this study suggests thermodynamics may be improved by using a variational method to determine discrete energy levels in the uncoupled limit, employing a PG scheme with a reference solved variationally, and accounting for multiple rotational and translational conformers by sampling the external mode-coupled potential energy surface.

6. Acknowledgements

We thank Adam J.C. Wells for helpful discussions on statistics. This work was supported by the Director, Office of Science, Office of Basic Energy Sciences of the US Department of Energy under contract No. DE-AC02-05CH11231.

References

- [1] Stefan Grimme. “Semiempirical Hybrid Density Functional with Perturbative Second-Order Correlation”. *J. Chem. Phys.* **124.3** (2006), 034108
- [2] Urs Zimmerli, Michele Parrinello, and Petros Koumoutsakos. “Dispersion Corrections to Density Functionals for Water Aromatic Interactions”. *J. Chem. Phys.* **120.6** (2004), 2693–2699
- [3] János G Angyán et al. “Van der Waals Forces in Density Functional Theory: Perturbational Long-Range Electron-Interaction Corrections”. *Phys. Rev. A* **72.1** (2005), 012510

- [4] O Anatole Von Lilienfeld et al. “Optimization of Effective Atom Centered Potentials for London Dispersion Forces in Density Functional Theory”. *Phys. Rev. Lett.* **93**.15 (2004), 153004
- [5] 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**.7 (2009), 073005
- [6] Luis Simón-Carballido et al. “Anharmonicity of Coupled Torsions: the Extended Two-dimensional Torsion Method and its Use to Assess More Approximate Methods”. *J. Chem. Theory Comput.* **13**.8 (2017), 3478–3492
- [7] Berend Smit and Theo LM Maesen. “Molecular Simulations of Zeolites: Adsorption, Diffusion, and Shape Selectivity”. *Chemical Reviews* **108**.10 (2008), 4125–4184
- [8] Yi-Pei Li, Alexis T Bell, and Martin Head-Gordon. “Thermodynamics of Anharmonic Systems: Uncoupled Mode Approximations for Molecules”. *J. Chem. Theory Comput.* **12**.6 (2016), 2861–2870
- [9] Keith T Butler et al. “Experimental Evidence for Vibrational Entropy as Driving Parameter of Flexibility in the Metal–Organic Framework ZIF-4 (Zn)”. *Chem. Mater.* **31**.20 (2019), 8366–8372
- [10] Shao-Nung Huang et al. “Absolute Entropy and Energy of Carbon Dioxide Using the Two-Phase Thermodynamic Model”. *J. Chem. Theory Comput.* **7**.6 (2011), 1893–1901
- [11] Tian Lu and Qinxue Chen. “Shermo: A general code for calculating molecular thermochemistry properties”. *Comput. Theor. Chem.* **1200** (2021), 113249
- [12] Stefan Grimme. “Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory”. *Chem. Eur. J.* **18**.32 (2012), 9955–9964
- [13] Anthony P Scott and Leo Radom. “Harmonic vibrational frequencies: an evaluation of Hartree-Fock, Møller-Plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors”. *J. Phys. Chem.* **100**.41 (1996), 16502–16513
- [14] Mattias P Andersson and Per Uvdal. “New scale factors for harmonic vibrational frequencies using the B3LYP density functional method with the triple- ζ basis set 6-311+G (d, p)”. *J. Phys. Chem. A* **109**.12 (2005), 2937–2941
- [15] GiovanniMaria Piccini and Joachim Sauer. “Effect of Anharmonicity on Adsorption Thermodynamics”. *J. Chem. Theory Comput.* **10**.6 (June 2014), 2479–2487. ISSN: 1549-9618. DOI: [10.1021/ct500291x](https://doi.org/10.1021/ct500291x). URL: <https://doi.org/10.1021/ct500291x>
- [16] GiovanniMaria Piccini et al. “Accurate Adsorption Thermodynamics of Small Alkanes in Zeolites. Ab Initio Theory and Experiment for H-Chabazite”. *J. Phys. Chem. C* **119**.11 (2015), 6128–6137
- [17] GiovanniMaria Piccini and Joachim Sauer. “Quantum Chemical Free Energies: Structure Optimization and Vibrational Frequencies in Normal Modes”. *J. Chem. Theory Comput.* **9**.11 (Nov. 2013), 5038–5045. ISSN: 1549-9618. DOI: [10.1021/ct4005504](https://doi.org/10.1021/ct4005504). URL: <https://doi.org/10.1021/ct4005504>

- [18] Alice Chung-Phillips. “A Study of the Fourier-series Representation for Internal Rotation”. *J. Chem. Phys.* **88.3** (1988), 1764–1774
- [19] Benjamin A Ellingson et al. “Statistical Thermodynamics of Bond Torsional Modes: Tests of Separable, Almost-Separable, and Improved Pitzer–Gwinn Approximations”. *J. Chem. Phys.* **125.8** (Aug. 2006), 84305. ISSN: 0021-9606. DOI: [10.1063/1.2219441](https://doi.org/10.1063/1.2219441). URL: <https://doi.org/10.1063/1.2219441>
- [20] Vanessa Audette Lynch, Steven L Mielke, and Donald G Truhlar. “High-Precision Quantum Thermochemistry on Nonquasiharmonic Potentials: Converged Path-Integral Free Energies and a Systematically Convergent Family of Generalized PitzerGwinn Approximations”. *J. Phys. Chem. A* **109.44** (Nov. 2005), 10092–10099. ISSN: 1089-5639. DOI: [10.1021/jp051742n](https://doi.org/10.1021/jp051742n). URL: <https://doi.org/10.1021/jp051742n>
- [21] Jingjing Zheng and Donald G Truhlar. “Quantum Thermochemistry: Multistructural Method with Torsional Anharmonicity Based on a Coupled Torsional Potential”. *J. Chem. Theory Comput.* **9.3** (2013), 1356–1367
- [22] Kenneth S Pitzer and William D Gwinn. “Thermodynamic Functions for Molecules with Internal Rotation”. *J. Chem. Phys.* **9.6** (June 1941), 485–486. ISSN: 0021-9606. DOI: [10.1063/1.1750941](https://doi.org/10.1063/1.1750941). URL: <https://doi.org/10.1063/1.1750941>
- [23] Steve Brooks. *Handbook of Markov Chain Monte Carlo*. Boca Raton London: CRC Press, 2011. ISBN: 9781420079418
- [24] Matthew D Hoffman, Andrew Gelman, et al. “The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo.” *J. Mach. Learn. Res.* **15.1** (2014), 1593–1623
- [25] Junpeng Lao et al. *tfp.mcmc: Modern Markov Chain Monte Carlo Tools Built for Modern Hardware*. 2020. arXiv: [2002.01184](https://arxiv.org/abs/2002.01184) [stat.CO]
- [26] Simon Duane et al. “Hybrid Monte Carlo”. *Physics letters B* **195.2** (1987), 216–222
- [27] Michael Betancourt. *A Conceptual Introduction to Hamiltonian Monte Carlo*. 2018. arXiv: [1701.02434](https://arxiv.org/abs/1701.02434) [stat.ME]
- [28] Adrian Barbu. *Monte Carlo methods*. Singapore: Springer, 2020. ISBN: 978-981-13-2970-8
- [29] Bosiljka Njegic and Mark S Gordon. “Predicting Accurate Vibrational Frequencies for Highly Anharmonic Systems”. *J. Chem. Phys.* **129.16** (2008), 164107
- [30] John E Kilpatrick and Kenneth S Pitzer. “Energy Levels and Thermodynamic Functions for Molecules with Internal Rotation. III. Compound Rotation”. *J. Chem. Phys.* **17.11** (1949), 1064–1075
- [31] Maxwell L Eidinoff and JG Aston. “The Rotational Entropy of Nonrigid Polyatomic Molecules”. *J. Chem. Phys.* **3.7** (1935), 379–383
- [32] Jingjing Zheng and Donald G Truhlar. “Quantum Thermochemistry: Multistructural Method with Torsional Anharmonicity Based on a Coupled Torsional Potential”. *J. Chem. Theory Comput.* **9.3** (2013), 1356–1367

- [33] Edgar Bright Wilson, John Courtney Decius, and Paul C Cross. *Molecular Vibrations: the Theory of Infrared and Raman Vibrational Spectra*. Courier Corporation, 1980
- [34] John Salvatier, Thomas V Wiecki, and Christopher Fonnesbeck. “Probabilistic Programming in Python using PyMC3”. *PeerJ Computer Science* **2** (2016), e55
- [35] Andrew Gelman, Daniel Lee, and Jiqiang Guo. “Stan: A probabilistic programming language for Bayesian inference and optimization”. *J. Educ. Behav. Stat.* **40.5** (2015), 530–543
- [36] Bernard W Silverman. “Monographs on Statistics and Applied Probability”. *Density Estimation for Statistics and Data Analysis* **26** (1986)
- [37] Jan Beirlant et al. “Nonparametric Entropy Estimation: An Overview”. *Int. J. Math. Stat. Sci.* **6.1** (1997), 17–39
- [38] Johannes Kraml et al. “X-Entropy: A Parallelized Kernel Density Estimator with Automated Bandwidth Selection to Calculate Entropy”. *J. Chem. Inf. Model.* **61.4** (2021), 1533–1538
- [39] David W Scott. *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley & Sons, 2015
- [40] Shih-Cheng Li. *Automated Property Estimator*. URL: <https://github.com/shihchengli/APE>
- [41] Evgeny Epifanovsky et al. “Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package”. *J. Chem. Phys.* **155.8** (2021), 084801
- [42] Stefan Grimme. “Semiempirical GGA-type Density Functional Constructed with a Long-range Dispersion Correction”. *J. Comput. Chem.* **27.15** (2006), 1787–1799
- [43] Jeng-Da Chai and Martin Head-Gordon. “Long-range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections”. *Phys. Chem. Chem. Phys.* **10.44** (2008), 6615–6620
- [44] Mengjie Liu et al. “Reaction Mechanism Generator v3.0: Advances in Automatic Mechanism Generation”. *J. Chem. Inf. Model.* **61.6** (2021). PMID: 34048230, 2686–2696. DOI: [10.1021/acs.jcim.0c01480](https://doi.org/10.1021/acs.jcim.0c01480). eprint: <https://doi.org/10.1021/acs.jcim.0c01480>. URL: <https://doi.org/10.1021/acs.jcim.0c01480>
- [45] Bob Carpenter et al. “Stan: A Probabilistic Programming Language”. *J. Stat. Softw.* **76.1** (2017), 1–32
- [46] Pauli Virtanen et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. *Nature Methods* **17** (2020), 261–272. DOI: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2). URL: <https://rdcu.be/b08Wh>
- [47] Russell D Johnson III. “NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101” (2020)
- [48] Shiwei Lan, Jeffrey Streets, and Babak Shahbaba. “Wormhole Hamiltonian Monte Carlo”. *Proc. AAAI Conf. Artif. Intell.* Vol. 28. 1. 2014

Chapter IV

Effects of Confinement on the Estimation of Enthalpies and Entropies of Small Molecule Adsorption in Zeolites

ROTATIONAL and translational contributions to the thermodynamics of adsorption are investigated using density functional theory in a QM/MM scheme for species in H-CHA and H-MFI. Conventional harmonic approaches ill-enumerate molecular motion of species affixed to surfaces, neglecting the anharmonic and non-local complexity of hindered rotational and translational degrees of freedom. By contrast, our approach explicitly considers the full-dimensional potential energy surfaces associated with orientation (rotation) and position (translation) of rigid molecules confined by the zeolite walls and interacting with a binding site. For rotations, classical, semi-classical, and quantum partition functions are considered. These methods involve sampling the rotational potential energy surface using an Euler angle parametrization of adsorbate orientations. Wigner D-Matrix basis functions are used to describe the rotational potential energy surface, classical rotational partition functions are obtained by numerical quadrature, and quantum rotational partition functions are determined using the variational method. These approximations are used together with translational partition functions approximated by local harmonic approximations and non-local classical approximations, the latter which is determined by the accessible volume of the zeolite. The methods presented provide a more rigorous description of relevant modes of adsorbed species, leading to a more accurate prediction of the thermodynamic properties of adsorbed molecules. In particular, when we compare against experimental adsorption data, we find that anharmonic rotational and translational approximations are favored for weakly bound adsorbates such as small alkanes in H-CHA, suggesting non-local potential energy surfaces should be used to describe their thermodynamics. For strongly bound molecules such as hydrogen-bonded adsorbates methanol and ethanol in H-MFI, non-local rotational approximations yield predictions that are comparable to those obtained using harmonic potentials. Conversely, local translational approximations such as the harmonic approximation outperform the non-local translational approximations when compared against experimental adsorption data. We thus find anharmonic rigid rotor approximations capable of describing the thermodynamics of rotation for both extremes, and suggest improvement/generalization of anharmonic translational approximations through a more informed description of the

non-local potential energy surface.

1. Introduction

Zeolites are used as adsorbents and catalysts for a wide range of applications, including hydrocarbon synthesis [1, 2], carbon capture [3, 4], molecular sieving [5, 6], and pollution control [7, 8]. All of these applications share a common feature that one or more components are adsorbed within the zeolite, where, in the case of catalysis, they may undergo chemical transformations to form products. Hence, understanding the thermodynamics of adsorption in terms of the properties of the adsorbate and the adsorbent, i.e., the zeolite, becomes highly important. Great progress has been achieved in determining the thermodynamics of weak molecular adsorption (physisorption) using classical potentials (e.g. Lennard-Jones and Coulombic potentials) and Monte Carlo or molecular dynamics methods [9, 10]. However, these approaches depend on the parametrization of potentials to fit measured enthalpies and entropies of adsorption. While the parameters thus obtained are transferable within a class of molecules (such as hydrocarbons), it is neither possible to readily determine how these parameters will change with changes in the adsorbate composition and structure, nor with changes in the composition and structure of the zeolite. A more rigorous and satisfying approach would be to use electronic structure methods, such as density functional theory (DFT) to determine the energy of adsorption at 0 K. However, generalization of electronic structure to finite temperature remains among the most difficult challenges towards developing a full ab initio description of the thermodynamics, namely the enthalpy and entropy of adsorption. Significant progress along these lines has been made, and today enthalpies of adsorption can be determined by such means with chemical accuracy ($\pm 5 \text{ kJ mol}^{-1}$ [11]) [9, 12, 13, 14]. Despite this, reliable methods for determining the entropy of adsorption are still under debate [12, 15, 16, 17, 18, 19, 20, 21].

The frequently-invoked harmonic oscillator (HO) approximation has been found to systematically underestimate the entropy of adsorption [12, 21, 22]. In this approximation, the internal motions in a system are represented as a collection of “normal modes” with localized, quadratic potential energies. Moreover, it should be recognized that the modes with the lowest vibrational frequencies carry the largest entropic contributions. Consequently, small errors in the predicted frequency consequently lead to large errors in the predicted thermodynamics. Such is the case for hindered translational and hindered rotational degrees of freedom, which tend to have low vibrational frequencies. For computational studies on systems whose reaction kinetics depend on adsorption free energies, the accuracy of available methods is therefore extremely limited.

Attempts have been made – often with good success – to correct entropies obtained by the HO approximation through quasi-harmonic interpolation approaches [9, 18] and anharmonic sampling along normal mode coordinates [19, 20]. However, these approaches also fail to account for the potentially non-local energy landscape associated with frustrated rotations and translations of adsorbates interacting with a binding site. This problem is reminiscent of gas phase internal rotations, where frustrated intramolecular rotations of chemical moieties are particularly anharmonic and non-local. Past contributions addressing the anharmonicity of these modes using uncoupled mode and Pitzer-Gwinn approximations have found that

local potential energy approximations, including anharmonic internal coordinate sampling, yield poor predictions of the gas phase thermodynamics when compared against experimental measurements [23]. Instead, better results are obtained through explicitly sampling along the dihedral coordinates of internal rotations, and leaving local approximations to the less thermodynamically-relevant vibrational modes.

In light of this observation, a first principles approach to surface chemistry should include direct enumeration of rotational and translational anharmonicity. In principle, this is possible through sampling the full-dimensional potential energy surface using molecular dynamics, Monte Carlo methods, or hybrid approaches. However, these methods can require millions of samples in order to reach convergence, and rapidly becomes intractable for large systems, especially at rigorous levels of electronic structure theory [24]. This motivates us to pursue general uncoupled descriptions for the potential energy surfaces of adsorbate rotations and translations, from which direct calculation of thermodynamic properties can be derived straightforwardly.

In this work, two uncoupled mode approximations are investigated: one for non-local rotations and one for non-local translations. The uncoupled rotational PES was found by treating the adsorbate as a rigid rotor. This is done by sampling rotational motion under an Euler angle parametrization of adsorbate orientations. A representation of the PES was then obtained by fitting the samples to the SO(3) Fourier series [25]. This allowed us to obtain the classical, quantum, and harmonic-reference Pitzer-Gwinn partition functions for rigid rotation. For translations, a non-local representation of the translational partition function was obtained using a description of the occupiable volume of unit cell of the zeolite. Using a previously-developed QM/MM framework for zeolites, adsorption thermodynamics of methane and ethane in H-CHA, and propane, methanol, and ethanol in H-MFI are considered and benchmarked against experimental references. We find considerable promise in directly modeling adsorbate rotations and non-local translations for relatively weakly-bound species, where substantial entropic contributions are retained from the gas phase. We find similar promise in direct rotational modeling for more strongly-bound species, such as hydrogen-bonded species, wherein multiple stable H-bond configurations can be considered compared to strictly local approximations. Additionally, we find translational approximations for strongly bound species to be best described by local approximations such as the HO approximation, while weakly bound species require explicit enumeration of translational freedom under a more rigorous approximation.

2. Theory and Methods

The thermodynamics of adsorption are derived from a description of the canonical partition function, which is an enumeration of the possible energetic states the adsorbate can occupy in the canonical ensemble. Classically, we solve this problem by taking an integral over the Boltzmann-distributed energies, leading to one momentum and one configurational integral

$$\begin{aligned}
 Q^{cl} &= \frac{1}{h^3} \int d^3p \int d^3q e^{-\beta H(p,q)} \\
 &= \frac{1}{h^3} \int d^3p e^{-\beta p \cdot p / 2m} \int d^3q e^{-\beta V(q)}
 \end{aligned}
 \tag{4.1}$$

where H is the Hamiltonian, p denotes the momentum of the adsorbate (interchangeable with angular momentum L for angular motion), q is the parametrization of the relevant motion in space, m is the mass of the adsorbate (or moment of inertia I), and V is the potential energy. Similarly, the quantum partition function is solved by summation over the energy eigenvalues of the Hamiltonian operator

$$\hat{H} |\Phi_j\rangle = \left[\frac{\hat{p}^2}{2m} + \hat{V} \right] |\Phi_j\rangle = E_j |\Phi_j\rangle \quad (4.2)$$

$$Q^{qu} = \sum_j g_j e^{-\beta E_j} \quad (4.3)$$

where E_j is the energy of the j th eigenstate $|\Phi_j\rangle$, and g_j is the degeneracy of the same.

The energy, and therefore the overall Hamiltonian, is often decomposed into nuclear and electronic contributions under the Born-Oppenheimer approximation, the former further decomposed into rotational, translational, and internal contributions. Consequently, the overall partition function is given as a product of these individual contributions, from which the thermodynamics of each can be solved independently.

$$E_{\text{tot}} = E_{\text{trans}} + E_{\text{rot}} + E_{\text{int}} + E_{\text{elec}} \quad (4.4)$$

$$Q_{\text{tot}} = Q_{\text{trans}} Q_{\text{rot}} Q_{\text{int}} Q_{\text{elec}} \quad (4.5)$$

Relevant thermodynamic state functions follow from standard thermodynamic relations derived from the partition functions,

$$A = -\frac{1}{\beta} \log Q \quad (4.6)$$

$$E = -\frac{\partial \log Q}{\partial \beta} \quad (4.7)$$

$$H = E + P\mathcal{V} \quad (4.8)$$

$$S = \frac{E - A}{T} \quad (4.9)$$

$$G = H - TS \quad (4.10)$$

where A , E , H , S , and G are respectively the Helmholtz free energy, average energy, enthalpy, entropy, and Gibbs free energy associated with the relevant motion whose partition function is Q . The parameters P and \mathcal{V} are the pressure and volume of the system. In the gas phase, the product $P\mathcal{V}$ is equivalent to the thermodynamic temperature β^{-1} .

In this work, the electronic energy contribution is solved by density functional theory, which considers only the ground state energy. The electronic partition function is not relevant because the energies are always assumed to be in the ground state, and the thermal contributions are contained within the other degrees of freedom. This is a good approximation in the low temperature limit, when electronic excitations can be neglected. Additionally, internal motions such as vibrations and internal rotations are approximated to be mostly retained from the gas phase to the adsorbed phase. Adsorption thermodynamics are determined by taking the difference between the adsorbed and gas phase quantities, resulting in

the cancellation of intramolecular contributions and leaving purely external rotational and translational contributions. The resulting adsorption energies, enthalpies, and entropies are

$$\Delta E_{\text{ads}} = \Delta E_{\text{trans}} + \Delta E_{\text{rot}} + \Delta E_{\text{elec}} \quad (4.11)$$

$$\Delta H_{\text{ads}} = \Delta E_{\text{ads}} + P\Delta\mathcal{V} \quad (4.12)$$

$$\Delta S_{\text{ads}} = \Delta S_{\text{trans}} + \Delta S_{\text{rot}} \quad (4.13)$$

The next sections will detail how electronic, rotational, and translational contributions are derived from first principles for models of molecules adsorbed in zeolites.

2.1. Molecular Model and Energy Calculations

The combinations of adsorbate and zeolite were chosen based on (1) the availability of experimental data against which to compare, namely the enthalpy and entropy, and (2), the presence of features that influence adsorption strength. While there is abundant data on the adsorption thermodynamics of weakly-bound alkanes in zeolites, there is a relative dearth on strongly-bound species, such as oxygenated compounds that interact with binding sites by hydrogen bonding. The generality of the approach developed herein is tested at both extremes. The chosen candidates are methane and ethane in H-chabazite, for which we expect to observe confinement effects, and methanol, ethanol, and propane in H-MFI, for which we expect to observe varying degrees of binding strength.

The charge-compensating proton accompanying a framework aluminum atom in H-CHA and H-MFI is considered as the primary binding site. In CHA, the acidic proton extends into pore vacuum of one of its straight channels, while in MFI, it extends into the intersection of two channels. In modeling both structures, the QM region was chosen to encompass as many of these features as computationally feasible. The extended zeolite framework surrounding the chosen QM region constitutes the MM region. These structures are depicted in Figures 1 and 2.

Adsorption geometries were generated by geometry optimization of each substrate within the QM/MM framework, with full relaxation of the QM region. Electronic structure calculations were performed using the ω B97X-D density functional [26] and the def2-SVP basis set [27], with non-zeolite MM contributions represented by the CHARMM force field [28], and framework MM contributions represented by improved Lennard-Jones parameters optimized for zeolites [9]. Thermal analysis, including potential energy sampling described in the sections below, was done at the same level of theory. Refined electronic energies for optimized geometries were performed at the more rigorous ω B97M-V / def2-TZVP level of theory [27, 29]. All electronic energy calculations were performed using a developmental version of the Q-Chem software package [30].

2.2. Adsorbate Rotations

Parametrization of a Molecular Top

The three-dimensional rotation group $\text{SO}(3)$ describes all rotations about the origin of three dimensional Euclidean space \mathbb{R}^3 , and has several candidate parametrization schemes, including Euler angles, Tait-Bryan angles, and quaternions [31]. We employ an Euler angle

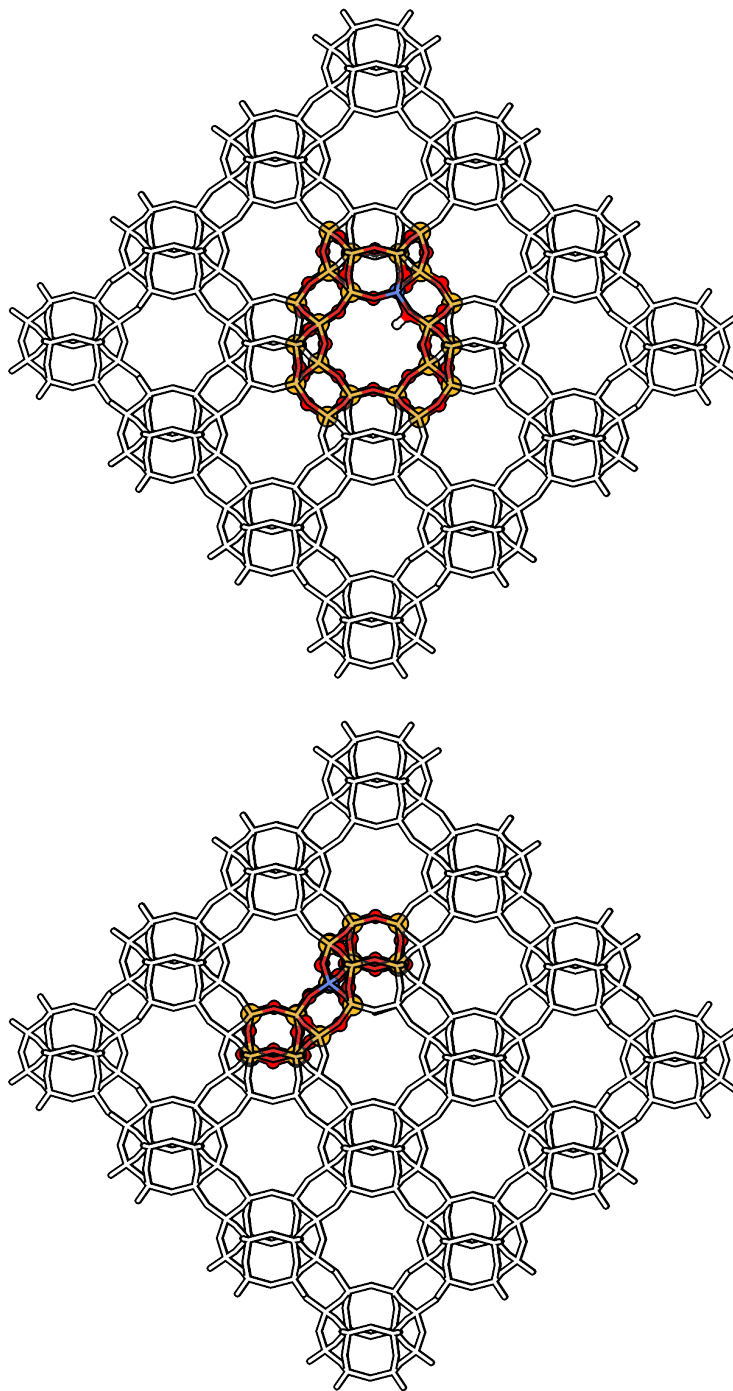


Figure 1: QM/MM Model for H-CHA. Yellow, red, blue, and white molecules represent silicon, oxygen, aluminum, and hydrogen atoms in the QM region. Gray silhouetted molecules represent the encompassing silica MM region. The top orthographic projection highlights the acidic proton and proximal aluminum dopant viewed down the straight channel of the structure. The bottom view depicts the QM region viewed from the side. Both structures are equivalent.

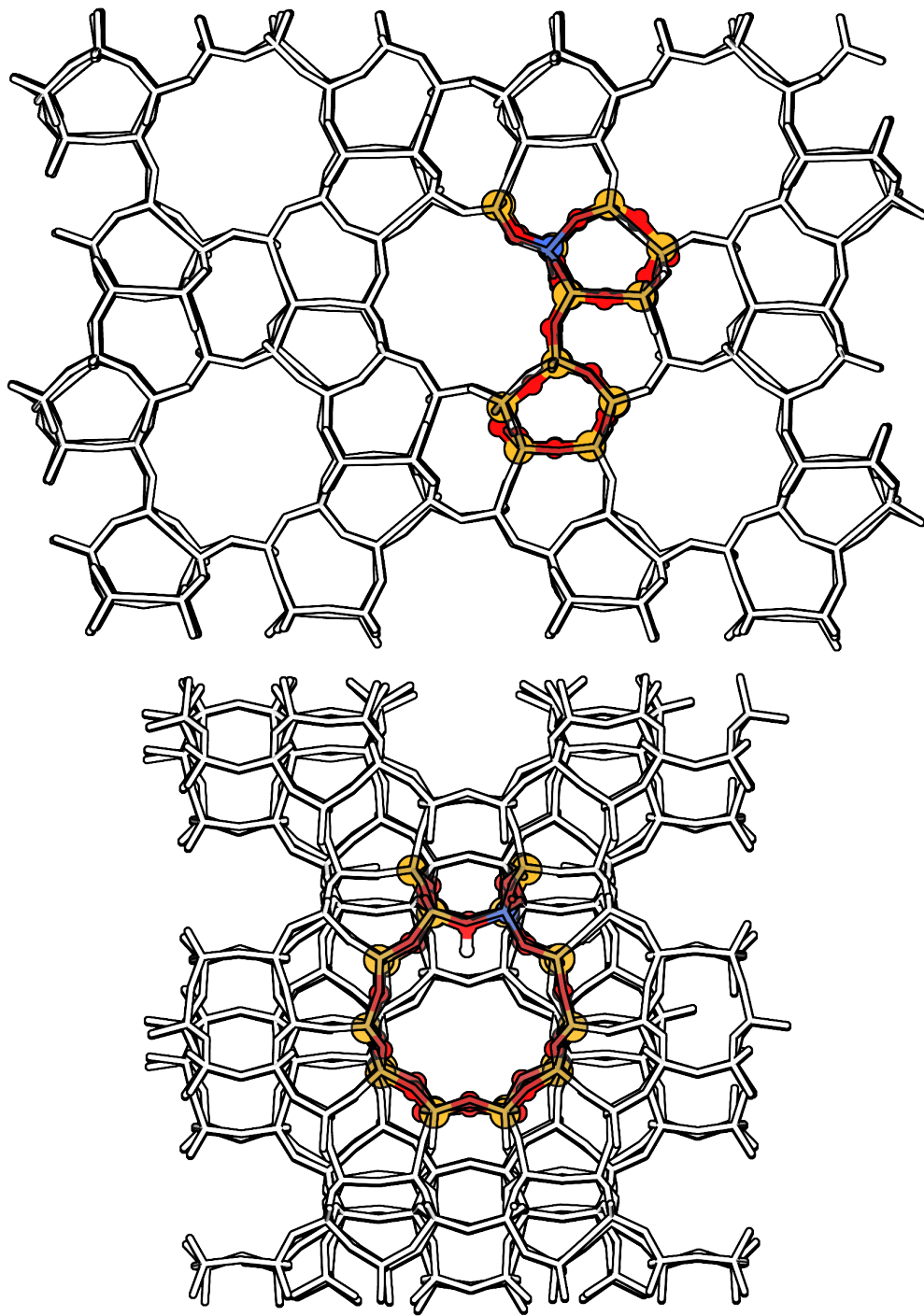


Figure 2: QM/MM Model for H-MFI. Yellow, red, blue, and white molecules represent silicon, oxygen, aluminum, and hydrogen atoms in the QM region. Gray silhouetted molecules represent the encompassing silica MM region. The top orthographic projection is viewed down the straight channel pore of the structure. The bottom view depicts the QM region viewed from the side, highlighting the acidic proton and proximal aluminum dopant. Both structures are equivalent.

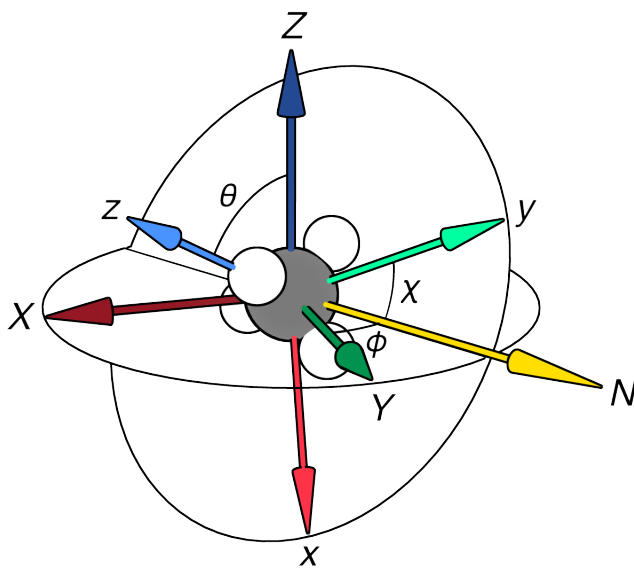


Figure 3: Geometric Interpretation of the Euler Angle Parametrization of a Molecular Top. Dark red, green, and blue arrows represent laboratory frame axes. Light red, green, and blue counterparts represent the molecular frame axes. The yellow vector represents the line of nodes. Azimuthal (ϕ), polar (θ), and intrinsic rotation (χ) angles are labeled accordingly.

parametrization, which is well-documented in angular momentum and harmonic analysis texts [25, 32, 33, 34]. Under this parametrization, the orientation of a rigid body is described by elemental rotations by angles ϕ , θ , and χ , with respect to a fixed coordinate system. For rotating molecules, the fixed coordinate system, often referred to as the laboratory frame, is denoted by XYZ , and the body-fixed frame, also known as the "molecular frame", by xyz . We employ the ZYZ convention shared by Zare [32], Kroto [33], and Brink and Satchler [34], wherein each net rotation is the result of three chained rotations by χ about Z , θ about Y , and ϕ about Z , in that order. This is an equivalent parametrization to the $zy'z''$ convention, where rotations are instead referenced to the molecular axes (ϕ about z , θ about y' , and χ about z'' , in that order). The equivalence is shown in Appendix C. Linear molecules contain only two rotational degrees of freedom, leading to a simple spherical parametrization of its rotations. In this chapter, we consider rotations of non-linear molecules, which are more general.

Potential Energy Surface

Determining the thermodynamics quantities for our systems of interest first requires a description of the potential energy surface. Computationally, this must be achieved by discretely sampling energies on the relevant domain, followed by numerical fitting, interpolation, or series expansion of sampled points. Previous studies solving the one-dimensional potential energy surfaces of internal rotations, for example, begin by sampling a uniform discretization of the dihedral angle. The resulting discrete function is fitted to a Fourier series or spline

interpolated [23, 35].

Analogously, obtaining the rotational potential energy surface begins by sampling a discrete set of adsorbate orientations which are obtained by rigid rotation of the atomic coordinates about the molecular center of mass. In the Euler angle parametrization scheme we adopt, angles ϕ and θ together span a spherical domain, and χ spans a 2π -periodic domain. A convenient discretization scheme should combine a spherical discretization for ϕ and θ parameters with a uniform discretization for χ . For spherical discretizations, popular methods include Gauss-Legendre [36, 37], Hierarchical Equal Area isoLatitude Pixelation (HEALPix) [38, 39], Chebyshev nodes [40, 41], and Lebedev grids [42, 43, 44]. For ease of numerical integration, we chose a Lebedev discretization of ϕ and θ parameters, for which the number and location of the grid points and set of numerical quadrature weights are determined by enforcing the exact integration of spherical harmonics up to a given order. The χ angular resolution was chosen to be approximately the average angular resolution of ϕ and θ , with the number of χ sampling points determined by $N_\chi = \lfloor \sqrt{\pi N_{\theta,\phi}} \rfloor$. A pictorial representation of this discretization scheme is given in Figure 4. Computation of the single-point energies at each point represents the computational bottleneck – accordingly, this method scales with $N_{\theta,\phi} \sqrt{N_{\theta,\phi}} / N_{\text{proc}}$, where N_{proc} is the number of processors that can independently perform the calculation at a gridpoint in parallel. The Lebedev grid and its associated integration weights were generated using the Numgrid library [45], with orientations achieved using a quaternionic rotation code with the origin defined as the orientation of the geometric minimum. We examined the effects of successively increasing Lebedev grid sizes with $N_{\theta,\phi} = 50, 74,$ and 110 quadrature points, finding 74 Lebedev quadrature points to be sufficient for modeling the rotational PES (see Table C.1).

A continuous representation of the potential from discretely sampled points is obtained by a harmonic expansion in the appropriate basis. If the adsorbate were a linear rotor that could be simply parameterized by ϕ and θ , we could expand the potential in the basis of the spherical harmonics,

$$V(\theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \hat{v}_\ell^m Y_\ell^m(\theta, \phi). \quad (4.14)$$

The presence of a third inertial axis in non-linear rotors requires a representation that includes the additional χ angular parameter. This is possible by introducing an arbitrary unitary rotation operator consistent with the ZYZ Euler angle convention previously detailed

$$\hat{R}(\phi, \theta, \chi) = e^{-i\phi \hat{L}_Z} e^{-i\theta \hat{L}_Y} e^{-i\chi \hat{L}_Z} \quad (4.15)$$

where \hat{L}_Y and \hat{L}_Z are the angular momentum operators for rotations about the laboratory frame Y and Z axes, respectively. From angular momentum eigenvalue relations discussed in Appendix C, it immediately follows that rotations in the spherical basis are described by the elements of the rotation matrix, well-known as the Wigner D-Matrix

$$D_{mk}^\ell(\phi, \theta, \chi) \equiv \langle \ell m | \hat{R}(\phi, \theta, \chi) | \ell k \rangle = e^{-im\phi} d_{mk}^\ell(\theta) e^{-ik\chi} \quad (4.16)$$

$$d_{mk}^\ell(\theta) \equiv \langle \ell m | e^{-i\theta \hat{J}_y} | \ell k \rangle = \sqrt{\frac{(\ell+k)!(\ell-k)!}{(\ell+m)!(\ell-m)!}} \left(\sin \frac{\theta}{2} \right)^{k-m} \left(\cos \frac{\theta}{2} \right)^{k+m} P_{\ell-k}^{(k-m, k+m)}(\cos \theta) \quad (4.17)$$

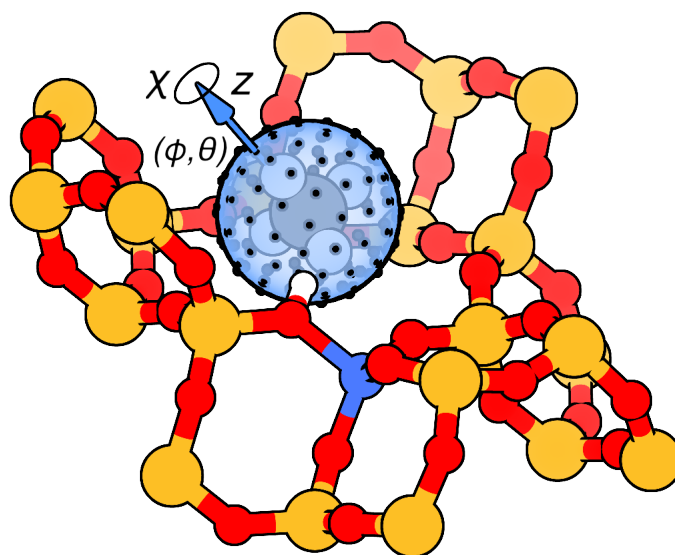


Figure 4: Schematic Representation of Potential Energy of Adsorbed Molecular Species Sampled on $SO(3)$. Pictured: centroid-represented, oriented methane embedded in the QM region. For clarity, the MM region is not pictured. Azimuthal and polar angles are discretely sampled on a Lebedev grid. Intrinsic spin angles are sampled at the same average angular resolution as the grid.

The function $d_{mk}^\ell(\theta)$ of Equation 4.17 is the reduced rotation matrix, also known as the Wigner (small) d-matrix, which is further expressed as a function of the Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$, which are a generalization of the Legendre polynomials.

Wigner D-Matrix elements form a complete set of orthogonal functions, with

$$\int_0^{2\pi} d\phi \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\chi D_{m'k'}^{\ell'*}(\phi, \theta, \chi) D_{mk}^\ell(\phi, \theta, \chi) = \frac{8\pi^2}{2\ell+1} \delta_{\ell'\ell} \delta_{m'm} \delta_{k'k} \quad (4.18)$$

Thus, any function of ϕ , θ , and χ can be expanded in the basis of Wigner D-Matrix elements to the bandlimit L_{\max} , which is the maximum degree ℓ for which the expansion coefficients are non-zero.

$$V(\phi, \theta, \chi) = \sum_{\ell=0}^{L_{\max}} \sum_{m=-\ell}^{\ell} \sum_{k=-\ell}^{\ell} \hat{v}_{mk}^\ell D_{mk}^\ell(\phi, \theta, \chi) \quad (4.19)$$

$$\hat{v}_{mk}^\ell = \frac{2\ell+1}{8\pi^2} \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\chi V(\phi, \theta, \chi) D_{mk}^{\ell*}(\phi, \theta, \chi) \quad (4.20)$$

The expansion coefficients relate to V by the SO(3) Fourier transform (SOFT) and are solved by evaluating the inner product given in Equation 4.20. This can be solved using quadrature rules for our discretization scheme. However, determining an appropriate bandwidth has been shown to be a challenging task, with increasing bandwidth often leading to diverging results or aliasing [25]. We instead cast SOFT(V) as the matrix vector product

$$\mathfrak{D} \hat{\mathbf{v}} = \mathbf{v} \quad (4.21)$$

where \mathfrak{D} is a matrix of all Wigner D-Matrix elements (up to the maximum order) at every sampling position. The rows of \mathfrak{D} are indexed by i , and correspond to the position of the sampling point $(\phi_i, \theta_i, \chi_i)$, while the columns of \mathfrak{D} are indexed by j , which is uniquely mapped to the lmk index by $j = \frac{4}{3}\ell^3 + 2\ell^2 + \frac{5}{3}\ell + 2\ell m + k$. \mathfrak{D}_{ij} is therefore the Wigner D-Matrix element at the lmk index corresponding to j , evaluated at the sampling point $(\phi_i, \theta_i, \chi_i)$. The dimensions of \mathfrak{D} are $N_{\phi,\theta,\chi} \times N_{lmk}$, where $N_{\phi,\theta,\chi} = N_{\theta,\phi} N_{\chi}$ and $N_{lmk} = \frac{1}{3}(\ell_{\max}+1)(2\ell_{\max}+1)(2\ell_{\max}+3)$. The vector $\hat{\mathbf{v}}$ is the vector of \hat{v}_{mk}^ℓ Fourier coefficients indexed by j , and the vector \mathbf{v} is the vector of sampled potential energies indexed by i , in other words $\mathbf{v}_i = V(\phi_i, \theta_i, \chi_i)$. Because \mathfrak{D} is generally not invertible, the Fourier coefficients are solved by the Moore-Penrose pseudoinverse (denoted by superscript $+$) through its singular value decomposition $\mathfrak{D} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\dagger$.

$$\begin{aligned} \hat{\mathbf{v}} &= \mathfrak{D}^+ \mathbf{v} \\ &= \mathbf{V}\mathbf{\Sigma}^+ \mathbf{U}^\dagger \mathbf{v} \end{aligned} \quad (4.22)$$

The factorization features the complex unitary matrices \mathbf{U} and \mathbf{V} , which have dimensions $N_{\phi,\theta,\chi} \times N_{\phi,\theta,\chi}$ and $N_{lmk} \times N_{lmk}$, respectively. The matrix $\mathbf{\Sigma}$ is rectangular diagonal with dimensions $N_{\phi,\theta,\chi} \times N_{lmk}$, whose diagonal entries $\varsigma_i = \Sigma_{ii}$ are the so-called singular values. $\mathbf{\Sigma}^+$ is easily solved by replacing the non-zero diagonal elements by their reciprocals and transposing the resulting matrix.

The Wigner D-Matrix functions used for fitting were calculated using open access code for spherical harmonic functions [46]. The SOFT coefficients for each system sampled under

this method were iteratively solved up to a coefficient of determination of $R^2 > 0.99$. This criterion was generally met at a maximum order $\ell_{\max} < 10$ for potential energy surfaces of molecules in this work.

Partition Functions

The classical partition function for adsorbate rigid rotation (RR) are given by

$$Q_{\text{rot}}^{\text{cl,RR}} = \frac{1}{\sigma} \sqrt{\frac{I_x I_y I_z}{(2\pi\beta\hbar)^3}} \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\chi e^{-\beta V(\phi,\theta,\chi)} \quad (4.23)$$

which is derived from Equation 4.1. In this expression, I_x is the moment of inertia about the molecular x axis, I_y about the y axis, and so forth. The parameter σ is the rotational symmetry number for our parametrization, and $V(\phi, \theta, \chi)$ is the potential energy given by Equation 4.19. The subscript RR denotes that this is the partition function under the rigid rotor approximation. In the limit of gas phase free rotation (i.e. zero potential), this equation reduces to the well-known free rotor (FR) partition function for non-linear rotors,

$$Q_{\text{rot}}^{\text{cl,FR}} = \frac{\sqrt{\pi}}{\sigma} \sqrt{\frac{(k_B T)^3}{ABC}} \quad (4.24)$$

where A , B , and C are the rotational constants associated with the molecular y , z , and x axes, respectively, a convention adapted from Kennerly [47]. The rotational symmetry numbers and rotational constants were determined using the Reaction Mechanism Generator (RMGPy) software package [48], which determines these quantities through analysis of the atomic coordinates of the adsorbate. The integral in Equation 4.23 is most conveniently solved using numerical quadrature for which integration weights w are known for Lebedev discretizations. The result is given by

$$\int_0^{2\pi} d\phi \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\chi e^{-\beta V(\phi,\theta,\chi)} = \frac{8\pi^2}{M_\chi} \sum_{p=1}^{M_\chi} \sum_{q=1}^{M_{\theta,\phi}} w_q e^{-\beta V(\phi_q,\theta_q,\chi_p)} \quad (4.25)$$

where M_χ and $M_{\theta,\phi}$ were chosen such that the integration grid is much finer than the sampling grid. Generation of grid points and integration weights was again facilitated by the Numgrid library [45], and Wigner D-Matrix functions are computed using spherical function software [46].

Pitzer-Gwinn [49] and variational solutions for rigid rotation were also considered. The former was originally formulated to describe the quantum effects of molecular internal rotations. This involved correcting classical torsional potential energies by the difference in the quantum and classical potential energies of a harmonic oscillator reference. In terms of partition functions, this results in the following expression for a given one dimensional torsion,

$$Q_{\text{tors}}^{\text{PGHO}} = \frac{Q_{\text{tors}}^{\text{qu,HO}}(\omega_\tau)}{Q_{\text{tors}}^{\text{cl,HO}}(\omega_\tau)} Q_{\text{tors}}^{\text{cl}} \quad (4.26)$$

The quantum and classical HO partition functions are functions of the normal mode frequency corresponding to the internal rotation, ω_τ . The expression for the quantum and classical HO partition functions are, for a general frequency ω ,

$$Q^{qu,HO}(\omega) = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}} \quad (4.27)$$

$$Q^{cl,HO}(\omega) = (\beta\hbar\omega)^{-1} \quad (4.28)$$

In the limit that the classical torsional potential resembles the reference classical HO, the quantum HO thermodynamics are thus recovered. The rotational analog employs a rigid rotor classical potential, keeping the HO reference correction factor.

$$Q_{\text{rot}}^{\text{PG}_{\text{RR}}^{\text{HO}}} = \frac{Q^{qu,HO}(\omega_\phi) Q^{qu,HO}(\omega_\theta) Q^{qu,HO}(\omega_\chi)}{Q^{cl,HO}(\omega_\phi) Q^{cl,HO}(\omega_\theta) Q^{cl,HO}(\omega_\chi)} Q_{\text{rot}}^{cl,RR} \quad (4.29)$$

Because non-linear rigid rotations are three-dimensional, harmonic reference partition functions are products of each one-dimensional normal mode frequency corresponding to rotational motion, ω_ϕ , ω_θ , ω_χ . These are calculated by considering only a partial Hessian that includes the atoms of the adsorbed species.

On the other hand, the variational solution to rigid rotation is Hessian-free, solved with normalized Wigner D-Matrix element basis functions given by

$$\psi_{mk}^\ell(\phi, \theta, \chi) \equiv \langle \phi, \theta, \chi | \ell m k \rangle = \sqrt{\frac{2\ell+1}{8\pi^2}} D_{mk}^{\ell*}(\phi, \theta, \chi) \quad (4.30)$$

The elements of the Hamiltonian matrix in this basis is constructed by the following expression, and then diagonalized to obtain the energy eigenvalues to solve Equation 4.3:

$$\begin{aligned} \langle \ell' m' k' | \hat{H} | \ell m k \rangle &= \frac{1}{2} [(A+C)\ell(\ell+1) + (A-C)\kappa k^2] \delta_{\ell'\ell} \delta_{m'm} \delta_{k'k} \\ &+ \frac{1}{4} (C-A) \sqrt{\ell(\ell+1) - k(k\pm 1)} \sqrt{\ell(\ell+1) - (k\pm 1)(k\pm 2)} \delta_{\ell'\ell} \delta_{m'm} \delta_{k'k\pm 2} \\ &+ \sqrt{\frac{2\ell'+1}{2\ell+1}} \sum_{\ell''} \sum_{m''} \sum_{k''} \hat{v}_{m''k''}^{\ell''} \langle \ell' m' \ell'' m'' | \ell m \rangle \langle \ell' k' \ell'' k'' | \ell k \rangle \end{aligned} \quad (4.31)$$

where $\kappa = \frac{2B-(A+C)}{A-C}$, A , B , and C are the rotational constants in the previously-defined convention, \hat{v}_{mk}^ℓ are the Fourier coefficients of Equation 4.20, and $\langle \ell' m' \ell'' m'' | \ell m \rangle$ are the Clebsch-Gordan coefficients. A derivation of the Hamiltonian matrix is given in Appendix C. The Hamiltonian matrix was constructed under an OpenMP parallelization scheme [50], with Clebsch-Gordan coefficients calculated using a C++ library by Dumont [51]. Eigen-decomposition was performed using Armadillo, a C++ library for scientific computing [52], after which the partition function was solved by the summation in Equation 4.3. An iteration convergence tolerance of 10^{-6} determined the minimum number of basis functions required.

Finally, the RR thermodynamics for all these methods were benchmarked against the HO approximation. The partition function for all the rotational harmonic frequencies is equal to the quantum HO reference partition function in the numerator of Equation 4.29, with each component HO partition function given by Equation 4.27 and each harmonic frequency again determined by diagonalization of the partial Hessian of only adsorbate atoms. Thermodynamic derivations under the HO, $\text{PG}_{\text{RR}}^{\text{HO}}$, classical rigid rotor (RR^{cl}) and quantum rigid rotor (RR^{qu}) approximations are given in Appendix C.

2.3. Translational Partition Functions

The partition functions for gas phase molecules are approximated classically, beginning from Equation 4.1. This results in the following well-known expression for free translation (FT):

$$Q_{\text{trans}}^{\text{cl,FT}} = \left(\frac{M}{2\pi\beta\hbar} \right)^{3/2} \mathcal{V} \quad (4.32)$$

where M is the mass of the molecule and \mathcal{V} is the volume. Similarly, we wish to evaluate partition functions for non-local modes of translation in zeolitic systems, for which local harmonic approximations significantly undercount. Just as ideal gas partition functions involve integration over the volume available to the gas, non-local translational partition functions involve integration over the accessible cell volume of the enclosing zeolite \mathcal{V}_{acc} . This leads to the following expression for non-local translation (NLT) of the adsorbate.

$$Q_{\text{trans}}^{\text{cl,NLT}} = \left(\frac{M}{2\pi\beta\hbar} \right)^{3/2} \mathcal{V}_{\text{acc}} \quad (4.33)$$

The accessible volume obviously depends not only on the zeolite framework, but also the size of the translator. Good estimates of the accessible volume for water in a number of zeolites have been attained computationally through sphere-packing methods [53], which provide a good starting point for estimating the accessible volume for more adsorbates of interest. More sophisticated sphere-packing sampling algorithms for micropore characterization have more recently been explored [54]; although not to a level of granularity for characterizing accessible volumes of significantly non-spherical molecules. Nevertheless, we proceed with the available data computed for water in H-CHA and H-MFI, and generalize the accessible volume of a generic adsorbate as

$$\mathcal{V}_{\text{acc}}(r) = \frac{r_{\text{H}_2\text{O}}^3}{r^3} \mathcal{V}_{\text{acc}}^{\text{H}_2\text{O}} \quad (4.34)$$

where r is the kinetic radius of the molecule of interest, $r_{\text{H}_2\text{O}}$ is the kinetic radius of water, and $\mathcal{V}_{\text{acc}}^{\text{H}_2\text{O}}$ is the accessible volume of water in a given zeolite. The rationale for this expression is that an (approximately spherical) adsorbate occupying x times as much space as (approximately spherical) water will have $1/x$ times the accessible volume in the unit cell. The values of $\mathcal{V}_{\text{acc}}^{\text{H}_2\text{O}}$ were obtained through the online database for zeolite structures compiled by Baerlocher and McCusker [55].

As with rotational methods developed in this work, the NLT method is benchmarked against the HO approximation, for which the partition function is the product of harmonic partition functions associated with translational normal mode frequencies ω_x , ω_y , and ω_z .

$$Q_{\text{trans}}^{\text{qu,HO}}(\omega_x, \omega_y, \omega_z) = Q^{\text{qu,HO}}(\omega_x) Q^{\text{qu,HO}}(\omega_y) Q^{\text{qu,HO}}(\omega_z) \quad (4.35)$$

These frequencies were also determined by visualization of the normal modes of the partial Hessian including only adsorbate atoms. The thermodynamics derived from these approximations are given in Appendix C.

Table 1: Generalized Procedure for Computing Translational and Rotational Thermodynamics of Adsorbed Species

	Trans. Method		Rot. Method			
	HO	NLT	HO	PG	RR ^{cl}	RR ^{qu}
Geometry optimization and single point energy calculation						
Mass-weighted Hessian diagonalization	×	·	×	×	·	·
Potential energy surface sampling	·	·	·	×	×	×
Hamiltonian matrix diagonalization	·	·	·	·	·	×
Computation of partition function and thermodynamics						

2.4. Computational Procedure and Additional Details

The synthesis of the concepts and protocols detailed above are summarized in a workflow diagram, Table 1, with necessary steps for each translational and rotational method indicated by a cross. Computational analysis for each system begins with geometry optimization of the system of interest, followed by single point electronic energy calculations. This provides the zero-Kelvin baseline to which subsequent thermal corrections are added. Depending on the combination of translational and rotational methods, the succeeding steps are vibrational analysis through diagonalization of the mass-weighted Hessian, sampling of the potential energy surface, and/or construction and diagonalization of the Hamiltonian matrix. Calculation of partition functions and the thermodynamics therefrom is the final step. Code for rotational potential energy sampling and Hamiltonian matrix diagonalization is available in public GitHub repositories [56, 57], which requires the previously-mentioned software dependencies, as well libraries containing wrapper functions facilitating Q-Chem job script creation [58].

3. Results and Discussion

3.1. Adsorption Structures and *ab Initio* Electronic Energies

Optimized structures representing the equilibrium positions and orientations of candidate adsorbates are shown in Figures 5 and 6. Electronic energies of adsorption accompanying these geometries appear in Table 2, along with previously reported values determined using different levels of theory. Alkanes exhibit relatively weak binding affinity to the acidic proton of their respective frameworks, and within the selection of alkanes, increasing chain length results in increased adsorption strength. Methane has the weakest binding strength with an adsorption energy of $-29.81 \text{ kJ mol}^{-1}$, followed by ethane at $-42.75 \text{ kJ mol}^{-1}$, and propane at $-72.27 \text{ kJ mol}^{-1}$. For these species, the interaction distance between the primary carbon of the molecule and the acidic proton is greater than 2 \AA . The predicted electronic adsorption

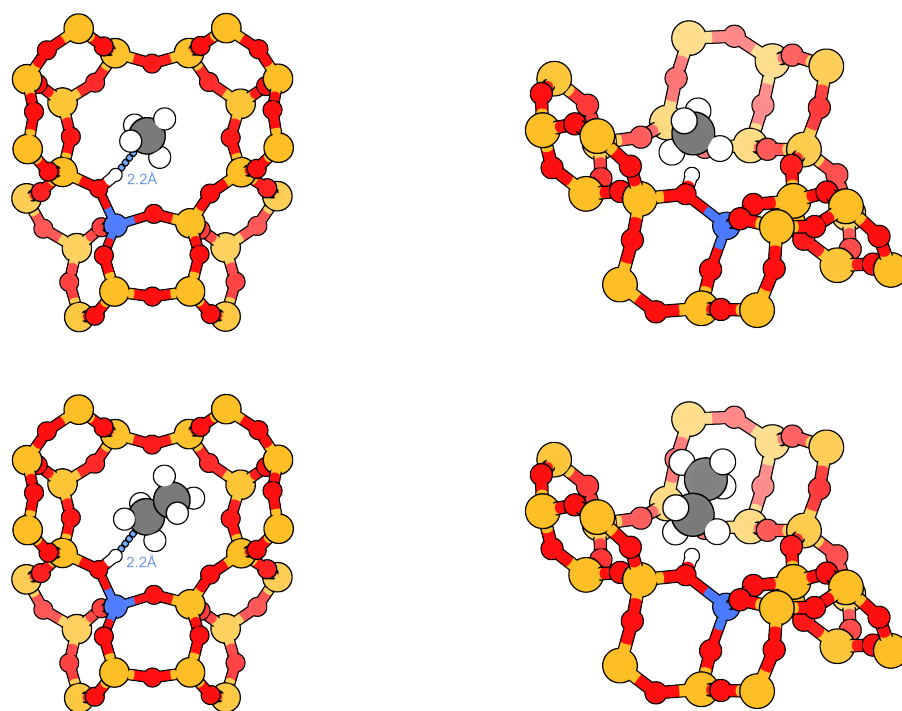


Figure 5: Adsorption geometries of methane (top row; two views) and ethane (bottom row; two views) in H-CHA. The QM region alone is pictured for clarity. Yellow atoms represent silicon; red, oxygen; blue, aluminum; black, carbon; and white, hydrogen. Bond lengths to the acidic proton are indicated in the left column.

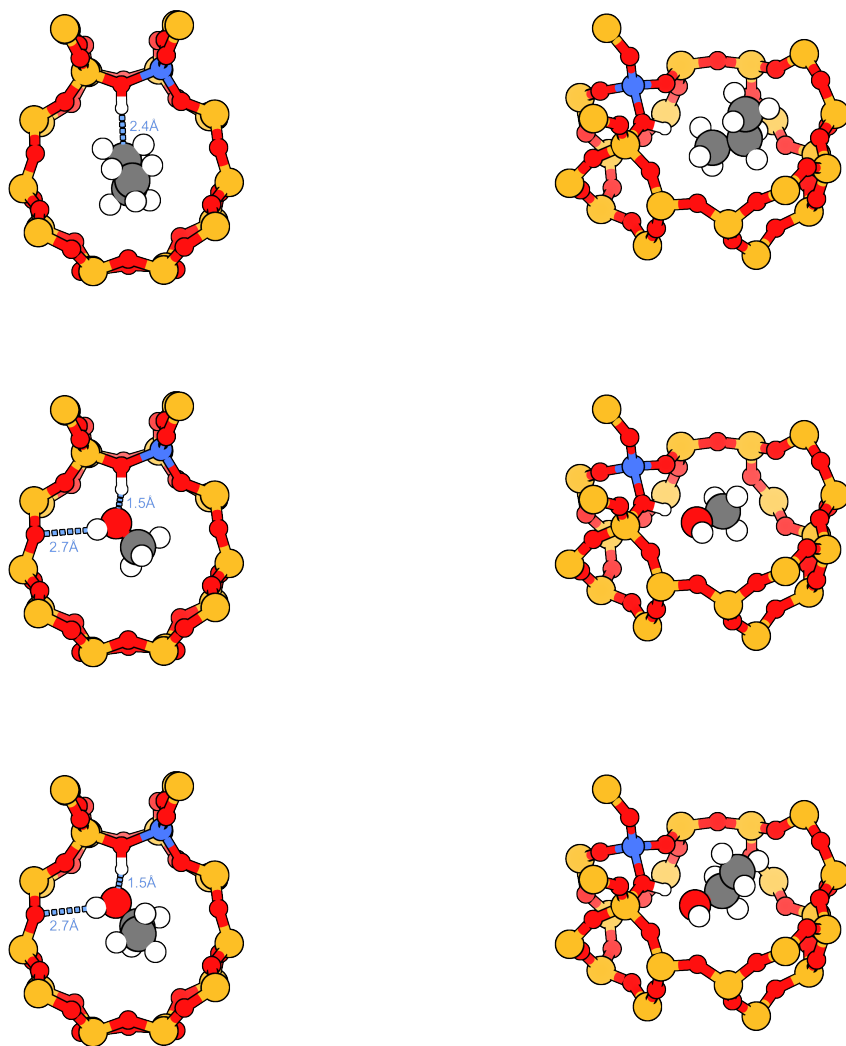


Figure 6: Adsorption geometries of propane (top row; two views), methanol (middle row; two views), and ethanol (bottom row; two views) in H-MFI. The QM region alone is pictured for clarity. Yellow atoms represent silicon; red, oxygen; blue, aluminum; black, carbon; and white, hydrogen. Bond lengths to the acidic proton are indicated in the left column.

Table 2: Calculated Electronic Adsorption Energies [kJ mol⁻¹] Species Adsorbed in Zeolites

Species	Zeolite	This work ^a	Literature refs.	
Methane	H-CHA	-29.81	-32.3 ^b	-25.3 ^c
Ethane	H-CHA	-42.75	-43.4 ^b	-36.2 ^c
Propane	H-MFI	-72.27	-68.5 ^d	-58.0 ^e
Methanol	H-MFI	-106.1	-117.7 ^f	-100.7 ^g
Ethanol	H-MFI	-117.9	-133.2 ^f	-112.8 ^g

^a ω B97M-V / def2-TZVP single point energy calculations at ω B97X-D / def2-SVP optimized geometries; ^bPBE+D at the PBE+D optimized structure, ref. [59]; ^chybrid MP2:PBE+D, ref. [19]; ^dPBE+D2 at the PBE+D2 optimized structure, ref. [60]; ^ehybrid MP2:PBE+D2 with a CCSD(T) correction, ref. [60]; ^fPBE+D2 at the PBE+D2 optimized structure, ref. [20]; ^ghybrid MP2(Counterpoise-corrected, complete basis set):PBE+D2 with a CCSD(T) correction at the PBE+D2 optimized structure, ref. [20].

energies for methane and ethane are in close agreement with those found by Göttl et al. [59], and those for propane are similar to those found by Berger et al. [60]. These DFT energies tend to be over-bound compared to electronic structure methods featuring higher level corrections, such as hybrid DFT / Møller-Plesset perturbation theory (MP2) approaches featuring coupled cluster single-double and perturbative triple (CCSD(T)) corrections [19, 60]. By contrast, electronic adsorption energies for methanol and ethanol are in closer alignment to the determinations at hybrid DFT/MP2 levels of theory [20].

3.2. Thermal Corrections to the Electronic Energy

Methane adsorbed in H-CHA was chosen as a representative system for in-depth analysis of competing methods before proceeding with the larger and significantly more complicated adsorbates. We begin with a visualization of the sampled rotational potential energy surface of methane given in Figure 7. This shows spherical representations of the potential energy at different orientations of χ , obtained after solving for the Fourier coefficients through Equation 4.22. The PES features significant barriers along the polar coordinate θ at the equilibrium position $(\phi, \theta, \chi) = \mathbf{0}$. These barriers become surmountable with rotation of the molecule by a small amount, which causes the spherical potential energy surface to flatten. As such, we anticipate a substantial retention of rotational freedom from the gas phase using the RR approximation, notwithstanding the confinement effects of the zeolite at the geometric minimum.

Indeed, we find a substantial degree of rotational entropy is retained from the gas phase to the adsorbed phase using RR and RR-like approximations, as shown in Table 3. For rotations, uncoupled classical, Pitzer-Gwinn, and variational solutions predict retention of more than 85% of the gas phase entropy, while harmonic approximations for rotations predict retention of about 57% of the rotational gas phase entropy. Expectedly, the HO approximation predicts a much more bound adsorption geometry. This is also the case for translational degrees of freedom, which account for the largest share of the total gas phase entropy. For these

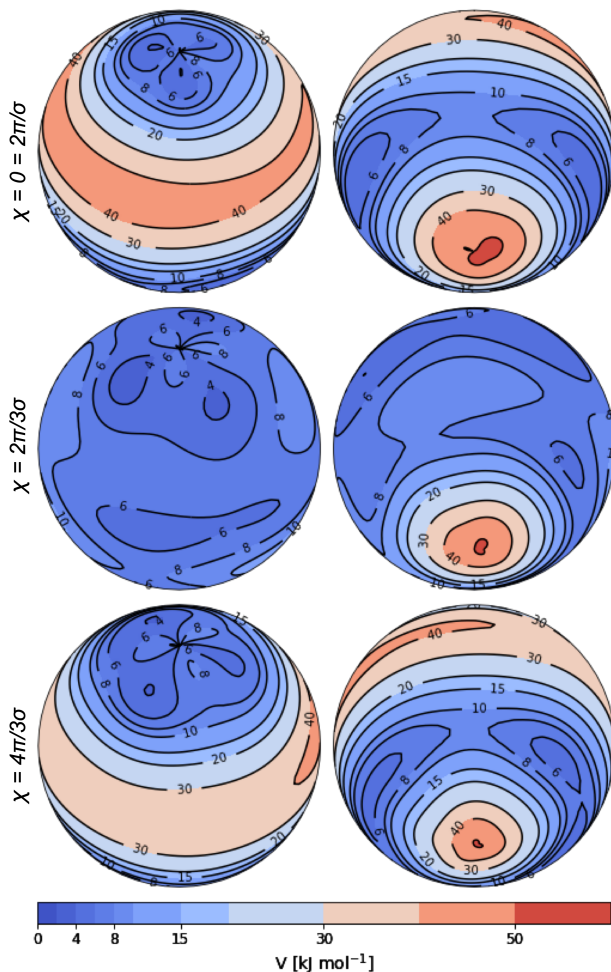


Figure 7: Spherical plots of the Wigner D-Matrix-fitted potential for methane adsorbed in H-CHA. Each row represents an intrinsic rotational orientation described by χ , with left spheres representing top-down views of the (ϕ, θ) potential energy surface, and right spheres representing the respective antipodal views.

modes, the HO approximations predict only 20.6% retention of the gas phase translational entropy, compared to 66.8% for NLT approximations. This is consistent with the analyses of Campbell and Sellers [22] and Dauenhauer and Abdelrahman [21], who found weakly-bound adsorbates were generally expected to retain two-thirds of their total gas phase entropies, substantially more than is predicted by HO approximations.

Combining rotational and translational methods, we compare the computed thermodynamics at experimental temperatures to the measured results, as shown in Table 4. We note across all combinations of methods tested, predicted enthalpies are remarkably stable. This is because the enthalpic term is dominated by the electronic energy, with variations in the thermal corrections across all methods small by comparison. By contrast, entropic terms are much more thermally sensitive, with a fully harmonic (HO/HO) entropic contribution to the free energy (i.e. $-T\Delta S$) of 38.0-41.7 kJ mol⁻¹ versus non-local translator/rigid

Table 3: Translational and Rotational Entropy Contributions: Methane in H-CHA

Mode	Method	$\Delta S_{\text{Mode}}^a [\text{J mol}^{-1} \text{K}^{-1}]$	% S Retained ^b
Translation	HO	-114.0	20.6
	NLT	-47.7	66.8
Rotation	HO	-23.5	56.7
	$\text{PG}_{\text{RR}}^{\text{HO}}$	-6.69	87.7
	RR^{cl}	-7.41	86.3
	RR^{qu}	-6.53	88.0

^aModal entropy contributions, calculated at $T = 303$ K.; ^bModal gas phase entropy retention at $T = 303$ K.

Table 4: Adsorption Thermodynamics of Methane in H-CHA for Translational/Rotational Approximation Methods

Methane	HO/HO	HO/PG	HO/RR ^{cl}	HO/RR ^{qu}	NLT/HO	NLT/PG	NLT/RR ^{cl}	NLT/RR ^{qu}	Exp.
ΔZPE	4.6	4.6	1.4	6.8	3.1	3.1	0	5.4	
ΔE_0	-25.3	-25.3	-28.4	-23.0	-26.7	-26.7	-29.8	-24.4	
$T = 273$ K									
ΔH	-24.7	-21.5	-22.0	-21.8	-25.9	-22.8	-23.2	-23.0	-20.4 ^a
$-T\Delta S$	38.0	33.2	33.5	33.2	20.6	15.8	16.1	15.8	17.7 ^a
ΔG	13.3	11.7	11.5	11.4	-5.3	-6.9	-7.2	-7.3	-2.7 ^a
$T = 303$ K									
ΔH	-24.2	-21.3	-21.7	-21.5	-25.6	-22.6	-23.0	-22.9	-17.0 ^b
$-T\Delta S$	41.7	36.6	36.8	36.5	22.8	17.7	17.9	17.6	19.2 ^b
ΔG	17.5	15.3	15.1	15.0	-2.8	-4.9	-5.2	-5.3	2.2 ^b

^aH-CHA, Si/Al = 2.6, $T = 273.2$, ref. [61]; ^bH-SSZ-13, Si/Al = 14.4, $T = 303$ K, ref. [19]

rotor (NLT/RR) entropic contributions of around 16-18 kJ mol⁻¹. To this end, NLT methods are invariably more accurate in predicting adsorption entropies, with more granular improvements achieved using $\text{PG}_{\text{RR}}^{\text{HO}}$ and RR approximations. In comparing the two experimental references, we achieve considerably better enthalpic agreement from the experiments of Barrer and Davies using H-CHA [61] compared to that of Piccini using H-SSZ-13, a more siliceous H-chabazite [19]. Discrepancies in entropies between the two experiments, however, are within the margin of chemical accuracy and are thus deemed insignificant.

An analysis of the effect of grid size on the methane adsorption showed no significant difference when decreasing grid size from 110 to 74, and 74 to 50 (see Table C.1). The savings in computational effort being significant, further computations of rotational potential energy surfaces were carried out using a Lebedev grid size of 74, for a total sampling set of 1036 and average angular resolution of 23.6 degrees. Additionally, thermodynamics calculations were not pursued through variational solutions to the Schrödinger equation for the remaining systems. This is primarily because direct diagonalization of the required Hamiltonian matrix for molecules larger than methane exceeds computer memory limitations. Because classical and Pitzer-Gwinn partition functions are significantly more tractable and comparably accurate, sparse matrix eigen-decompositions were also not pursued. The results of these determinations are presented in Table 5.

Table 5: Adsorption Thermodynamics of Molecules in Zeolites for Translational/Rotational Approximation Methods

Ethane	HO/HO	HO/PG _{RR} ^{HO}	HO/RR ^{cl}	NLT/HO	NLT/PG _{RR} ^{HO}	NLT/RR ^{cl}	Exp.
ΔZPE	3.3	3.3	1.1	2.2	2.2	0	
ΔE_0	-39.5	-39.5	-41.7	-40.6	-40.6	-42.8	
$T = 313$ K							
ΔH	-37.3	-35.8	-36.1	-38.6	-37.2	-37.4	-27.5 ^a , -30.8 ^b
$-T\Delta S$	48.9	42.1	42.2	31.1	24.3	24.4	23.82 ^a , 20.33 ^b
ΔG	11.6	6.2	6.1	-7.5	-12.9	-13.0	-3.68 ^a , -10.47 ^b
Propane	HO/HO	HO/PG _{RR} ^{HO}	HO/RR ^{cl}	NLT/HO	NLT/PG _{RR} ^{HO}	NLT/RR ^{cl}	Exp.
ΔZPE	2.5	2.5	0.6	1.9	1.9	0	
ΔE_0	-69.8	-69.8	-71.7	-70.4	-70.4	-72.3	
$T = 320$ K							
ΔH	-66.8	-63.5	-63.6	-68.1	-64.8	-65.0	-45.0 ^c
$-T\Delta S$	50.6	43.3	43.4	31.6	24.2	24.3	32.6 ^c
ΔG	-16.1	-20.2	-20.2	-36.6	-40.6	-40.7	-12.4 ^c
$T = 340$ K							
ΔH	-66.4	-62.9	-63.0	-67.9	-64.3	-64.5	-41 ^d
$-T\Delta S$	53.5	45.4	45.5	33.5	25.4	25.5	32 ^d
ΔG	-13.0	-17.5	-17.6	-34.4	-38.9	-39.0	-9 ^d
Methanol	HO/HO	HO/PG _{RR} ^{HO}	HO/RR ^{cl}	NLT/HO	NLT/PG _{RR} ^{HO}	NLT/RR ^{cl}	Exp.
ΔZPE	4.4	4.4	1.0	3.4	3.4	0	
ΔE_0	-101.7	-101.7	-105.1	-102.7	-102.7	-106.1	
$T = 323$ K							
ΔH	-100.2	-105.4	-105.9	-101.6	-106.8	-107.3	-107 ^e , -74 ^f
$-T\Delta S$	56.1	52.0	52.2	31.8	27.7	28.0	55 ^e , 47 ^f
ΔG	-44.1	-53.4	-53.7	-69.7	-79.0	-79.3	-52 ^e , -27 ^f
Ethanol	HO/HO	HO/PG _{RR} ^{HO}	HO/RR ^{cl}	NLT/HO	NLT/PG _{RR} ^{HO}	NLT/RR ^{cl}	Exp.
ΔZPE	3.9	3.9	0.8	3.1	3.1	0	
ΔE_0	-114.1	-114.1	-117.2	-114.8	-114.8	-117.9	
$T = 313$ K							
ΔH	-112.3	-120.6	-121.0	-113.6	-121.9	-122.4	-89 \pm 1 ^g
$-T\Delta S$	58.8	56.9	57.1	35.9	34.1	34.3	62 \pm 1 ^g
ΔG	-53.5	-63.7	-63.9	-77.7	-87.8	-88.1	-27 \pm 0.1 ^g
$T = 400$ K							
ΔH	-110.9	-119.1	-119.4	-112.6	-120.8	-121.1	-130 \pm 5 ^h
$-T\Delta S$	73.5	71.0	71.1	45.6	43.1	43.2	
ΔG	-37.4	-48.1	-48.3	-67.0	-77.7	-77.9	

^aZeolite H-SSZ-13, Si/Al = 14.4, ref. [19]; ^bZeolite H-CHA, Si/Al = 2.6, ref. [61]; ^crefs. [62, 63]; ^dObtained 300-400 K, ref. [64]; ^eData extracted from Fig. 1 of ref. [65], coverage of 0.4; ^fData extracted from Fig. 1 of ref. [65], coverage of 0.75; ^gmonomeric ethanol adsorption, H-MFI, Si/Al = 15, [66]; ^hcoverage of 1.0, ref. [67]. Statistical errors are given where reported.

We first comment on the enthalpies, which indicate the predicted binding strength. For alkanes ethane and propane, we find that our methods predict more negative adsorption enthalpies relative to experiments. As with methane, these quantities are dominated by the predicted energies at zero Kelvin, and consequently represent limitations in electronic structure methods. For methanol and ethanol, the deviation from experimental enthalpies is largely dependent on the experimental determinations themselves, with different adsorbate coverages and experimental methods producing widely different results. We find generally better agreement in the adsorption enthalpy of ethanol, for example, comparing against the calorimetry experiments of Lee [67] relative to those of Alexopoulos [66]. Similarly, we find our models’ predicted adsorption enthalpies of methanol to be much closer to the experimentally-determined enthalpy at a coverage of 0.4 relative to that of 0.75.

The entropies paint a clearer picture. As with methane, alkanes ethane and propane are unmistakably best predicted using the NLT approximation. RR and $\text{PG}_{\text{RR}}^{\text{HO}}$ rotational methods improve the NLT for ethane, for a predicted entropy contribution of about 24 kJ mol^{-1} compared to $20.3\text{-}23.8 \text{ kJ mol}^{-1}$ predicted by experiments. Propane is a slightly different story, where NLT/RR and NLT/ $\text{PG}_{\text{RR}}^{\text{HO}}$ methods stray further than NLT/HO from experimental values. The degree to which this result is due to an over-approximation of rotational entropy retention predicted by $\text{PG}_{\text{RR}}^{\text{HO}}$ and RR methods versus an over-approximation of translational entropy retention predicted by NLT is unclear. When it comes to methanol and ethanol, adsorption entropies are primarily determined by the translational contribution. The highly local HO approximation expectedly predicts higher entropy loss from the gas phase to the adsorbed phase, which for H-bonded methanol and ethanol species proves to be a good approximation. The NLT fails in this limit, systematically underestimating the entropic contribution to the free energy. Within the translational approximations, however, differing rotational methods produce remarkably similar results, indicating RR and $\text{PG}_{\text{RR}}^{\text{HO}}$ thermodynamics obtained from sampling also predict a highly localized potential energy surface, notwithstanding the non-locality of the sampling domain.

Free energies are reported in Table 5. Beginning with ethane, we find that we can attain remarkable agreement against the experimental determinations of Barrer and Davies [61] using NLT/RR and NLT/ $\text{PG}_{\text{RR}}^{\text{HO}}$ approximations, a consequence of a modest underestimation of enthalpic contributions and overestimation of entropic contributions. If we were to use higher level of theory electronic energy calculations pursued by Piccini et al. [19], we would underestimate enthalpic contributions by 6.55 kJ mol^{-1} less (see Table 2), predicting more reliably a free energy that lies between both experimental values. This is similarly the case for propane, whose binding enthalpy using our QM/MM scheme is largely underestimated. Using the higher level of theory electronic adsorption energy, we underestimate enthalpies of propane adsorption by $14.27 \text{ kJ mol}^{-1}$ less, placing predicted enthalpies within the reach of chemical accuracy. Nevertheless, apparently errors in the enthalpies of adsorption for propane lead to artificial support of HO translational methods when benchmarking against experimental free energies, despite having considerably less accurate entropies relative to NLT counterparts. Such a cancellation of errors can be a constant theme in using primitive approximations to support experimental observations sensitive to the free energy, such as the rate coefficients or equilibrium constants. A general, system-agnostic method should ideally minimize the errors of its component parts. Therein lies the importance of considering both contributions to the free energy separately.

In the case of methanol adsorption, we find quite good agreement using the HO/ $\text{PG}_{\text{RR}}^{\text{HO}}$ and HO/ $\text{PG}_{\text{RR}}^{\text{HO}}$ approximations with experimental free energies of adsorption at a site coverage of 0.4 [65]. This is attributed to (1) RR and $\text{PG}_{\text{RR}}^{\text{HO}}$ methods predicting lower binding enthalpies relative to HO rotational counterparts, likely due to their consideration of multiple H-bonding conformers, (2) the comparable accuracy of our QM/MM methods in methanol and ethanol systems with higher level of theory determinations, and (3) the accuracy of HO translation entropy approximations for strongly bound systems. However, as discussed above, experimental data for methanol and ethanol in H-MFI are highly uncertain. Therefore, the same agreement in free energies cannot be found for higher methanol coverages, nor for ethanol adsorption altogether. This finding is similar to that of Piccini et al. [20], who postulated variability of experimental results is due to surface heterogeneity, differences in Si/Al ratios, or lattice imperfections, all of which are not explicitly accounted for in our idealized framework model of H-MFI.

4. Conclusion

The methods developed in this work were undertaken for the purpose of improving the anharmonic description of rotational and translation motions of molecules adsorbed in zeolites, particularly with regard to the entropy. To this end, NLT approximations and more rigorous RR and $\text{PG}_{\text{RR}}^{\text{HO}}$ approximations are marked improvements in describing the adsorption of alkanes in zeolites, although the same principles can be applied in other weakly-bound systems in heterogeneous catalysis.

More strongly bound adsorbates are also found to potentially benefit from direct consideration of rotational degrees of freedom, with RR and $\text{PG}_{\text{RR}}^{\text{HO}}$ approximations capable of accounting for multiple stable adsorption conformers. Moreover, RR and $\text{PG}_{\text{RR}}^{\text{HO}}$ methods are exceptionally scalable, with the number of samples needed independent of molecule size. This opens the possibility of applying higher levels of electronic structure theory in describing the potential energy surface of rotations. These results lead us to advocate for the general applicability of the rigid rotor methods developed herein.

But while rotational descriptions of the adsorbed molecule are rigorous in the uncoupled limit, considerable improvements can be made for translational contributions, which as of yet lack the same generality. Our crude NLT approximation assumes free translation of a sphere in the vacuum space of the zeolite, whereas realistically finer features of the unit cell environment, and the molecule, can significantly hinder translational freedom. This is particularly the case for the methanol, ethanol, and propane molecules investigated in this work. A more general technique might improve on the description of the occupiable volume, including more sophisticated determinations of occupiable volume [54] and improved non-local translational sampling techniques in three dimensions.

5. Future Directions

The result of this study brings theory closer to the objective of chemical accuracy in modeling adsorption in zeolites, with a rigorous method for obtaining the potential energy surface for

whole-adsorbate rotations parameterized by three non-separable degrees of freedom. While the harmonic and NLT approximations can offer reasonable results in the extremes of strong and weak adsorption, respectively, a unified protocol for obtaining translational potential energy surface and the combined rotational-translational thermodynamics remains to be formulated. We propose the following work can be done to this end:

In this work, translations are assumed to be uncoupled from rotations, such that the translational potential energy surface can be found distinctly from the rotational potential energy surface. An uncoupled mode treatment for translations would include jointly sampling finite displacements of the adsorbate center of mass from its equilibrium position in x , y , and z spatial coordinates. The resulting surface can be modeled using three-dimensional Fourier basis functions, the resulting thermodynamics of which can be solved using the variational method, classical integration, or harmonic-reference $\text{PG}_{\text{RR}}^{\text{HO}}$ methods.

A non-separable approach can also be taken, which would include joint translational and rotational sampling of the adsorbate. This process would thus require sampling x , y , z , ϕ , θ , and χ together. This six-dimensional problem approaches the limit of computational feasibility, particularly when it comes to obtaining a variational solution to the Schrodinger equation obtained by matrix diagonalization. Instead, the classical translational-rotational partition function can be solved by Monte Carlo integration. A quantum corrective factor obtained using uncoupled mode reference potentials for translation and rotation can be applied to the mode coupled result for an improved description of coupled adsorbate motion, analogous to the investigation of torsional mode coupling undertaken in Chapter III.

6. Acknowledgements

Special thanks to Brian Lee Bettinson for his insight on the mathematics of rigid body rotations. We also thank Justin Talbot and Romit Chakraborty for helpful discussions. This work was supported by the Director, Office of Science, Office of Basic Energy Sciences of the US Department of Energy under contract No. DE-AC02-05CH11231.

References

- [1] Kunhao Li, Julia Valla, and Javier Garcia-Martinez. “Realizing the Commercial Potential of Hierarchical Zeolites: New Opportunities in Catalytic Cracking”. *ChemCatChem* **6.1** (2014), 46–66
- [2] Vincent Blay et al. “Engineering Zeolites for Catalytic Cracking to Light Olefins”. *ACS Catal.* **7.10** (2017), 6542–6566
- [3] Ahmed I Osman et al. “Recent Advances in Carbon Capture Storage and Utilisation Technologies: a Review”. *Environ. Chem. Lett.* **19.2** (2021), 797–849
- [4] Cong Chao et al. “Post-Combustion Carbon Capture”. *Renewable and Sustainable Energy Reviews* **138** (2021), 110490. ISSN: 1364-0321. DOI: <https://doi.org/10.1016/j.rser.2020.110490>. URL: <https://www.sciencedirect.com/science/article/pii/S1364032120307760>

- [5] Helge Bux et al. “Zeolitic Imidazolate Framework Membrane with Molecular Sieving Properties by Microwave-Assisted Solvothermal Synthesis”. *J. Am. Chem. Soc.* **131**.44 (2009), 16000–16001
- [6] Chen Zhang et al. “Unexpected Molecular Sieving Properties of Zeolitic Imidazolate Framework-8”. *J. Phys. Chem. Lett.* **3**.16 (2012), 2130–2134
- [7] Sabéha Kesraoui-Ouki, Christopher R. Cheeseman, and Roger Perry. “Natural Zeolite Utilisation in Pollution Control: A Review of Applications to Metals’ Effluents”. *J. Chem. Technol. Biotechnol.* **59**.2 (1994), 121–126. DOI: <https://doi.org/10.1002/jctb.280590202>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/jctb.280590202>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jctb.280590202>
- [8] Rolando Roque-Malherbe. “Applications of Natural Zeolites in Pollution Abatement and Industry”. Dec. 2001, pp. 495–522. DOI: [10.1016/B978-012513910-6/50069-4](https://doi.org/10.1016/B978-012513910-6/50069-4)
- [9] Yi-Pei Li et al. “Improved Force-field Parameters for QM/MM Simulations of the Energies of Adsorption for Molecules in Zeolites and a Free Rotor Correction to the Rigid Rotor Harmonic Oscillator Model for Adsorption Enthalpies”. *J. Phys. Chem. C* **119**.4 (2015), 1840–1850
- [10] Amber Janda et al. “Effects of Zeolite Structural Confinement on Adsorption Thermodynamics and Reaction Kinetics for Monomolecular Cracking and Dehydrogenation of n-Butane”. *J. Am. Chem. Soc.* **138**.14 (2016), 4739–4756
- [11] Branko Ruscic. “Uncertainty Quantification in Thermochemistry, Benchmarking Electronic Structure Computations, and Active Thermochemical Tables”. *Int. J. Quantum Chem.* **114**.17 (Sept. 2014), 1097–1101. ISSN: 0020-7608. DOI: [10.1002/qua.24605](https://doi.org/10.1002/qua.24605). URL: <https://doi.org/10.1002/qua.24605>
- [12] Sankaranarayananpillai Shylesh et al. “Experimental and Computational Studies of Carbon–Carbon Bond Formation via Ketonization and Aldol Condensation over Site-isolated Zirconium Catalysts”. *ACS Catal.* **10**.8 (2020), 4566–4579
- [13] Christopher R Ho et al. “Zeolite-Catalyzed Isobutene Amination: Mechanism and Kinetics”. *ACS Catal.* **9**.8 (2019), 7012–7022
- [14] David G Hanna et al. “Experimental and Theoretical Study of n-Butanal Self-condensation over Ti Species Supported on Silica”. *ACS catalysis* **4**.9 (2014), 2908–2916
- [15] Florian Eder and Johannes A Lercher. “Alkane Sorption in Molecular Sieves: The Contribution of Ordering, Intermolecular Interactions, and Sorption on Brønsted Acid Sites”. *Zeolites* **18**.1 (1997), 75–81
- [16] Aditya Bhan et al. “Entropy Considerations in Monomolecular Cracking of Alkanes on Acidic Zeolites”. *J. Catal.* **253**.1 (2008), 221–224
- [17] Rajamani Gounder and Enrique Iglesia. “The Roles of Entropy and Enthalpy in Stabilizing Ion-Pairs at Transition States in Zeolite Acid Catalysis”. *Acc. Chem. Res.* **45**.2 (2012), 229–238
- [18] Stefan Grimme. “Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory”. *Chem. Eur. J.* **18**.32 (2012), 9955–9964

- [19] GiovanniMaria Piccini et al. “Accurate Adsorption Thermodynamics of Small Alkanes in Zeolites. Ab Initio Theory and Experiment for H-Chabazite”. *J. Phys. Chem. C* **119.11** (2015), 6128–6137
- [20] GiovanniMaria Piccini, Maristella Alessio, and Joachim Sauer. “Ab Initio Study of Methanol and Ethanol Adsorption on Brønsted Sites in Zeolite H-MFI”. *Phys. Chem. Chem. Phys.* **20.30** (2018), 19964–19970
- [21] Paul J Dauenhauer and Omar A Abdelrahman. “A Universal Descriptor for the Entropy of Adsorbed Molecules in Confined Spaces”. *ACS Central Science* **4.9** (2018), 1235–1243
- [22] Charles T Campbell and Jason RV Sellers. “The Entropies of Adsorbed Molecules”. *J. Am. Chem. Soc.* **134.43** (2012), 18109–18115
- [23] Yi-Pei Li, Alexis T Bell, and Martin Head-Gordon. “Thermodynamics of Anharmonic Systems: Uncoupled Mode Approximations for Molecules”. *J. Chem. Theory Comput.* **12.6** (2016), 2861–2870
- [24] Berend Smit and Theo LM Maesen. “Molecular Simulations of Zeolites: Adsorption, Diffusion, and Shape Selectivity”. *Chemical Reviews* **108.10** (2008), 4125–4184
- [25] Peter J Kostelec and Daniel N Rockmore. “FFTs on the Rotation Group”. *J. Fourier Anal. Appl.* **14.2** (2008), 145–179
- [26] Jeng-Da Chai and Martin Head-Gordon. “Long-range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections”. *Phys. Chem. Chem. Phys.* **10.44** (2008), 6615–6620
- [27] Florian Weigend and Reinhart Ahlrichs. “Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy”. *Phys. Chem. Chem. Phys.* **7.18** (2005), 3297–3305
- [28] Kenno Vanommeslaeghe et al. “CHARMM General Force Field: A Force Field for Drug-like Molecules Compatible with the CHARMM All-atom Additive Biological Force Fields”. *J. Comput. Chem.* **31.4** (2010), 671–690
- [29] Narbe Mardirossian and Martin Head-Gordon. “ ω B97M-V: A Combinatorially Optimized, Range-separated Hybrid, Meta-GGA Density Functional with VV10 Nonlocal Correlation”. *J. Chem. Phys.* **144.21** (2016), 214110
- [30] Evgeny Epifanovsky et al. “Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package”. *J. Chem. Phys.* **155.8** (2021), 084801
- [31] Hashim A Hashim. “Special Orthogonal Group SO (3), Euler Angles, Angle-Axis, Rodriguez Vector and Unit-Quaternion: Overview, Mapping and Challenges”. *arXiv preprint arXiv:1909.06669* (2019)
- [32] Richard Zare. *Angular Momentum : Understanding Spatial Aspects in Chemistry and Physics*. New York: Wiley, 1988. ISBN: 978-0-471-85892-8
- [33] H. W. Kroto. *Molecular Rotation Spectra*. London, New York: Wiley, 1975. ISBN: 9780486495408
- [34] D. M. Brink. *Angular Momentum*. Oxford New York: Clarendon Press Oxford University Press, 1993. ISBN: 978-0198517597

- [35] Luis Simón-Carballido et al. “Anharmonicity of Coupled Torsions: the Extended Two-dimensional Torsion Method and its Use to Assess More Approximate Methods”. *J. Chem. Theory Comput.* **13.8** (2017), 3478–3492
- [36] Paul N Swarztrauber and William F Spitz. “Generalized Discrete Spherical Harmonic Transforms”. *J. Comput. Phys.* **159.2** (2000), 213–230
- [37] Paul N Swarztrauber. “On the Spectral Approximation of Discrete Scalar and Vector Functions on the Sphere”. *SIAM J. Numer. Anal.* **16.6** (1979), 934–949
- [38] Krzysztof M Gorski et al. “HEALPix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere”. *Astrophys. J.* **622.2** (2005), 759
- [39] Randy C Hoover, Anthony A Maciejewski, and Rodney G Roberts. “Pose detection of 3-D objects using images sampled on SO (3), spherical harmonics, and Wigner-D matrices”. *2008 IEEE International Conference on Automation Science and Engineering*. IEEE. 2008, pp. 47–52
- [40] James R Driscoll and Dennis M Healy. “Computing Fourier Transforms and Convolutions on the 2-Sphere”. *Advances in applied mathematics* **15.2** (1994), 202–250
- [41] Dennis M Healy et al. “FFTs for the 2-Sphere-Improvements and Variations”. *J. Fourier Anal. Appl.* **9.4** (2003), 341–385
- [42] Gavin A McCarver and Robert J Hinde. “High Accuracy ab Initio Potential Energy Surface for the H₂O–H van der Waals Dimer”. *J. Chem. Phys.* **155.11** (2021), 114302
- [43] Thomas R Furlani, Jing Kong, and Peter MW Gill. “Parallelization of SCF Calculations within Q-Chem”. *Computer physics communications* **128.1-2** (2000), 170–177
- [44] Christopher W Murray, Nicholas C Handy, and Gregory J Laming. “Quadrature Schemes for Integrals of Density Functional Theory”. *Molecular Physics* **78.4** (1993), 997–1014
- [45] Radovan Bast. *Numgrid: Numerical integration grid for molecules*. Version v2.1.0. Jan. 2021. DOI: [10.5281/zenodo.1470276](https://doi.org/10.5281/zenodo.1470276). URL: <https://doi.org/10.5281/zenodo.1470276>
- [46] Mike Boyle. *moble/spherical: Release v1.0.10*. 2021. DOI: [10.5281/ZENODO.5590582](https://doi.org/10.5281/ZENODO.5590582). URL: <https://zenodo.org/record/5590582>
- [47] William W Kennerly. *Molecules Rotating in Electric Fields by Quantum and Semi-Quantum Mechanics*. Cornell University, 2005
- [48] Mengjie Liu et al. “Reaction Mechanism Generator v3.0: Advances in Automatic Mechanism Generation”. *J. Chem. Inf. Model.* **61.6** (2021). PMID: 34048230, 2686–2696. DOI: [10.1021/acs.jcim.0c01480](https://doi.org/10.1021/acs.jcim.0c01480). eprint: <https://doi.org/10.1021/acs.jcim.0c01480>. URL: <https://doi.org/10.1021/acs.jcim.0c01480>
- [49] Kenneth S Pitzer and William D Gwinn. “Thermodynamic Functions for Molecules with Internal Rotation”. *J. Chem. Phys.* **9.6** (June 1941), 485–486. ISSN: 0021-9606. DOI: [10.1063/1.1750941](https://doi.org/10.1063/1.1750941). URL: <https://doi.org/10.1063/1.1750941>
- [50] Rohit Chandra et al. *Parallel Programming in OpenMP*. Morgan Kaufmann, 2001
- [51] Joey Dumont. *Wignersymbols: Release 0.1.3*. 2014. DOI: [10.5281/ZENODO.12580](https://doi.org/10.5281/ZENODO.12580). URL: <https://zenodo.org/record/12580>

- [52] Conrad Sanderson and Ryan Curtin. “Armadillo: a Template-Based C++ Library for Linear Algebra”. *J. Open Source Softw.* **1.2** (2016), 26
- [53] M.D. Foster et al. “A Geometric Solution to the Largest-Free-Sphere Problem in Zeolite Frameworks”. *Microporous Mesoporous Mater.* **90.1** (2006). Dedicated to the late Denise Barthomeuf, George Kokotailo and Sergey P. Zhdanov in appreciation of their outstanding contributions to zeolite science, 32–38. ISSN: 1387-1811. DOI: <https://doi.org/10.1016/j.micromeso.2005.08.025>. URL: <https://www.sciencedirect.com/science/article/pii/S1387181105003914>
- [54] Daniele Ongari et al. “Accurate Characterization of the Pore Volume in Microporous Crystalline Materials”. *Langmuir* **33.51** (2017), 14529–14538
- [55] Ch. Baerlocher and L.B. McCusker. *Database of Zeolite Structures*. URL: <http://www.iza-structure.org/databases/>
- [56] Lance A. Bettinson. *Rotational Potential Energy Surface Sampler*. URL: https://github.com/lbettins/spherical_schrodinger
- [57] Lance A. Bettinson. *Hamiltonian Matrix Diagonalization*. URL: <https://github.com/lbettins/rotational-hamiltonian>
- [58] Lance A. Bettinson. *No U-Turn Sampler for Intramolecular Torsions*. URL: <https://github.com/lbettins/T-NUTS>
- [59] Florian Göttl et al. “Van der Waals interactions between hydrocarbon molecules and zeolites: Periodic calculations at different levels of theory, from density functional theory to the random phase approximation and Møller-Plesset perturbation theory”. *J. Chem. Phys.* **137.11** (2012), 114111
- [60] Fabian Berger, Marcin Rybicki, and Joachim Sauer. “Adsorption and Cracking of Propane by Zeolites of Different Pore Size”. *J. Catal.* **395** (2021), 117–128
- [61] Richard Maling Barrer and JA Davies. “Sorption in Decationated Zeolites II. Simple Paraffins in H-Forms of Chabazite and Zeolite L”. *Proc. R. Soc. Lond. A Math. Phys. Sci.* **322.1548** (1971), 1–19
- [62] Florian Eder and Johannes A Lercher. “Alkane Sorption in Molecular Sieves: the Contribution of Ordering, Intermolecular Interactions, and Sorption on Brønsted Acid Sites”. *Zeolites* **18.1** (1997), 75–81
- [63] F Eder, M Stockenhuber, and JA Lercher. “Brønsted Acid Site and Pore Controlled Siting of Alkane Sorption in Acidic Molecular Sieves”. *J. Phys. Chem. B* **101.27** (1997), 5414–5419
- [64] Bart A De Moor et al. “Adsorption of C2- C8 n-Alkanes in Zeolites”. *J. Phys. Chem. C* **115.4** (2011), 1204–1219
- [65] Christopher G Pope. “Adsorption of Methanol and Related Molecules on Zeolite H-ZSM-5 and Silicalite”. *J. Chem. Soc. Faraday Trans.* **89.7** (1993), 1139–1141
- [66] Konstantinos Alexopoulos et al. “Anharmonicity and Confinement in Zeolites: Structure, Spectroscopy, and Adsorption Free Energy of Ethanol in H-ZSM-5”. *J. Phys. Chem. C* **120.13** (2016), 7172–7182

- [67] C-C Lee, RJ Gorte, and WE Farneth. “Calorimetric Study of Alcohol and Nitrile Adsorption Complexes in H-ZSM-5”. *J. Phys. Chem. B* **101**.19 (1997), 3811–3817

Appendices

Appendix A

Supplementary Information for Chapter II

1. Experimental Methods

1.1. Catalyst Synthesis

Zirconium was grafted onto a silica support using the following procedure. One gram of amorphous silica (Silicycle, surface area: $500 \text{ m}^2 \text{ g}^{-1}$) was dried in vacuum at 373 K for 24 h, or pretreated at 1023 K for 5 h, and then stored in vacuum prior to use. Zr was introduced onto the silica surface either by grafting Cp_2ZrCl_2 (Aldrich, 97%) dissolved in toluene (Alfa Aesar, anhydrous 99.8% pure) or by impregnation with a toluene solution of $\text{Zr}(\text{O}^i\text{Pr})_4$ (Aldrich, 99.999%). These catalysts are referred to as $x\text{ZrCp}/\text{SiO}_2$ and $x\text{ZrPr}/\text{SiO}_2$, respectively, where x denotes the dispersion of zirconium (atoms nm^{-2}). A similar synthesis procedure was utilized for the synthesis of titanium-silica and tin-silica catalysts using Cp_2TiCl_2 and $(\text{CH}_3)_2\text{SnCl}_2$ as the respective metal precursors. All impregnated and grafted catalysts were dried at 393 K for 12 h and calcined in $100 \text{ cm}^3 \text{ min}^{-1}$ of air (Praxair) for 823 K for 6 h. A zirconium dimer complex, $[(i\text{-PrCp})_2\text{ZrH}(\mu\text{-H})]_2$ was also prepared according to literature procedures [1].

1.2. Catalyst Activity

Reaction rates were measured using a 6.35 mm OD quartz reactor containing an expanded section ($\approx 12.7 \text{ mm OD}$, $\approx 20 \text{ mm length}$). A plug of quartz wool placed below the catalyst bed to hold the powder in place. The catalyst bed temperature was measured with a K-type thermocouple sheathed in a quartz capillary placed in direct contact with the catalyst bed. Prior to reaction, the catalyst was heated to the reaction temperature at a rate of 2 K min^{-1} in pure He (Praxair, 99.999%) flowing at $100 \text{ cm}^3 \text{ min}^{-1}$ at STP. A filled 1 mL syringe connected to a syringe pump (Cole-Palmer, 74900 series) was used to inject propanoic acid or propan-2-one into a heated port through which helium was flowed continuously. All experiments were carried out at a total gas pressure of 1 atm. The total gas flow rate was typically $100 \text{ cm}^3 \text{ min}^{-1}$ at STP. Under these conditions, the conversion of carboxylic acids or propan-2-one was less than 20%. Reaction products were analyzed using an Agilent 6890N

gas chromatograph containing a bonded and cross-linked (5%-phenyl)-methylpolysiloxane capillary column (Agilent, HP-1) connected to a flame ionization detector (FID). The rate of product formation per Zr atom, or turnover frequency (TOF) was calculated by dividing the measured rate of reaction per gram of catalyst by the measured amount of Zr per gram of catalyst.

1.3. Catalyst Characterization

The metal content of the catalysts was determined by inductively coupled plasma optical emission spectroscopy (ICP-OES) at Galbraith Laboratories (Knoxville, TN). Nitrogen adsorption isotherms were performed using a Micromeritics Gemini VII surface area and pore volume analyzer. The specific surface area and pore size were calculated using the Brunauer-Emmet-Teller (BET) equation and Barrett-Joyner-Halenda (BJH) equations. Raman spectra were acquired at ambient conditions using a confocal Raman microscope (LabRam HR, Horiba Jobin Yvon) equipped with a 532 nm HeNe laser operated at a power of 50 mW. Diffuse reflectance UV-Vis spectra were acquired using a Fischer Scientific EVO 300 spectrometer equipped with a Praying Mantis reflectance chamber. Spectra were referenced to the diffuse reflectance spectrum of Teflon.

2. Experimental Results

2.1. Catalyst Characterization

The properties of each catalyst are shown in Table A.1. The surface area of the Cp_2ZrCl_2 grafted silica sample is very similar to that of the silica support at low Zr loadings and decreases progressively with increasing weight loading of Zr on the silica support. We note that a part of this loss is apparent and is associated with the increasing mass of the deposited Zr relative to that of the silica support.

Literature XANES and EXAFS analysis of Zr dispersed on silica derived from Cp_2ZrCl_2 and $\text{Zr}(\text{O}^i\text{Pr})_4$ showed that Zr forms in coordinatively unsaturated, tetragonal structures [1]. We further probed the coordination and connectivity of Zr by Raman and UV-Vis spectroscopy. Raman spectra of the silica support and various $x\text{ZrCp}/\text{SiO}_2$ samples are shown in Figure A.6. The spectrum of the support exhibits peaks at 485 cm^{-1} and 975 cm^{-1} attributable to the stretching vibrations of four-membered siloxane linkages ($\text{Si}-\text{O}-\text{Si}$) and surface silanol groups ($\text{O}_3\text{Si}-\text{OH}$), respectively [2]. For loading of up to 0.1 Zr nm^{-2} , the spectrum of ZrCp/SiO_2 is indistinguishable from that of silica because of the low weight loading of Zr. Samples with higher Zr loadings exhibit the characteristic peaks of tetragonal zirconia at 600 cm^{-1} and 800 cm^{-1} due to Ag vibrations of $\text{Zr}-\text{O}-\text{Zr}$ bonds [3]. Although XRD patterns do not show the presence of bulk oxides at any Zr loading levels, the Raman spectra suggest that a fraction of the Zr is present as ZrO_2 nanoparticles, possibly in particles smaller than 5 nm in diameter for loadings of more than $0.5\text{ Zr atoms nm}^{-2}$.

The UV-Vis edge energy of $x\text{ZrCp}/\text{SiO}_2$ prepared with different Zr surface densities is shown in Figure A.1 together with those of model compounds. The observed decrease in the edge energy of $x\text{ZrCp}/\text{SiO}_2$ with increasing Zr surface density is attributable to an increase

Table A.1: Chemical and textural properties of Zr containing silica samples

Sample	Zr content (wt.%)	BET surface area [m ² g ⁻¹]	TOF [10 ⁻⁴ , s ⁻¹]
0.07ZrCp/SiO ₂	0.61	495	3.2
0.1ZrCp/SiO ₂	0.82	490	3.3
0.1ZrCp/SiO ₂ a	0.82	482	7.3
0.5ZrCp/SiO ₂	3.85	435	2.5
0.7ZrCp/SiO ₂	4.30	413	1.5
1.0ZrCp/SiO ₂	6.10	385	0.5
0.1ZrPr/SiO ₂	0.86	473	2.1
SiO ₂	-	500	0

Sample prepared on silica calcined at 1023 K prior to grafting with Cp₂ZrCl₂. Reaction conditions: $T = 573$ K, $P_{\text{Tot.}} = 1$ atm, $Q_{\text{Tot.}} = 100$ cm³ min⁻¹, $P_{\text{acid}} = 0.2$ kPa, $M_{\text{cat.}} = 0.1$ g.

in the connectivity of Zr atoms. These results suggest that, at low Zr loading (0.1 Zr atoms nm⁻¹), Zr exists as isolated tetrahedral structures, at intermediate Zr loadings (0.5 Zr atoms nm⁻¹), possibly as dimeric or oligomeric species, and at high loadings (1.0 Zr atoms nm⁻¹), as bulk-like ZrO₂ structures. The trend in Zr speciation with increasing Zr surface density is supported by measured edge energies for the model compounds Cp₂ZrCl₂ (isolated Zr atoms) and [(*i*-PrCp)₂ZrH(μ -H)]₂ (isolated Zr dimers) and for bulk ZrO₂.

2.2. Catalytic Activity of Silica-Supported Zr Catalysts

The catalytic activity of silica-supported zirconia was evaluated for the ketonic decarboxylation of propanoic acid at 573 K. Experiments were conducted under conditions of differential conversion (< 2%) so that catalyst deactivation and secondary condensation could be minimized. The only product observed under the conditions investigated was pentan-3-one. The rate of formation of the product was zero order in the partial pressure of propanoic acid for partial pressures between 0.1 kPa and 0.4 kPa (Figure A.7). Control experiments with silica displayed no catalytic activity, indicating that surface-grafted zirconium sites are responsible for the observed catalytic activity.

The dependence of the specific activity of ZrCp/SiO₂ on the Zr loading is shown in Figure A.2. As the surface density of Zr increases, the turnover frequency decreases monotonically. The catalyst prepared with a low surface concentration of zirconium (0.1 Zr nm⁻²) is approximately six times more active than that prepared with 1.0 Zr nm⁻², whereas the activity of the catalyst prepared with a high zirconium, such as 2.0 Zr nm⁻² is negligible. The high activity of 0.1ZrCp/SiO₂ suggests that isolated tetrahedrally coordinated Zr sites are more active than Zr atoms in clusters.

FT-IR spectra of adsorbed CH₃CN were acquired as a function of temperature to probe the strength of Lewis and Brønsted acidic sites. As shown in Figure A.8, the peaks observed at 2315 cm⁻¹ and 2290 cm⁻¹ correspond to the $\nu(\text{CN})$ mode of CH₃CN, N-coordinated to Zr

split by the coupling with the $\nu(\text{CC}) + \delta_{\text{sym}}(\text{CH}_3)$ combination, while the peak at 2260 cm^{-1} corresponds to physisorbed or H-bonded CH_3CN associated with surface hydroxyl groups [4]. A strong absorption band is not observed in the region of $2200\text{-}2050 \text{ cm}^{-1}$, characteristic of $\nu(\text{CN})$ modes of CH_2CN^- species and polymeric anions such as $[\text{CH}_3\text{C}(\text{NH})\text{CHCN}]^-$, suggesting the absence of strongly Lewis basic sites on ZrCp/SiO_2 [5].

The data presented in Figure A.8 show that CH_3CN desorbs at a lower temperature from Zr oligomers (1.0 Zr nm^{-2}) than from isolated $\equiv\text{Zr-OH}$ species (0.1 Zr nm^{-2}). These results suggest that more highly coordinated Zr sites are less Lewis acidic than isolated four-coordinated Zr sites, from which we infer that the adsorption of propanoic acid via binding of the carbonyl group with the Lewis acid center will be stronger with isolated Zr atoms sites than with Zr atoms that are part of a ZrO_x oligomer. Thus we attribute the decrease in the ketonization activity for catalysts with higher Zr site densities to the formation of ZrO_x dimers and oligomers, in a manner similar to what we previously reported for aldol condensation on TiO_x species [2].

To obtain additional support for the above hypothesis, the silica support was calcined at progressively higher temperatures in order to achieve spatial separation of the silanol groups prior to grafting Cp_2ZrCl_2 onto the support. ^{29}Si MAS NMR indicates that silica pretreated at 373 K and 1023 K possess nearly 4 and 1.6 SiOH nm^{-2} , respectively [6]. As shown in Table A.1, the TOF of $0.1\text{ZrCp}/\text{SiO}_2$ catalyst prepared on silica pretreated at 1023 K is roughly twice as high as that for $0.1\text{ZrCp}/\text{SiO}_2$ prepared on silica pretreated at 373 K. This inverse correlation between the activity and the density of support silanol groups supports the idea that isolated ZrOH centers are most active for carboxylic acid ketonization. By contrast, Zr catalysts synthesized by incipient wetness impregnation using $\text{Zr}(\text{O}^i\text{Pr})_4$, showed a much lower TOF than those prepared by grafting Cp_2ZrCl_2 for similar Zr loadings (Table A.1). These findings are consistent with previous studies done on titania dispersed on silica, which show that a metal alkoxide precursor such as $\text{Ti}(\text{O}^i\text{Pr})_4$ reacts with surface silanols to form dinuclear Ti complexes [7].

Further experiments were undertaken to obtain information about the mechanism of the reaction. The first of these involved examination of the inhibiting effects of ketone, CO_2 and water, the products formed during the ketonization of propanoic acid over ZrCp/SiO_2 . We also explored the effect of co-feeding pyridine, a Lewis base expected to inhibit the reaction. As illustrated in Figure A.4, co-feeding pyridine or water with the reactants produced an approximately two-fold decrease in the rate of ketone formation. However, co-feeding CO_2 decreased the rate of ketone formation only marginally. These results show that activity of $0.1\text{ZrCp}/\text{SiO}_2$ for the ketonization of propanoic acid decreases in the order: $\text{H}_2\text{O} > \text{C}_6\text{H}_5\text{N} > \text{CO}_2$. The data in Figure A.4 also show that upon removal of CO_2 from the feed, the activity returned to that observed prior to the addition of this compound; however, in the cases of water and pyridine addition, the original activity was only partially restored. The incomplete recovery of activity after the removal of pyridine or water from the feed suggests that Lewis acidic sites are more relevant than the Brønsted basic sites (e.g., Zr-OH) for the ketonization of propanoic acid over isolated sites of Zr supported on SiO_2 .

To confirm that isolated $\equiv\text{Zr-OH}$ sites only contain Lewis acidic sites, and not Brønsted acid sites, we acquired IR spectra of adsorbed pyridine on $0.1\text{ZrCp}/\text{SiO}_2$. Prior work has shown that IR peaks at 1445 , 1575 , and 1605 cm^{-1} are attributable to strong Lewis acid sites; a peak at 1490 cm^{-1} , to a combination of Lewis and Brønsted acid sites; and peaks at

1540 and 1640 cm^{-1} , to Brønsted acid sites [8]. The IR spectra shown in Figure A.9 indicates that 0.1ZrCp/SiO₂ possess pyridine bonded to strong Lewis acid sites (peaks at 1445, 1575, and 1605 cm^{-1}) and no evidence for pyridine interacting with Brønsted acid sites.

In situ IR spectroscopy was used to probe the form in which a carboxylic acid interacts with isolated Zr sites on 0.1ZrCp/SiO₂. As seen in Figure A.5, upon passing pulses of acetic acid over 0.1ZrCp/SiO₂ at 373 K, a peak was observed at 1745 cm^{-1} for acetic acid adsorbed on the silanol groups of silica, and a second peak was observed at 1680 cm^{-1} attributable to the interaction of -C=O species in the carboxylic acids with the Lewis acidic Zr sites. Interestingly, no peak was observed near 1580 cm^{-1} characteristic of the formation of bidentate species. By contrast, peaks were observed at 1745, 1680 cm^{-1} and 1580 cm^{-1} for 1.0ZrCp/SiO₂ [9, 10, 11]. As mentioned, the first two peaks are attributable to acetic acid adsorbed on silanol groups and on Lewis-acidic Zr sites, whereas the third peak is assigned to the formation of bidentate carboxylate species formed by the dissociative adsorption of carboxylic acids on Lewis acid pairs. Previous studies have proposed that bidentate carboxylates are the most abundant surface species formed on the surface of metal oxides during ketonization [12, 13, 14, 15, 16]. However, our results suggest that for isolated Zr supported on silica, monodentate carboxylate species are the active sites for ketonization of carboxylic acids. This interpretation is also in accord with recent DFT calculations over anatase TiO₂, which show that bidentate carboxylate species are not involved in ketonization, but act instead as spectators [17].

Additional insights into the reaction mechanism were obtained by exploring the reaction of a number of substituted carboxylic acids. Substituting the α -H of propanoic acid with one or two methyl groups decreased the rate of reaction, and carboxylic acids without any α -H were almost inactive (Figure A.10). Moving the methyl group from the α - to the β -position decreased the rate of ketonization by a factor of two relative to that for carboxylic acids without any substituents. These results indicate that the presence of an α -H is crucial for ketonic decarboxylation to occur [18].

H/D isotopic substitution was utilized to further probe the mechanism of propanoic acid ketonization over 0.1ZrCp/SiO₂ and 1.0ZrCp/SiO₂. The kinetic isotope effect, $k_{\text{H}}/k_{\text{D}}$, for ketonic decarboxylation of CH₃COOH versus CD₃COOD was nearly unity at 573 K (Figure A.11), suggesting that abstraction of α -H is not rate limiting and that instead C-C bond formation is more critical. The conclusion that a bimolecular reaction is rate limiting for ketonization of carboxylic acids differs from previous reports that concluded that the formation of a ketene intermediate or abstraction of α -H, are rate-limiting [19, 20, 21]. Nevertheless, our findings are consistent with recent DFT calculations, which suggest that C-C bond formation to produce a β -keto acid is the kinetically relevant step in the ketonic decarboxylation of monocarboxylic acids [13].

3. Supporting Mathematics

3.1. The Energetic Span Model

We model the turnover frequency using the energetic span model derived by Kozuch, which considers all intermediate states, transition states, and effects of reactant and product partial

pressures [22, 23]:

$$\begin{aligned}
\text{TOF} &= \frac{k_B T}{h} \frac{e^{-\beta \Delta G_{\text{rxn}}} \prod_k [R_k] - \prod_k [\mathcal{P}_k]}{\sum_{ij} e^{\beta(T_i - I_j - \delta G'_{ij})} \prod_h \delta \mathcal{R}'_{k,ij} \delta \mathcal{P}_{k,ij}} \\
&= \frac{k_B T}{h} \frac{e^{-\beta \Delta G_{\text{rxn}}} P_{\text{acid}}^2 - P_{\text{H}_2\text{O}} P_{\text{CO}_2} P_{\text{ketone}}}{\sum_{ij} e^{\beta(T_i - I_j - \delta G'_{ij})} \prod_h \delta \mathcal{R}'_{k,ij} \delta \mathcal{P}_{k,ij}} \tag{A.1} \\
\delta G'_{ij} &= \begin{cases} \Delta G_{\text{rxn}} & \text{if } i > j \\ 0 & \text{if } i \leq j \end{cases}
\end{aligned}$$

Here, T_i refers to the free energy of the i th transition state before the i th intermediate, and I_j refers to the free energy of the j th intermediate state after the j th transition state. All values of T_i and I_j are energies with respect to a single reference state. Where there is no explicit transition state for T_i , the transition state is considered to be the higher of the previous intermediate I_{i-1} and the following intermediate I_i .

The term $\delta \mathcal{R}'_{ij}$ is equal to the partial pressure of reactant k if, between intermediate I_j and T_i , the reactant is *not* consumed, and is equal to unity otherwise. Reactants that enter the catalytic cycle are considered to do so on the uphill of the relevant transition state's formation, while products that leave the catalytic cycle are considered to do so on the downhill from the relevant transition state. The result for the partial pressure term in the denominator for all i and j in this reaction is given by the matrix

$$\prod_k \delta \mathcal{R}'_{k,ij} \delta \mathcal{P}_{k,ij} = \begin{bmatrix} cbd & bcd & bcd & cd & cd & d & ad & ad & a \\ a^2 & bcd & bcd & cd & cd & cd & d & ad & ad & a \\ a^2 & a^2 & bcd & cd & cd & cd & d & ad & ad & a \\ a^2 & a^2 & a^2 & cd & cd & cd & d & ad & ad & a \\ a^2 b & a^2 b & a^2 b & a^2 & bcd & bcd & bcd & abd & abd & ab \\ a^2 b & a^2 b & a^2 b & a^2 & a^2 & bcd & bcd & abd & abd & ab \\ a^2 b & a^2 b & a^2 b & a^2 & a^2 & a^2 & bcd & abd & abd & ab \\ abc & abc & abc & ac & ac & ac & a & bcd & bcd & ab \\ abc & abc & abc & ac & ac & ac & a & a^2 & bcd & ab \\ abc & abc & abc & ac & ac & ac & a & a^2 & a^2 & ab \end{bmatrix} \tag{A.2}$$

where a corresponds to the partial pressure of the acid, b corresponds to the partial pressure of H_2O byproduct, c to that of CO_2 byproduct, and d to that of the ketone product. Equation A.2 simplifies when product partial pressures are approximated as zero. This assumption is deemed appropriate because the reaction is considered under differential conditions, meaning the conversion of products to reactants is less than 2%. Therefore, only a few terms in the matrix need to be considered:

$$\prod_k \delta \mathcal{R}'_{k,ij} \delta \mathcal{P}_{k,ij} = \begin{bmatrix} 0 & & & & & & & & & a \\ a^2 & & & & & & & & & a \\ a^2 & a^2 & & & & & & & & a \\ a^2 & a^2 & a^2 & & & & & & & a \\ & & & a^2 & & & & & & \\ & & & a^2 & a^2 & & & & & \\ & & & a^2 & a^2 & a^2 & & & & \\ & & & & & & a & & & \\ & & & & & & a & a^2 & & \\ & & & & & & a & a^2 & a^2 & 0 \end{bmatrix} \quad (\text{A.3})$$

The exponential terms in the denominator of Equation A.1 are dominated by the combinations of transition state T_i and intermediate I_j free energies whose differences are greatest. Therefore, the TOF can be further simplified considering only the dominant combinations of T_i and I_j . In our case, the free energy difference of the C-C bond formation transition state D^\ddagger and the singly adsorbed propanoic acid state B ($T_3 - I_{10/0}$), and the difference in free energy of D^\ddagger and the doubly adsorbed propanoic acid state C ($T_3 - I_1$) are sufficiently large that their exponential values dominate the denominator of Equation A.1. Their associated partial pressure weights are indicated in red text in Equation A.3. The TOF simplifies to

$$\text{TOF} = \frac{k_B T}{h} \frac{P_{\text{acid}} e^{-\beta \Delta G_{\text{rxn}}}}{P_{\text{acid}} e^{\beta(T_3 - I_1 - \Delta G_{\text{rxn}})} + e^{\beta(T_3 - I_{10})}} \quad (\text{A.4})$$

The percent error of this simplification to the full implementation of the model is on the order of 10^{-3} s^{-1} for all relevant values of reactant partial pressure.

3.2. Generalized Degrees of Rate Control and Apparent Activation Energy

Using our expression of TOF derived from the energetic span model, we quantified the relative importance of each intermediate and transition state (denoted by index i) to the total rate using an expression derived by Campbell [24].

$$X_i = \frac{\partial \log \text{TOF}}{\partial (-\beta G_i)} \quad (\text{A.5})$$

Since our TOF expression depends on the energies of three intermediates and transition states, DRCs for these intermediates were calculated using Equation A.1.

$$X_B = -\frac{e^{\beta(G_C + \Delta G_{\text{rxn}})}}{P_{\text{acid}} e^{\beta G_B} + e^{\beta(G_C + \Delta G_{\text{rxn}})}} \quad (\text{A.6})$$

$$X_C = -\frac{P_{\text{acid}} e^{\beta G_B}}{P_{\text{acid}} e^{\beta G_B} + e^{\beta(G_C + \Delta G_{\text{rxn}})}} \quad (\text{A.7})$$

$$X_{D^\ddagger} = 1 \quad (\text{A.8})$$

The dependence of the intermediate DRCs on reactant partial pressure is shown in Figure A.14. The apparent activation energy is then calculated as a sum of RT and each intermediate and transition state enthalpy, weighted by their respective DRCs [25]. The result is an apparent activation energy as a function of reactant partial pressure.

$$E_{\text{app}} = RT + X_B H_B^0 + X_C H_C^0 + X_{D^\ddagger} H_{D^\ddagger}^0 \quad (\text{A.9})$$

3.3. Gibbs Activation Energy

As with the derivation for Kozuch's energetic span model, the effective rate of the reaction can be related to an effective Gibbs free energy of activation for the reaction.

$$\text{TOF} = \frac{k_B T}{h} e^{-\beta \Delta G^\ddagger} \quad (\text{A.10})$$

$$\Delta G^\ddagger = -RT \log \left(\frac{h}{k_B T} \text{TOF} \right) \quad (\text{A.11})$$

3.4. Energetic Span Model, Generalized Degrees of Rate Control, and Apparent Activation Energy: Aldol Condensation

Given the reaction mechanism for aldol condensation depicted in Scheme II.2, and assuming negligible product partial pressures under differential conversion, the energetic span model simplifies to

$$\text{TOF} = \frac{k_B T}{h} \frac{P_{\text{ketone}}^2 e^{-\beta \Delta G_{\text{rxn}}}}{\mathcal{Z}} \quad (\text{A.12})$$

$$\begin{aligned} \mathcal{Z} = & P_{\text{ketone}}^2 \left(e^{\beta(G_{D^\ddagger} - G_C - \Delta G_{\text{rxn}})} + e^{\beta(G_{C^\ddagger} - G_C - \Delta G_{\text{rxn}})} \right) + \\ & P_{\text{ketone}} \left(e^{\beta(G_{D^\ddagger} - G_C - \Delta G_{\text{rxn}})} + e^{\beta(G_{C^\ddagger} - G_C - \Delta G_{\text{rxn}})} \right) + \\ & e^{\beta(G_{D^\ddagger} - G_A)} + e^{\beta(G_{C^\ddagger} - G_A)} \end{aligned}$$

The associated generalized degrees of rate controls and the apparent activation energy are as follows

$$X_A = - \frac{e^{\beta(G_B + G_C + \Delta G_{\text{rxn}})}}{P_{\text{ketone}}^2 e^{\beta(G_A + G_B)} + P_{\text{ketone}} e^{\beta(G_A + G_C)} + e^{\beta(G_B + G_C + \Delta G_{\text{rxn}})}} \quad (\text{A.13})$$

$$X_B = - \frac{P_{\text{ketone}} e^{\beta(G_A + G_C)}}{P_{\text{ketone}}^2 e^{\beta(G_A + G_B)} + P_{\text{ketone}} e^{\beta(G_A + G_C)} + e^{\beta(G_B + G_C + \Delta G_{\text{rxn}})}} \quad (\text{A.14})$$

$$X_C = - \frac{P_{\text{ketone}}^2 e^{\beta(G_A + G_B)}}{P_{\text{ketone}}^2 e^{\beta(G_A + G_B)} + P_{\text{ketone}} e^{\beta(G_A + G_C)} + e^{\beta(G_B + G_C + \Delta G_{\text{rxn}})}} \quad (\text{A.15})$$

$$X_{C^\ddagger} = \frac{e^{\beta G_{C^\ddagger}}}{e^{\beta G_{C^\ddagger}} + e^{\beta G_{D^\ddagger}}} \quad (\text{A.16})$$

$$X_{D^\ddagger} = \frac{e^{\beta G_{D^\ddagger}}}{e^{\beta G_{C^\ddagger}} + e^{\beta G_{D^\ddagger}}} \quad (\text{A.17})$$

$$E_{\text{app}} = RT + X_A H_A^0 + X_B H_B^0 + X_C H_C^0 + X_{C^\ddagger} H_{C^\ddagger}^0 + X_{D^\ddagger} H_{D^\ddagger}^0 \quad (\text{A.18})$$

4. Supporting Figures

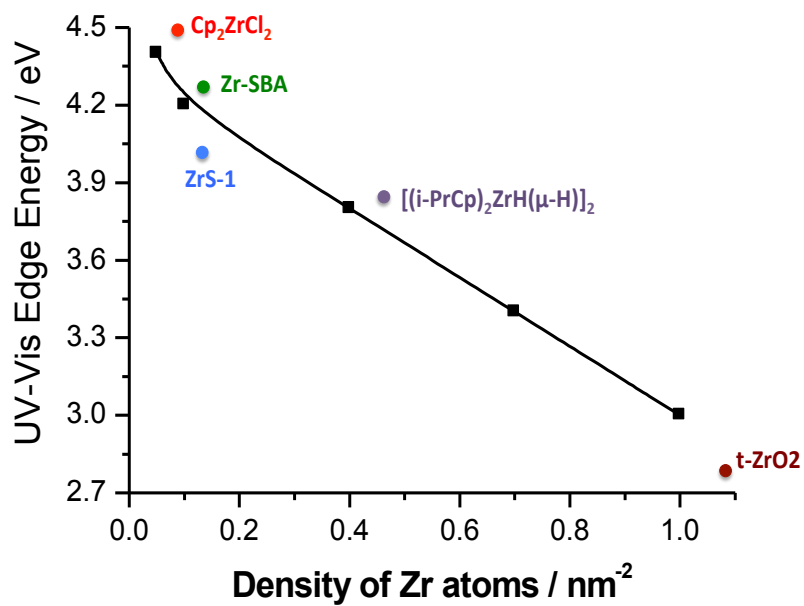


Figure A.1: UV-Vis edge energies of $x\text{ZrCp}/\text{SiO}_2$ as a function of Zr surface density and Zr-containing model compounds.

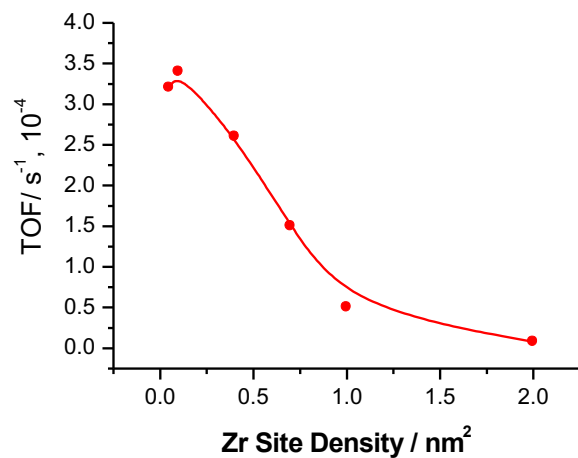


Figure A.2: Effect of Zr site density in the activity of $x\text{ZrCp}/\text{SiO}_2$ catalysts for ketonization of propanoic acid. Reaction conditions: $T = 573 \text{ K}$, $P_{\text{Tot.}} = 1 \text{ atm}$, $Q_{\text{Tot.}} = 100 \text{ cm}^3 \text{ min}^{-1}$, $P_{\text{acid}} \approx 0.2 \text{ kPa}$, $M_{\text{Cat}} = 0.1 \text{ g}$.

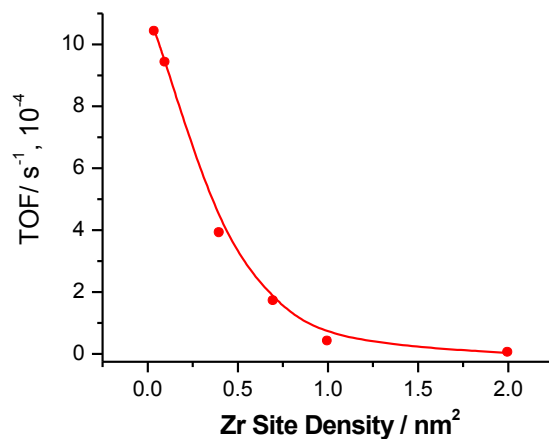


Figure A.3: Effect of Zr site density in the activity of $x\text{ZrCp}/\text{SiO}_2$ catalysts for the aldol condensation of propan-2-one. Reaction conditions: $T = 573 \text{ K}$, $P_{\text{Tot.}} = 1 \text{ atm}$, $Q_{\text{Tot.}} = 100 \text{ cm}^3 \text{ min}^{-1}$, $P_{\text{acid}} \approx 0.2 \text{ kPa}$, $M_{\text{Cat}} = 0.1 \text{ g}$.

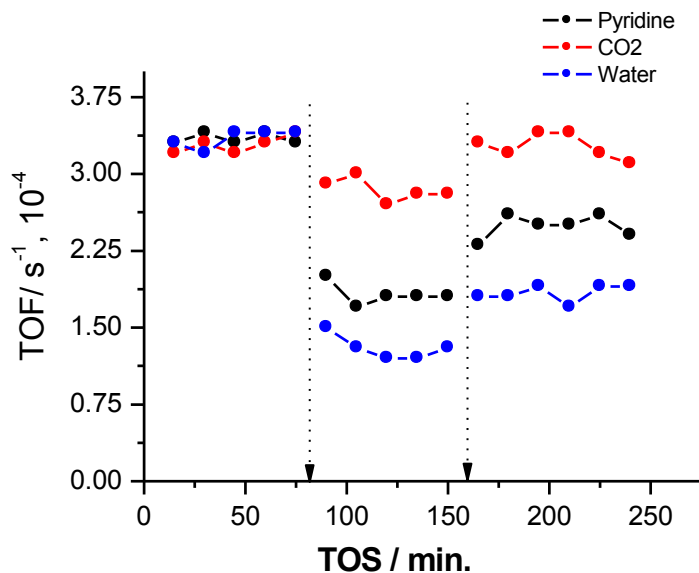


Figure A.4: Poisoning experiments conducted over $0.1\text{ZrCp}/\text{SiO}_2$ catalysts in the ketonic decarboxylation of propanoic acid. Dotted lines represent the time at which 10 wt.% of the poisoning agent was co-fed together with propanoic acid in the gas stream. Reaction conditions: $T = 573 \text{ K}$, $P_{\text{Tot.}} = 1 \text{ atm}$, $Q_{\text{Tot.}} = 100 \text{ cm}^3 \text{ min}^{-1}$, $P_{\text{acid}} \approx 0.2 \text{ kPa}$, $M_{\text{Cat}} = 0.1 \text{ g}$.

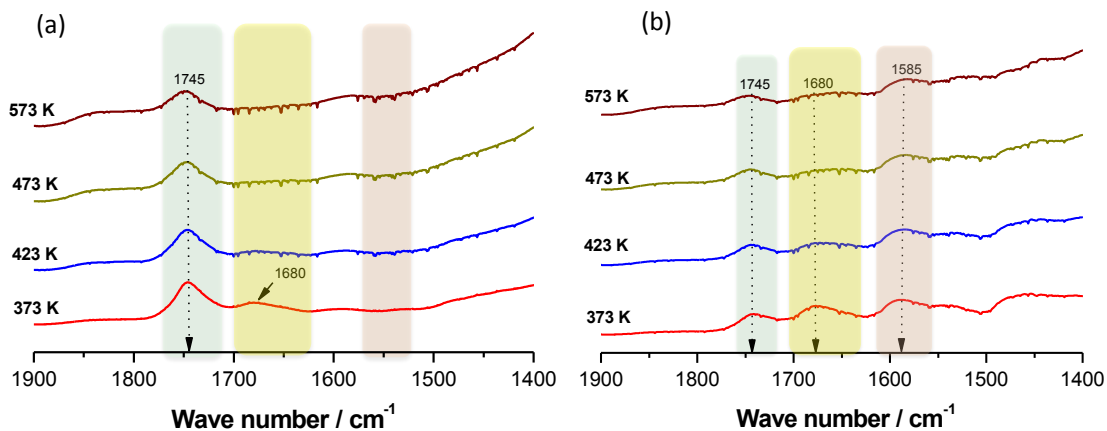
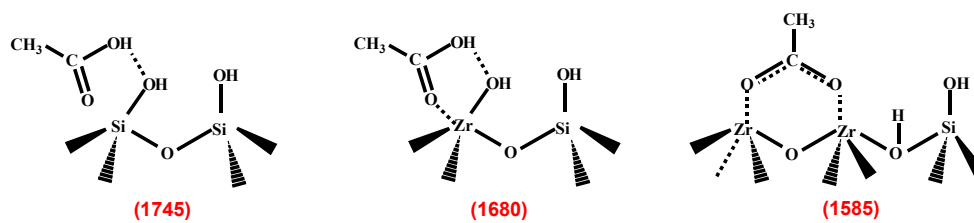


Figure A.5: In situ FT-IR spectra of acetic acid adsorption recorded at room temperature and during desorption at various temperatures. (a) $0.1\text{ZrCp}/\text{SiO}_2$ and (b) $1.0\text{ZrCp}/\text{SiO}_2$ catalyst.



Scheme A.1: Possible surface species formed in the ketonic decarboxylation reaction over ZrCp/SiO₂ catalysts and their respective FT-IR frequencies.

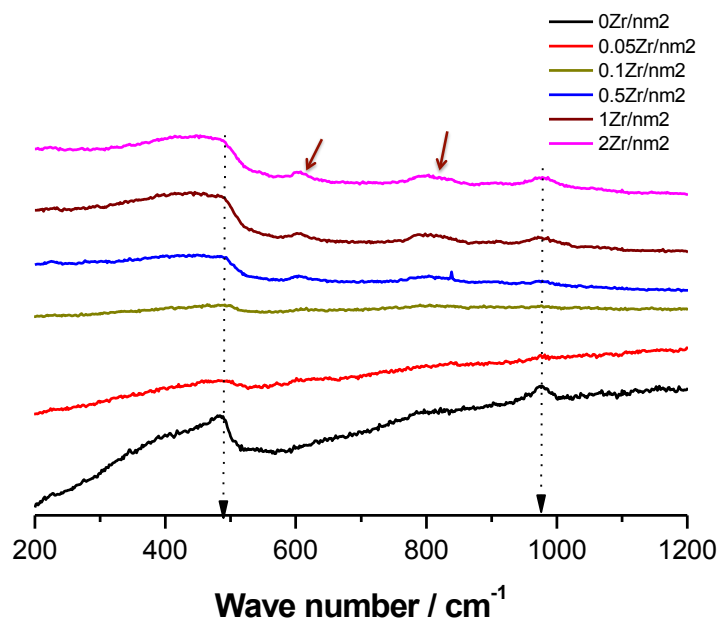


Figure A.6: Raman spectra of support silica and various Zr loaded silica samples.

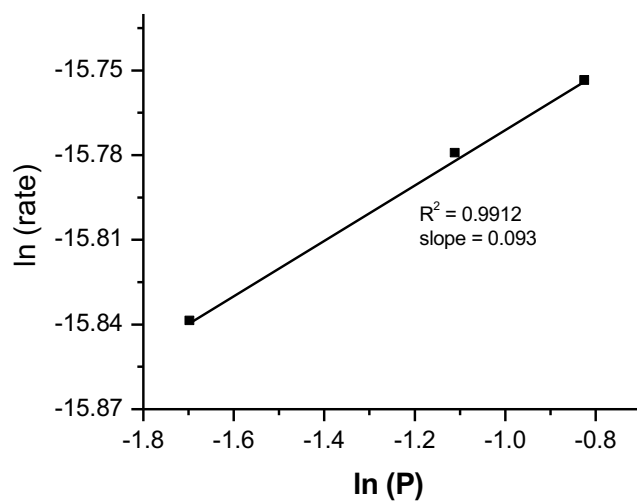


Figure A.7: Partial pressure dependency of 0.1ZrCp/SiO₂ catalysts in the ketonization reaction of propanoic acid.

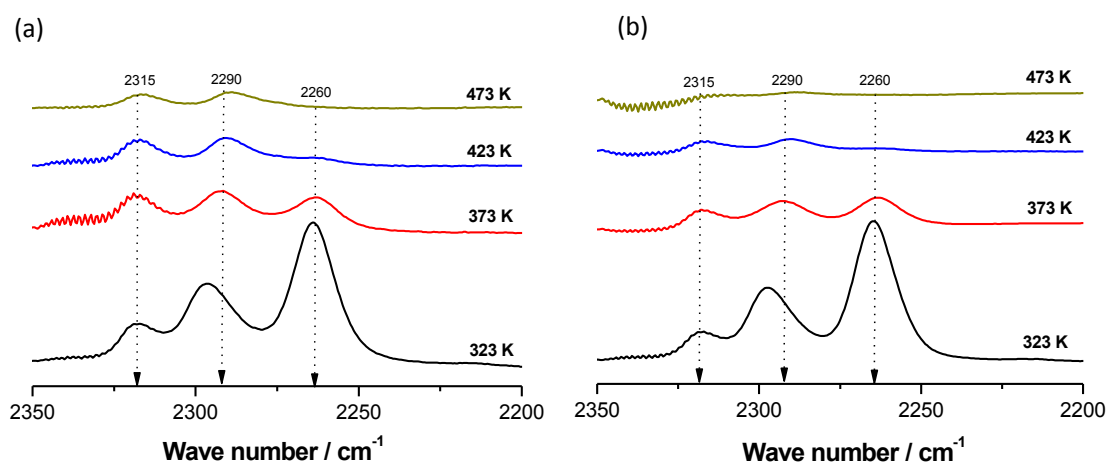


Figure A.8: In situ FT-IR spectra of CH₃CN adsorption at room temperature and its desorption at various temperatures. (a) 0.1ZrCp/SiO₂ and (b) 1.0ZrCp/SiO₂ catalyst.

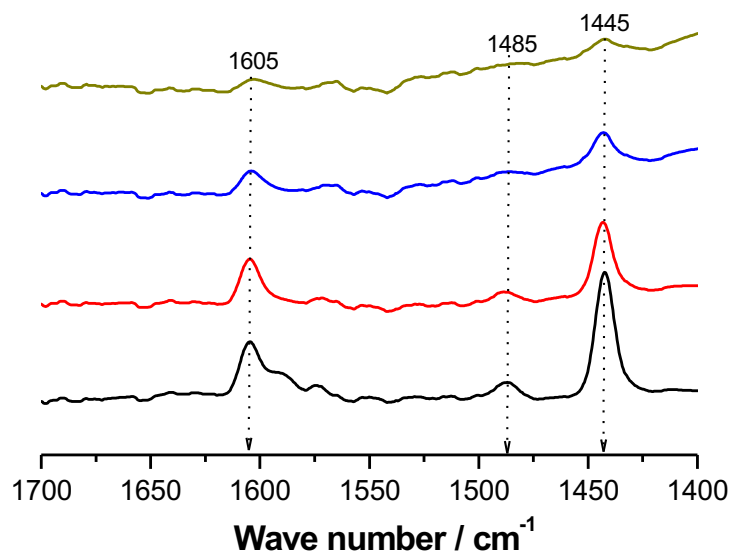


Figure A.9: Pyridine IR studies over a ZrCp/SiO₂ catalyst. Lewis acid sites are noted in the IR spectrum and a decrease in peak intensity is noted with an increase in temperature from 50 to 200 °C (bottom to top).

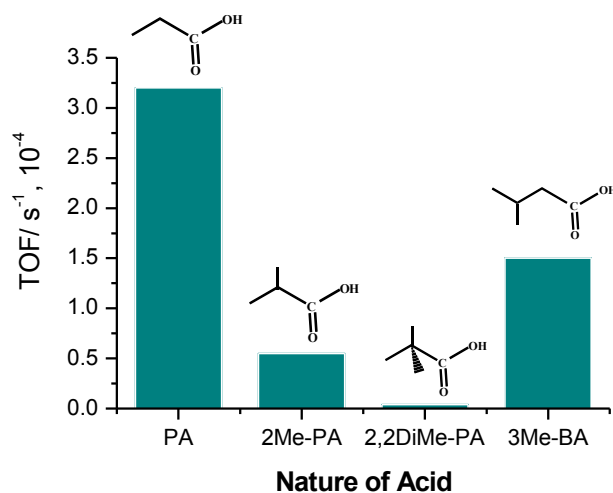


Figure A.10: Catalytic activity of propanoic acid and various substituted carboxylic acids in the ketonization reactions, where PA = propanoic acid, 2Me-PA = 2-methyl propanoic acid, 2,2DiMe-PA = 2,2-dimethyl propanoic acid, and 3Me-BA = 3-methyl butyric acid. Reaction conditions: $T = 573$ K, $P_{\text{Tot.}} = 1$ atm, $Q_{\text{Tot.}} = 100$ cm³ min⁻¹, $P_{\text{acid}} \approx 0.2$ kPa, $M_{\text{Cat}} = 0.1$ g.

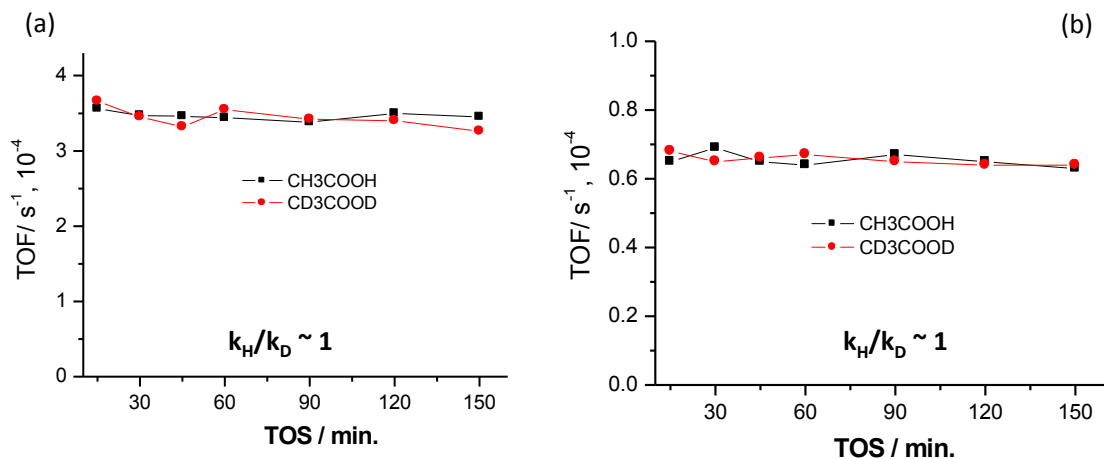


Figure A.11: Kinetic isotope effect over (a) 0.1ZrCp/SiO₂ and (b) 1.0ZrCp/SiO₂ catalyst using acetic acid and deuterated acetic acid substrate. Reaction conditions: $T = 573 \text{ K}$, $P_{\text{Tot.}} = 1 \text{ atm}$, $Q_{\text{Tot.}} = 100 \text{ cm}^3 \text{ min}^{-1}$, $P_{\text{acid}} \approx 0.2 \text{ kPa}$, $M_{\text{Cat}} = 0.1 \text{ g}$.

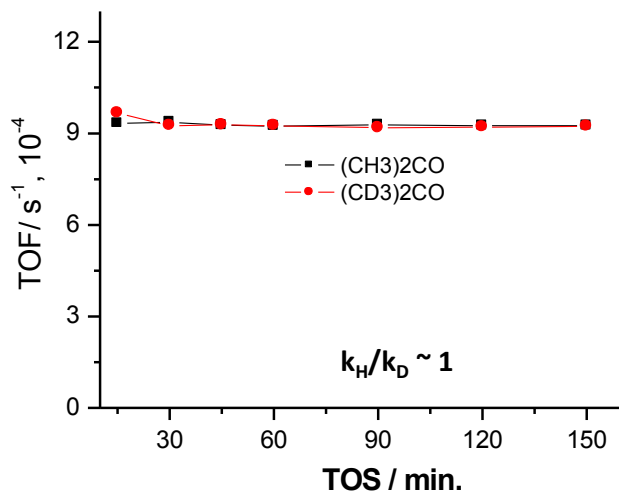


Figure A.12: Kinetic isotope effect over a 1.0ZrCp/SiO₂ catalyst using propan-2-one and deuterated propan-2-one substrate. Reaction conditions: $T = 473 \text{ K}$, $P_{\text{Tot.}} = 1 \text{ atm}$, $Q_{\text{Tot.}} = 100 \text{ cm}^3 \text{ min}^{-1}$, $P_{\text{ketone}} \approx 0.2 \text{ kPa}$, $M_{\text{Cat}} = 0.1 \text{ g}$.

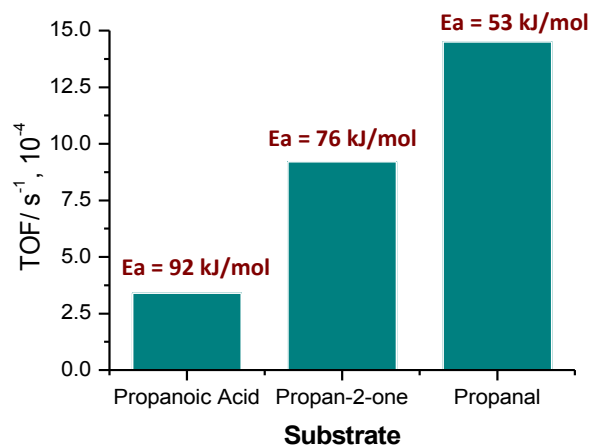


Figure A.13: Comparison in TOF and activation energy of carboxylic acid, ketone, and aldehyde. Reaction conditions: $T = 573$ K, $P_{\text{Tot.}} = 1$ atm, $Q_{\text{Tot.}} = 100$ cm³ min⁻¹, $P_{\text{acid}} \approx 0.2$ kPa, $M_{\text{Cat}} = 0.1$ g.

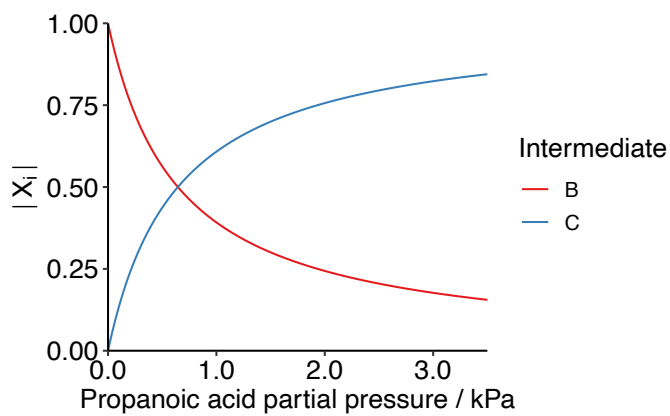


Figure A.14: Relative degrees of TOF control of ketonization intermediates B and C as a function of reactant partial pressure, based on the approximation of TOF given in Equation 2.1.

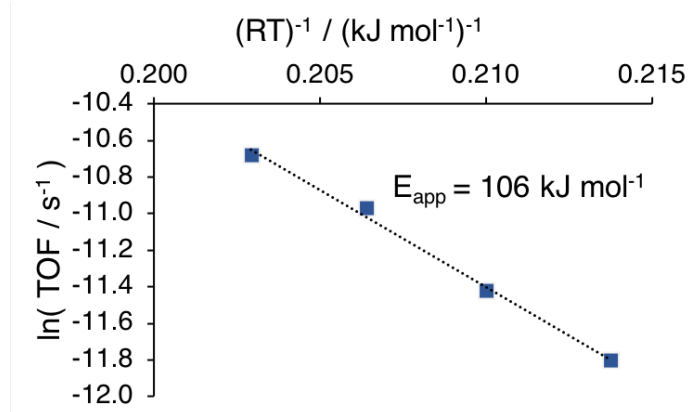


Figure A.15: Arrhenius plot for ketonization of propanoic acid on ZrOH sites using the free energy pathway of Figure 2. Values of TOF were calculated using the full expression of the energetic span model [22].

References

- [1] K Okumura and Y Iwasawa. “Zirconium Oxides Dispersed on Silica Derived from Cp_2ZrCl_2 , $[(i\text{-PrCp})_2\text{ZrH}(\mu\text{-H})_2]$, and $\text{Zr}(\text{OEt})_4$ Characterized by X-Ray Absorption Fine Structure and Catalytic Ketonization of Acetic Acid”. *J. Catal.* **164.2** (1996), 440–448
- [2] David G Hanna et al. “Experimental and Theoretical Study of n-Butanal Self-Condensation over Ti Species Supported on Silica”. *ACS Catal.* **4.9** (2014), 2908–2916
- [3] Meijun Li et al. “Phase Transformation in the Surface Region of Zirconia Detected by UV Raman Spectroscopy”. *J. Phys. Chem. B* **105.34** (2001), 8107–8111
- [4] MB Jensen et al. “FT-IR Characterization of Supported Ni-catalysts: Influence of Different Supports on the Metal Phase Properties”. *Catal. Today* **197.1** (2012), 38–49
- [5] JC Lavalley. “Infrared Spectrometric Studies of the Surface Basicity of Metal Oxides and Zeolites using Adsorbed Probe Molecules”. *Catal. Today* **27.3-4** (1996), 377–401
- [6] Sankaranarayananpillai Shylesh et al. “Factors Influencing the Activity, Selectivity, and Stability of Rh-based Supported Ionic Liquid Phase (SILP) Catalysts for Hydroformylation of Propene”. *ACS Catal.* **2.4** (2012), 487–493
- [7] Abdillahi Omar Bouh, Gordon L Rice, and Susannah L Scott. “Mono-and Dinuclear Silica-supported Titanium (IV) Complexes and the Effect of Ti-O-Ti Connectivity on Reactivity”. *J. Am. Chem. Soc.* **121.31** (1999), 7201–7210
- [8] Sankaranarayananpillai Shylesh et al. “Effects of Composition and Structure of Mg/Al Oxides on their Activity and Selectivity for the Condensation of Methyl Ketones”. *Ind. Eng. Chem. Res.* **55.40** (2016), 10635–10644
- [9] Gerald Jödicke et al. “Developing Environmentally-sound Processes in the Chemical Industry: a Case Study on Pharmaceutical Intermediates”. *J. Clean. Prod.* **7.2** (1999), 159–166
- [10] VN Panchenko et al. “DRIFTS and UV–vis DRS Study of Valeric Acid Ketonization Mechanism over ZrO_2 in Hydrogen Atmosphere”. *J. Mol. Catal. A: Chemical* **388** (2014), 133–140
- [11] Z-F Pei and V Ponec. “On the Intermediates of the Acetic Acid Reactions on Oxides: an IR Study”. *Appl. Surf. Sci.* **103.2** (1996), 171–182
- [12] Tu N Pham et al. “Ketonization of Carboxylic Acids: Mechanisms, Catalysts, and Implications for Biomass Conversion”. *ACS Catal.* **3.11** (2013), 2456–2473
- [13] James A Bennett et al. “Acetic Acid Ketonization over $\text{Fe}_3\text{O}_4/\text{SiO}_2$ for pyrolysis bio-oil upgrading”. *ChemCatChem* **9.9** (2017), 1648
- [14] Madhesan Balakrishnan et al. “Novel Pathways for Fuels and Lubricants from Biomass Optimized using Life-cycle Greenhouse Gas Assessment”. *Proc. Natl. Acad. Sci.* **112.25** (2015), 7645–7649

- [15] Gianfranco Pacchioni. “Ketonization of Carboxylic Acids in Biomass Conversion over TiO_2 and ZrO_2 Surfaces: a DFT Perspective”. *ACS Catal.* **4.9** (2014), 2874–2888
- [16] Angeles Pulido et al. “Ketonic Decarboxylation Reaction Mechanism: a Combined Experimental and DFT Study”. *ChemSusChem* **6.1** (2013), 141–151
- [17] Shuai Wang and Enrique Iglesia. “Experimental and Theoretical Assessment of the Mechanism and Site Requirements for Ketonization of Carboxylic Acids on Oxides”. *J. Catal.* **345** (2017), 183–206
- [18] Osamu Nagashima et al. “Ketonization of Carboxylic Acids over CeO_2 -based Composite Oxides”. *J. Mol. Catal. A. Chem.* **227.1-2** (2005), 231–239
- [19] Sandeep D. Randery, Jack S. Warren, and Kerry M. Dooley. “Cerium Oxide-based Catalysts for Production of Ketones by Acid Condensation”. *Appl. Catal. A: Gen.* **226.1** (2002), 265–280. ISSN: 0926-860X. DOI: [https://doi.org/10.1016/S0926-860X\(01\)00912-7](https://doi.org/10.1016/S0926-860X(01)00912-7). URL: <https://www.sciencedirect.com/science/article/pii/S0926860X01009127>
- [20] Kerry M. Dooley et al. “Ketones from Acid Condensation using Supported CeO_2 Catalysts: Effect of Additives”. *Appl. Catal. A: Gen.* **320** (2007), 122–133. ISSN: 0926-860X. DOI: <https://doi.org/10.1016/j.apcata.2007.01.021>. URL: <https://www.sciencedirect.com/science/article/pii/S0926860X07000361>
- [21] Travis S. Hendren and Kerry M. Dooley. “Kinetics of Catalyzed Acid/acid and Acid/aldehyde Condensation Reactions to Non-symmetric Ketones”. *Catal. Today* **85.2** (2003). Metallic Oxides: Filling the Gap between Catalysis and Surface Science, 333–351. ISSN: 0920-5861. DOI: [https://doi.org/10.1016/S0920-5861\(03\)00399-7](https://doi.org/10.1016/S0920-5861(03)00399-7). URL: <https://www.sciencedirect.com/science/article/pii/S0920586103003997>
- [22] Sebastian Kozuch and Sason Shaik. “How to Conceptualize Catalytic Cycles? The Energetic Span Model”. *Acc. Chem. Res.* **44.2** (2011), 101–110
- [23] Sebastian Kozuch. “A Refinement of Everyday Thinking: the Energetic Span Model for Kinetic Assessment of Catalytic Cycles”. *Wiley Interdiscip. Rev. Comp. Mol. Sci.* **2.5** (2012), 795–815
- [24] Charles T Campbell. *The Degree of Rate Control: a Powerful Tool for Catalysis Research*. 2017
- [25] Zhongtian Mao and Charles T Campbell. “Apparent Activation Energies in Complex Reaction Mechanisms: a Simple Relationship via Degrees of Rate Control”. *ACS Catal.* **9.10** (2019), 9465–9473

Appendix B

Supplementary Information for Chapter III

1. Thermodynamic Derivations

Thermodynamic quantities are statistically-rooted values that can be derived from the partition function Q , which in the canonical ensemble is defined by a summation of the allowed states. The Helmholtz Free Energy, for example, is computed through the logarithm of the partition function as

$$A = -RT \log Q, \quad (\text{B.1})$$

the internal energy is given by

$$E = -\frac{\partial \log Q}{\partial \beta}, \quad (\text{B.2})$$

the entropy is computed by the relationship

$$S = \frac{E - A}{T}, \quad (\text{B.3})$$

and the heat capacity is computed as

$$C_V = \frac{\partial E}{\partial T} = \frac{1}{RT^2} \frac{\partial^2 \log Q}{\partial \beta^2}. \quad (\text{B.4})$$

Under the assumption of separable translational, rotational, and internal degrees of freedom, Q is defined as the product of each separable component

$$Q = Q_{\text{trans}} Q_{\text{rot}} Q_{\text{int}} \quad (\text{B.5})$$

This work explores the internal degrees of freedom, which can be further separated into stretches, bends, and torsions

$$Q_{\text{int}} = Q_{\text{sb}} Q_{\text{tors}} \quad (\text{B.6})$$

The torsional contributions to the free energy, internal energy, and heat capacity are reported for several approximation schemes in the following subsections. The entropy relationship, while not explicitly stated, is calculated from Equation B.3.

1.1. Quantum Harmonic Oscillator

The quantum harmonic oscillator Schrödinger Equation has equally spaced energy levels

$$E_v = \hbar\omega\left(v + \frac{1}{2}\right) \quad \forall v \in \mathbb{Z}^{\geq} \quad (\text{B.7})$$

which leads directly to the zero point energy and quantum harmonic oscillator partition function.

$$E_0^{\text{HO}} = \frac{1}{2}\hbar\omega \quad (\text{B.8})$$

$$Q^{\text{qu,HO}} = \frac{1}{2 \sinh(\beta E_0)} \quad (\text{B.9})$$

$$(\text{B.10})$$

The latter is a direct result of simplifying the geometric series to its closed form. The internal energies and heat capacities follow by derivation, algebraically simplified:

$$E^{\text{qu,HO}} = \frac{E_0}{\tanh(\beta E_0)} \quad (\text{B.11})$$

$$C_V^{\text{qu,HO}} = R \left(\frac{\beta E_0}{\sinh(\beta E_0)} \right)^2 \quad (\text{B.12})$$

1.2. Classical Harmonic Oscillator

The classical harmonic oscillator partition function is simply calculated by two Gaussian integrals, one momentum integral and one position integral. This simplifies to

$$Q^{\text{cl,HO}} = (\beta\hbar\omega)^{-1} \quad (\text{B.13})$$

The derivation from Equations B.2 and B.4 is straightforward.

$$E^{\text{cl,HO}} = RT \quad (\text{B.14})$$

$$C_V^{\text{cl,HO}} = R \quad (\text{B.15})$$

1.3. Classical Uncoupled Mode Approximation

Beginning with the classical partition function for an uncoupled torsion under the UM-T approximations given in the manuscript,

$$Q_{\text{tors},\tau}^{\text{cl,UM-T}} = \left(\frac{I_\tau}{2\pi\beta\hbar^2} \right)^{1/2} \int_0^{2\pi/\sigma_\tau} d\phi_\tau e^{-\beta V_\tau(\phi_\tau)}, \quad (\text{B.16})$$

the internal energy and heat capacities are then given by the equations

$$E_{\text{tors},\tau}^{\text{cl,UM-T}} = \frac{1}{2\beta} + \langle V_\tau \rangle \quad (\text{B.17})$$

$$C_{V,\text{tors},\tau}^{\text{cl,UM-T}} = \frac{1}{2}R + \frac{1}{RT^2} \int_0^{2\pi/\sigma_\tau} d\phi_\tau (V_\tau(\phi_\tau) - \langle V_\tau \rangle)^2 f(\phi) \quad (\text{B.18})$$

where $\langle V_\tau \rangle$ is the average potential energy of the τ th torsion,

$$\langle V_\tau \rangle = \int_0^{2\pi/\sigma_\tau} d\phi_\tau V_\tau(\phi_\tau) f(\phi_\tau) \quad (\text{B.19})$$

and f is the associated probability density function in the canonical ensemble,

$$f(\phi) = \frac{e^{-\beta V(\phi)}}{\int_0^{2\pi/\sigma} d\phi e^{-\beta V(\phi)}} \quad (\text{B.20})$$

Bounds of integration correct for symmetry effects by division by the torsional symmetry factor σ_τ . These integrals are solved numerically, with expressions for V_τ obtained by cubic spline interpolation of points sampled along the dihedral coordinate.

1.4. Classical Coupled Mode Approximation

The CM-T thermodynamics derived from $Q_{\text{tors}}^{\text{cl,CM-T}}$ are given in the main text. The energy and heat capacities are straightforward, and can be compared against the uncoupled mode quantities whose mathematics are similar. They are repeated here, with $\langle V \rangle$ representing the expected value of the coupled torsional potential energy:

$$E_{\text{tors}}^{\text{cl,CM-T}} = \frac{N}{2\beta} + \langle V \rangle \quad (\text{B.21})$$

$$C_{V,\text{tors}}^{\text{cl,CM-T}} = \frac{N}{2}R + \frac{\langle V^2 \rangle - \langle V \rangle^2}{RT^2} \quad (\text{B.22})$$

The entropy term of Equation 3.28 is derived beginning with Equations B.1 and B.3:

$$S_{\text{tors}}^{\text{cl,CM-T}} = \frac{E_{\text{tors}}^{\text{cl,CM-T}}}{T} + R \ln Q_{\text{tors}}^{\text{cl,CM-T}} \quad (\text{B.23})$$

where $Q_{\text{tors}}^{\text{cl,CM-T}}$ is the classical CM-T partition function given by Equation 3.17 in the main text. Using

$$\langle |D|^{1/2} \rangle = \frac{\int d^N \Phi |D|^{1/2} e^{-\beta V(\Phi)}}{\int d^N \Phi e^{-\beta V(\Phi)}}, \quad (\text{B.24})$$

Equation 3.17 becomes

$$Q_{\text{tors}}^{\text{cl,CM-T}} = \left(\frac{1}{2\pi\beta\hbar^2} \right)^{N/2} \langle |D|^{1/2} \rangle \int d^N \Phi e^{-\beta V(\Phi)} \quad (\text{B.25})$$

Using this result, we separate Equation B.23 into kinetic (\mathcal{S}) and potential (V) contributions, arriving at

$$S_{\text{tors}}^{\text{cl,CM-T}} = \left[\frac{\langle \mathcal{S} \rangle}{T} + R \ln \left(\frac{1}{2\pi\beta\hbar^2} \right)^{N/2} + R \ln \langle |D|^{1/2} \rangle \right] + \left[\frac{\langle V \rangle}{T} + R \ln \int d^N \Phi e^{-\beta V(\Phi)} \right] \quad (\text{B.26})$$

$$= R \left[\frac{N}{2} - \frac{N}{2} \ln(2\pi\beta\hbar^2) + \ln\langle |D|^{1/2} \rangle \right] - \langle R \ln f(\Phi) \rangle \quad (\text{B.27})$$

We now show the equivalence of $\langle -R \ln f(\Phi) \rangle$ with the potential contribution to the entropy. First, we have probability density function given by

$$f(\Phi) = \frac{e^{-\beta V(\Phi)}}{\int d^N \Phi e^{-\beta V(\Phi)}} = \frac{e^{-\beta V(\Phi)}}{\mathcal{Q}} \quad (\text{B.28})$$

Inserting this result into the logarithm term of the expression for the expected value, we obtain

$$\langle -R \ln f(\Phi) \rangle = -R \int d^N \Phi f(\Phi) \ln f(\Phi) \quad (\text{B.29})$$

$$= -R \int d^N \Phi \ln \left(\frac{e^{-\beta V(\Phi)}}{\mathcal{Q}} \right) f(\Phi) \quad (\text{B.30})$$

$$= -R \int d^N \Phi [(-\beta V(\Phi)) - \ln \mathcal{Q}] f(\Phi) \quad (\text{B.31})$$

$$= \frac{\langle V \rangle}{T} + R \ln \mathcal{Q} \int d^N \Phi f(\Phi) \quad (\text{B.32})$$

$$= \frac{\langle V \rangle}{T} + R \ln \int d^N \Phi e^{-\beta V(\Phi)} \quad (\text{B.33})$$

which is identically the potential contribution, as desired.

2. No U-Turn Sampling

The No U-Turn Sampling (NUTS) algorithm requires a positive integer number of leap frog steps L and a positive time step size Δt . Beginning at $t = 0$, for a position Φ of sample k in the Markov chain:

$$\Phi_k(0) = \Phi_k \quad (\text{B.34})$$

Hamiltonian Monte Carlo employs a random walk in momentum (rather than position, as is the case for Metropolis-Hastings), so that the momentum vector at $t = 0$ is drawn from

$$\mathbf{p}_k(0) \sim \mathcal{N}(\mathbf{0}, M) \quad (\text{B.35})$$

where M is the mass matrix appearing in the Hamiltonian and approximated during in the tuning schedule. Then, the position will evolve under Hamiltonian dynamics for time $L\Delta t$, solved numerically using the leap frog algorithm. The position and momentum vectors after time Δt using the leap frog algorithm are

$$\mathbf{p}_k \left(t + \frac{\Delta t}{2} \right) = \mathbf{p}_k(t) - \frac{\Delta t}{2} \nabla V(\Phi)|_{\Phi=\Phi_k(t)} \quad (\text{B.36})$$

$$\Phi_k(t + \Delta t) = \Phi_k(t) + \Delta t M^{-1} \mathbf{p}_k \left(t + \frac{\Delta t}{2} \right) \quad (\text{B.37})$$

$$\mathbf{p}_k(t + \Delta t) = \mathbf{p}_k \left(t + \frac{\Delta t}{2} \right) - \frac{\Delta t}{2} \nabla V(\Phi) |_{\Phi = \Phi_k(t + \Delta t)} \quad (\text{B.38})$$

These equations are applied to Φ_k and \mathbf{p}_k in the forward (Φ_k^+ and \mathbf{p}_k^+) and backward (Φ_k^- and \mathbf{p}_k^-) directions until the U-Turn condition is met. This is given by either criterion

$$(\Phi_k^+ - \Phi_k^-) \cdot \mathbf{p}_k^+ < 0 \quad (\text{B.39})$$

$$(\Phi_k^+ - \Phi_k^-) \cdot \mathbf{p}_k^- < 0 \quad (\text{B.40})$$

Once the U-Turn condition is met, the next MCMC sample, Φ_{k+1} , is obtained by sampling uniformly the leap frog path traced out by the binary tree

$$\{\Phi_k^-, \dots, \Phi_k(-\Delta t), \Phi_k(0), \Phi_k(\Delta t), \dots, \Phi_k^+\}$$

which satisfies

$$U_k < e^{-H(\Phi_{k+1}, \mathbf{p}_{k+1})} \quad (\text{B.41})$$

where $U_k \sim \text{Uniform}(0, \exp(-H(\Phi_k(0), \mathbf{p}_k(0))))$ is sampled. This procedure is covered comprehensively in the original literature [1].

3. Data Visualization

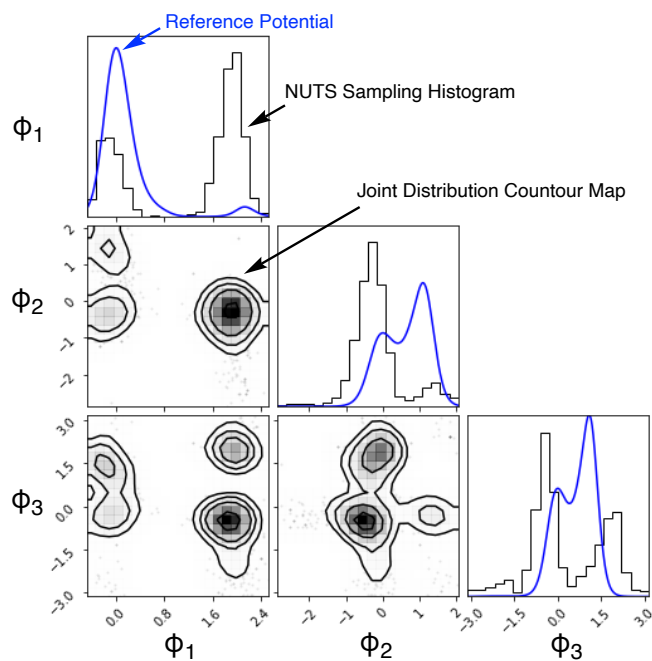


Figure B.1: Corner plot of internal rotation sampling for ethylene glycol. On the diagonal, sampling histograms for each internal rotation. Blue lines represent the uncoupled mode distributions for each internal rotation. On the off-diagonal, density plots of samples drawn from the joint distribution. Substantial coupling accounts for the difference between the reference potential and the mode-coupled potential.

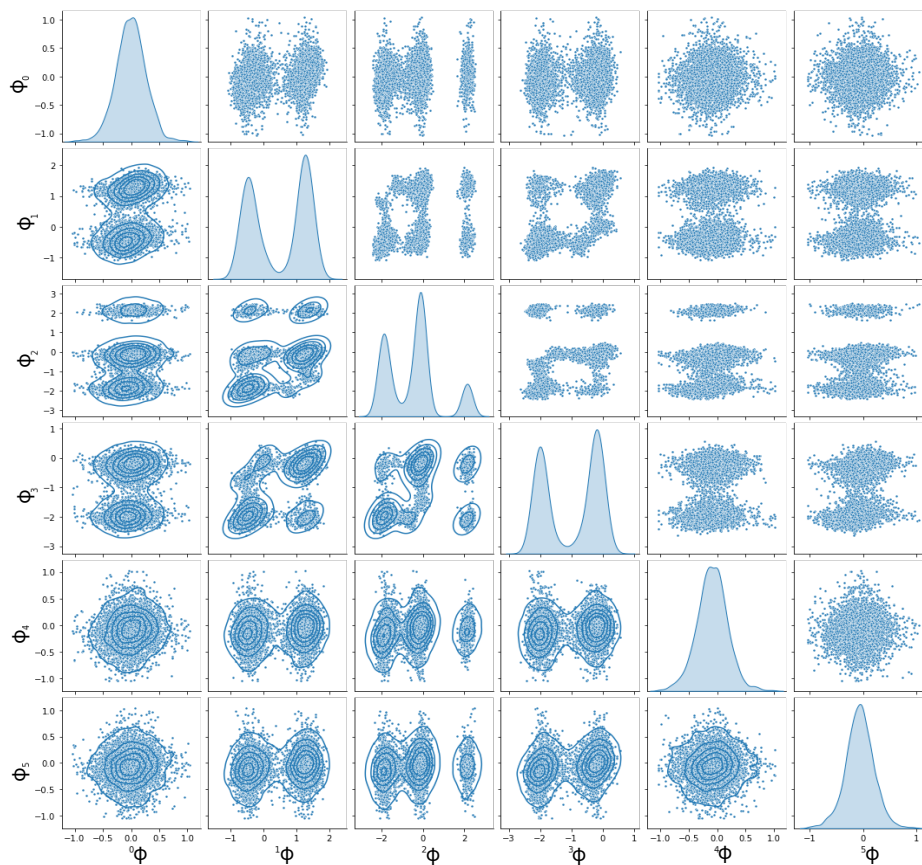


Figure B.2: Corner plot of internal rotation sampling for 3-ethylpentane. On the diagonal, marginal distributions estimated by Gaussian kernels are shown for each internal rotation. Scatter plots of samples are given on the upper and lower triangles. Contours of the densities are indicated on the lower triangle.

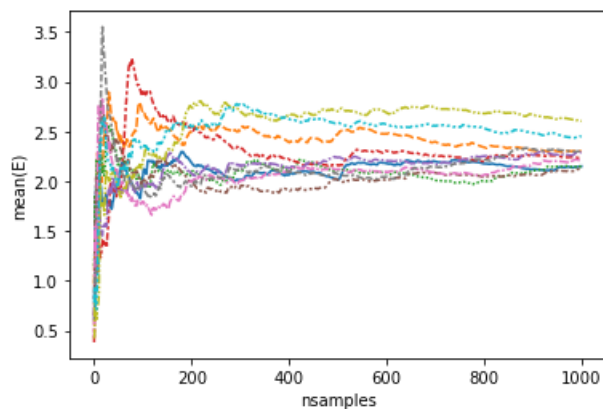


Figure B.3: Mean unpooled energy for internal rotations of isobutane (in kcal mol⁻¹) versus number of samples taken. Each line color represents an independently-sampled chain.

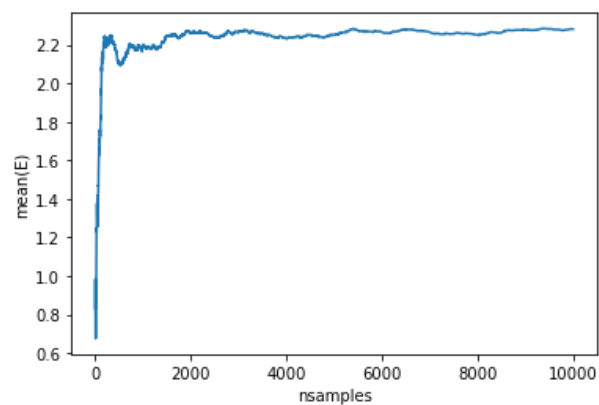


Figure B.4: Mean pooled energy for internal rotations of isobutane (in kcal mol⁻¹) versus number of samples taken. Samples were pooled maintaining parallelism across chains.

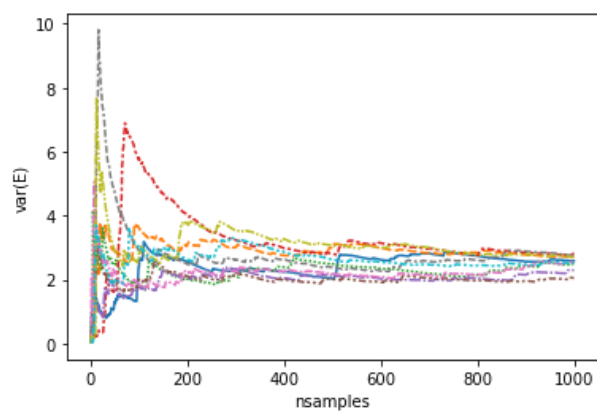


Figure B.5: Variance of the unpooled energy for internal rotations of isobutane (in kcal² mol⁻²) versus number of samples taken. Each line color represents an independently-sampled chain.

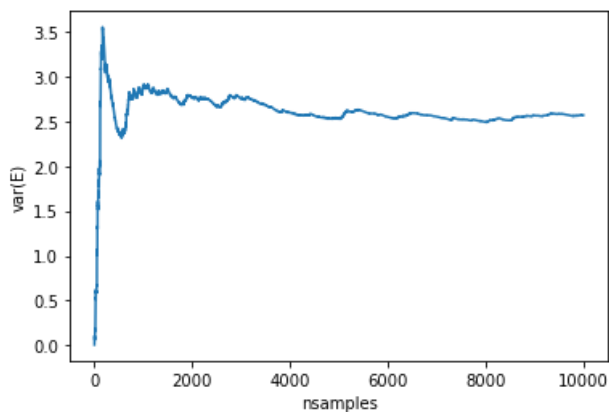


Figure B.6: Variance of the pooled energy for internal rotations of isobutane (in $\text{kcal}^2 \text{mol}^{-2}$) versus number of samples taken. Samples were pooled maintaining parallelism across chains.

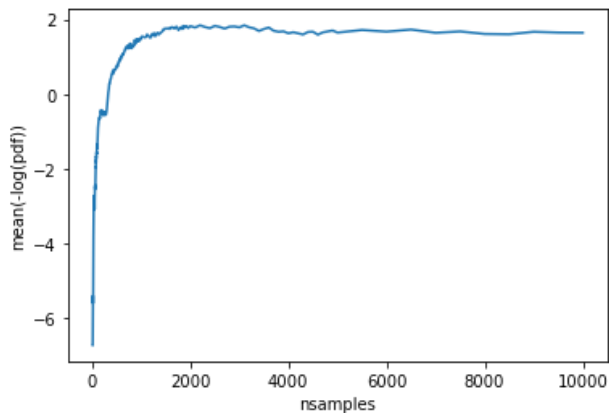


Figure B.7: Mean of the negative logarithm of the probability density function for internal rotations of isobutane. versus number of samples taken. This was obtained by computing the probability density function through a Gaussian kernel density estimator and resampling the resulting distribution. Samples were pooled maintaining parallelism across chains.

References

- [1] Matthew D Hoffman, Andrew Gelman, et al. “The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo.” *J. Mach. Learn. Res.* **15.1** (2014), 1593–1623

Appendix C

Supplementary Information for Chapter IV

1. Molecular Rotations

1.1. Euler Angle Addendum

As discussed in Section §2.2.2, Euler angles are a convenient parametrization to describe the orientation of a rigid body in space. This section clarifies the mathematics appertaining to the adopted ZYZ convention.

Rotations can be classified under two categories: intrinsic and extrinsic. Extrinsic rotations are defined with respect to the laboratory frame axes XYZ , whereas intrinsic rotations are defined with respect to the molecular frame axes xyz . Because laboratory frame axes are definitionally fixed, extrinsic rotations are perhaps the more intuitive. Extrinsicly, the target orientation (ϕ, θ, χ) is attained as follows

1. The molecule rotates by χ about the Z axis, s.t. the x axis is at an angle χ relative to the X axis (and the y axis is at an angle χ relative to the Y axis).
2. The newly-oriented molecule rotates by θ about the Y axis, s.t. the z axis is at an angle θ relative to the Z axis.
3. The molecule then rotates by ϕ , again about the Z axis.

This sequence is clearly depicted with the rotation operator defined in Equation 4.15, and directly leads to the Wigner D-Matrix definition. This operation can also be cast as the chained matrix multiplication

$$R(\phi, \theta, \chi) = R_Z(\phi)R_Y(\theta)R_Z(\chi) \tag{C.1}$$

where $R_Z(\phi)$ is the matrix for a rotation about the Z axis by ϕ , and so forth. Equivalently, the intrinsic rotation sequence to achieve the orientation (ϕ, θ, χ) is given by

1. The molecule rotates ϕ about the z axis.
2. The molecule rotates θ about the y' axis.

3. The molecule rotates χ about the z'' axis.

where y' and z'' are the molecular frame axes after the first and second rotations, respectively. The rotation matrix is again given by a chained matrix multiplication

$$R(\phi, \theta, \chi) = R_{z''}(\chi)R_{y'}(\theta)R_z(\phi) \quad (\text{C.2})$$

The equivalence of Equations C.1 and C.2 is not necessarily obvious, and is now shown. That is,

$$R_Z(\phi)R_Y(\theta)R_Z(\chi) = R_{z''}(\chi)R_{y'}(\theta)R_z(\phi) \quad (\text{C.3})$$

First, we note that z and Z are definitionally the same because the molecular axes prior to rotation are aligned with the laboratory axes, leading to

$$R_Z(\phi) = R_z(\phi) \quad (\text{C.4})$$

Next, let us consider the y' axis, which is the molecular y axis after rotation by ϕ about z . The same orientation for y' can be achieved by rotating the molecule back to the original xyz orientation, rotating the molecule by θ about the Y axis, and rotating the molecule back again to resume the $x'y'z'$ orientation. Mathematically:

$$R_{y'}(\theta) = R_z(\phi)R_Y(\theta)R_z^{-1}(\phi) \quad (\text{C.5})$$

And by the same logic, but with two more rotation operations:

$$R_{z''}(\phi) = R_{y'}(\theta)R_z(\phi)R_Z(\chi)R_z^{-1}(\phi)R_{y'}^{-1}(\theta) \quad (\text{C.6})$$

And therefore,

$$\begin{aligned} R_{z''}(\chi)R_{y'}(\theta)R_z(\phi) &= R_{y'}(\theta)R_z(\phi)R_Z(\chi)R_z^{-1}(\phi)\cancel{R_{y'}^{-1}(\theta)R_{y'}(\theta)}\overset{\mathbb{1}}{R_z(\phi)} \\ &= R_{y'}(\theta)R_z(\phi)R_Z(\chi)\cancel{R_z^{-1}(\phi)R_z(\phi)}\overset{\mathbb{1}}{} \\ &= R_z(\phi)R_Y(\theta)\cancel{R_z^{-1}(\phi)R_z(\phi)}\overset{\mathbb{1}}{R_Z(\chi)} \\ &= R_Z(\phi)R_Y(\theta)R_Z(\chi) \end{aligned}$$

which proves Equation C.3, as desired. Explicitly, the rotation matrix is given by the 3×3 matrix

$$R(\phi, \theta, \chi) = \begin{bmatrix} \cos \phi \cos \theta \cos \chi - \sin \phi \sin \theta & -\cos \chi \sin \phi - \cos \phi \cos \theta \sin \chi & \cos \phi \sin \theta \\ \cos \phi \sin \chi + \cos \theta \cos \chi \sin \phi & \cos \phi \cos \chi - \cos \theta \sin \phi \sin \chi & \sin \phi \sin \theta \\ -\cos \chi \sin \theta & \sin \theta \sin \chi & \cos \theta \end{bmatrix} \quad (\text{C.7})$$

1.2. Variational Method for Non-linear Rotors

The linear variational method is a procedure for finding the best possible approximate solutions to the eigenvalue problem

$$\hat{H}|\Phi\rangle = E|\Phi\rangle \quad (\text{C.8})$$

from a set of orthonormal basis functions. Following the derivation of Szabo and Ostlund [1], $|\Phi\rangle$ is approximated in the symmetric top basis as

$$|\Phi\rangle = \sum_{\ell=0}^{L_{\max}} \sum_{m=-\ell}^{\ell} \sum_{k=-\ell}^{\ell} c_{mk}^{\ell} |\ell mk\rangle \quad (\text{C.9})$$

where $\langle\phi, \theta, \chi|\ell mk\rangle \equiv \sqrt{\frac{2\ell+1}{8\pi^2}} D_{mk}^{\ell*}(\phi, \theta, \chi)$. Substituting the expansion into Equation C.8 and multiplying by $\langle\ell' m' k'|$, the matrix representation of \hat{H} becomes apparent.

$$\sum_{\ell=0}^{L_{\max}} \sum_{m=-\ell}^{\ell} \sum_{k=-\ell}^{\ell} c_{mk}^{\ell} \langle\ell' m' k'|\hat{H}|\ell mk\rangle = E \sum_{\ell=0}^{L_{\max}} \sum_{m=-\ell}^{\ell} \sum_{k=-\ell}^{\ell} c_{mk}^{\ell} \langle\ell' m' k'|\ell mk\rangle = E c_{m'k'}^{\ell'} \quad (\text{C.10})$$

$$\mathbf{Hc} = E\mathbf{c} \quad (\text{C.11})$$

Kinetic Energy Matrix Elements

The following paragraphs are a derivation of the kinetic energy matrix elements. We use a rotational constant convention after Kennerly [2], with

$$a = \frac{1}{2I_y}, \quad b = \frac{1}{2I_z}, \quad c = \frac{1}{2I_x}$$

The kinetic energy operator under this convention is given by

$$\begin{aligned} \hat{T} &= a\hat{L}_y^2 + b\hat{L}_z^2 + c\hat{L}_x^2 \\ &= \frac{1}{2}(a+c)\hat{L}^2 + \frac{1}{2}(a-c)\left(\hat{L}_y^2 + \kappa\hat{L}_z^2 - \hat{L}_x^2\right) \end{aligned} \quad (\text{C.12})$$

where an asymmetry parameter is defined as $\kappa = \frac{2b-(a+c)}{a-c}$. Each component will be solved individually below, and pieced together to give the final result.

First, we state the known eigenvalues of the operators \hat{L}^2 , \hat{L}_Z , and \hat{L}_z , drawing attention to the difference between operators of the molecular frame and the laboratory frame. Specifically, \hat{L}_z manifests in observations of k (multiples of \hbar), and \hat{L}_Z in observations of m .

$$\begin{aligned} \hat{L}^2 |\ell mk\rangle &= \hbar^2 \ell(\ell+1) |\ell mk\rangle \\ \hat{L}_Z |\ell mk\rangle &= \hbar m |\ell mk\rangle \\ \hat{L}_z |\ell mk\rangle &= \hbar k |\ell mk\rangle \end{aligned}$$

Intuitively, only molecular frame operations appear in the kinetic energy operator, and the quantum number m will therefore not appear in the description of the kinetic energy. The relevant contributions for total angular momentum and z angular momentum are thus

$$\langle\ell' m' k'|\hat{L}^2|\ell mk\rangle = \hbar^2 \ell(\ell+1) \delta_{\ell'\ell} \delta_{m'm} \delta_{k'k} \quad (\text{C.13})$$

$$\langle\ell' m' k'|\hat{L}_z^2|\ell mk\rangle = \hbar^2 k^2 \delta_{\ell'\ell} \delta_{m'm} \delta_{k'k} \quad (\text{C.14})$$

For \hat{L}_x and \hat{L}_y contributions, we define angular momentum creation and annihilation operators $\hat{L}^\pm = \hat{L}_x \pm i\hat{L}_y$, which act on the Wigner D-Matrix kets as [3]

$$\hat{L}^\pm |l m k\rangle = \hbar \sqrt{\ell(\ell+1) - k(k \mp 1)} |l m k \mp 1\rangle \quad (\text{C.15})$$

(Note that \hat{L}^+ is the *annihilation* operator and \hat{L}^- is the *creation* operator in the molecular frame, which lower and raise the quantum number k only). With $\hat{L}_x = \frac{1}{2}(\hat{L}^+ + \hat{L}^-)$ and $\hat{L}_y = \frac{1}{2i}(\hat{L}^+ - \hat{L}^-)$,

$$\hat{L}_x^2 = \frac{1}{4} \left(\hat{L}^+ \hat{L}^+ + \hat{L}^- \hat{L}^- + \hat{L}^+ \hat{L}^- + \hat{L}^- \hat{L}^+ \right) \quad (\text{C.16})$$

$$\hat{L}_y^2 = -\frac{1}{4} \left(\hat{L}^+ \hat{L}^+ + \hat{L}^- \hat{L}^- - \hat{L}^+ \hat{L}^- - \hat{L}^- \hat{L}^+ \right), \quad (\text{C.17})$$

and the selection rules $\Delta k = 0, \pm 2$ follow. By Equations C.15-C.17, we obtain the \hat{L}_x^2 and \hat{L}_y^2 matrix elements after some simple algebra:

$$\begin{aligned} \langle \ell' m' k' | \hat{L}_x^2 | \ell m k \rangle &= \frac{\hbar^2}{2} [\ell(\ell+1) - k^2] \delta_{\ell'\ell} \delta_{m'm} \delta_{k'k} \\ &+ \frac{\hbar^2}{4} \sqrt{\ell(\ell+1) - k(k \pm 1)} \sqrt{\ell(\ell+1) - (k \pm 1)(k \pm 2)} \delta_{\ell'\ell} \delta_{m'm} \delta_{k'k \pm 2} \end{aligned} \quad (\text{C.18})$$

$$\begin{aligned} \langle \ell' m' k' | \hat{L}_y^2 | \ell m k \rangle &= \frac{\hbar^2}{2} [\ell(\ell+1) - k^2] \delta_{\ell'\ell} \delta_{m'm} \delta_{k'k} \\ &- \frac{\hbar^2}{4} \sqrt{\ell(\ell+1) - k(k \pm 1)} \sqrt{\ell(\ell+1) - (k \pm 1)(k \pm 2)} \delta_{\ell'\ell} \delta_{m'm} \delta_{k'k \pm 2} \end{aligned} \quad (\text{C.19})$$

Combining Equations C.12-C.14 and C.18-C.19, we arrive at the following result for the kinetic energy matrix elements, again after some light algebra.

$$\begin{aligned} \langle \ell' m' k' | \hat{T} | \ell m k \rangle &= \frac{1}{2} [(A+C)\ell(\ell+1) + (A-C)\kappa k^2] \delta_{\ell'\ell} \delta_{m'm} \delta_{k'k} \\ &+ \frac{1}{4} (C-A) \sqrt{\ell(\ell+1) - k(k \pm 1)} \sqrt{\ell(\ell+1) - (k \pm 1)(k \pm 2)} \delta_{\ell'\ell} \delta_{m'm} \delta_{k'k \pm 2} \end{aligned} \quad (\text{C.20})$$

The products of the \hbar^2 factors and constants a and c lead to the canonical rotational constants A and C , convention preserved. Note the dependence of the kinetic energy on I_z is contained within b , which is contained within κ . The rotational constant B therefore does not explicitly appear in Equation C.20.

Potential Energy Matrix Elements

Matrix elements of the potential energy component of the Hamiltonian matrix are derived here. As discussed in the main text, sampled potential energy surfaces can be approximated as continuous functions of Euler angles ϕ , θ , and χ by expansion into a finite basis of Wigner

D-Matrix elements (see Equations 4.18-4.20). Including this result in the potential energy matrix in the same basis yields

$$\begin{aligned}\langle \ell' m' k' | \hat{V} | \ell m k \rangle &= \langle \ell' m' k' | \sum_{\ell''} \sum_{m''} \sum_{k''} \hat{v}_{m'' k''}^{\ell''} D_{m'' k''}^{\ell''}(\phi, \theta, \chi) | \ell m k \rangle \\ &= \frac{\sqrt{2\ell' + 1} \sqrt{2\ell + 1}}{8\pi^2} \sum_{\ell''} \sum_{m''} \sum_{k''} \hat{v}_{m'' k''}^{\ell''} \int_{\Omega} d\Omega D_{m' k'}^{\ell'}(\Omega) D_{m'' k''}^{\ell''}(\Omega) D_{m k}^{\ell*}(\Omega)\end{aligned}\quad (\text{C.21})$$

where Ω specifies the orientation of the molecule in Euler angles, and

$$\int_{\Omega} d\Omega \equiv \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin \theta \int_0^{2\pi} d\chi \quad (\text{C.22})$$

The Kronecker product $\mathbf{D}^{\ell'} \otimes \mathbf{D}^{\ell''}$ can be reduced by the Clebsch-Gordan series [4]:

$$D_{m' k'}^{\ell'}(\Omega) D_{m'' k''}^{\ell''}(\Omega) = \sum_{L=|\ell' - \ell''|}^{\ell' + \ell''} \langle \ell' m' \ell'' m'' | LM \rangle \langle \ell' k' \ell'' k'' | LK \rangle D_{MK}^L(\Omega) \quad (\text{C.23})$$

where $M \equiv m' + m''$ and $K \equiv k' + k''$. The quantity $\langle \ell' m' \ell'' m'' | LM \rangle$ and its k counterpart are the well-known Clebsch-Gordan coefficients. By orthogonality (Equation 4.18) and Equation C.23, the triple integral over the three Wigner D-Matrices of Equation C.21 is reduced to

$$\int_{\Omega} d\Omega D_{m' k'}^{\ell'}(\Omega) D_{m'' k''}^{\ell''}(\Omega) D_{m k}^{\ell*}(\Omega) = \frac{8\pi^2}{2\ell + 1} \langle \ell' m' \ell'' m'' | \ell m \rangle \langle \ell' k' \ell'' k'' | \ell k \rangle \quad (\text{C.24})$$

Equation C.21 simplifies to the final result:

$$\langle \ell' m' k' | \hat{V} | \ell m k \rangle = \sqrt{\frac{2\ell' + 1}{2\ell + 1}} \sum_{\ell''} \sum_{m''} \sum_{k''} \hat{v}_{m'' k''}^{\ell''} \langle \ell' m' \ell'' m'' | \ell m \rangle \langle \ell' k' \ell'' k'' | \ell k \rangle \quad (\text{C.25})$$

1.3. Thermodynamic Derivations

Thermodynamic quantities are statistical quantities derived from the partition function Q . In all cases, the Helmholtz Free Energy is computed by

$$A = -\frac{1}{\beta} \log Q \quad (\text{C.26})$$

and the entropy is computed by the relationship

$$S = \frac{E - A}{T} \quad (\text{C.27})$$

The partition functions and energetic contributions derived therefrom are given by the following relationships, stemming from

$$E = -\frac{\partial \log Q}{\partial \beta} \quad (\text{C.28})$$

Partition functions are included in the main text, but are included in subsequent equations (often in different algebraic forms) for additional clarity. Superscripts and subscripts are retained for disambiguation across methods.

Quantum Harmonic Oscillator

$$E_0 = \frac{1}{2}\hbar\omega \quad (\text{C.29})$$

$$Q_{\text{HO}}^{qu} = \frac{1}{2 \sinh(\beta E_0)} \quad (\text{C.30})$$

$$E_{\text{HO}}^{qu} = \frac{E_0}{\tanh(\beta E_0)} \quad (\text{C.31})$$

Classical Harmonic Oscillator

$$Q_{\text{HO}}^{cl} = (\beta\hbar\omega)^{-1} \quad (\text{C.32})$$

$$E_{\text{HO}}^{cl} = \beta^{-1} \quad (\text{C.33})$$

Classical Rigid Rotor

$$Q_{\text{RR}}^{cl} = \frac{1}{\sigma} \sqrt{\frac{I_x I_y I_z}{(2\pi\beta\hbar)^3}} \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\chi e^{-\beta V(\phi, \theta, \chi)} \quad (\text{C.34})$$

Linearization of the partition function by log and subsequent derivation with respect to β gives

$$E_{\text{RR}}^{cl} = \frac{3}{2\beta} + \langle V \rangle \quad (\text{C.35})$$

where the average potential energy is given by the integral

$$\langle V \rangle = \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\chi V(\phi, \theta, \chi) f(\phi, \theta, \chi) \quad (\text{C.36})$$

where f is the probability density function in the canonical ensemble, given by

$$f(\phi, \theta, \chi) = \frac{e^{-\beta V(\phi, \theta, \chi)}}{\int_0^{2\pi} d\phi \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\chi e^{-\beta V(\phi, \theta, \chi)}} \quad (\text{C.37})$$

Quantum Rigid Rotor

The thermodynamics of the quantum rigid rotor are solved after diagonalization of the Hamiltonian matrix to obtain a set of eigenvalues E_j . The partition function is obtained by direct summation of the Boltzmann-distributed eigenvalues, with symmetry corrected:

$$Q_{\text{RR}}^{qu} = \frac{1}{\sigma} \sum_j e^{-\beta E_j} \quad (\text{C.38})$$

Degeneracies are implicitly included in this expression, stipulating computed eigenvalues can be identical. The average energy under a discrete probability distribution is therefore

$$E_{\text{RR}}^{qu} = \frac{\sum_j E_j e^{-\beta E_j}}{\sum_j e^{-\beta E_j}}. \quad (\text{C.39})$$

Zero point energies under this method are given by lowest eigenvalue, E_0 .

Pitzer-Gwinn Approximations

The thermodynamics of the PG approximation are straightforward, since the product of the partition function linearizes under logarithmic operation. As such,

$$Q_{\text{RR}}^{\text{PG}} = \frac{Q_{\text{HO}}^{\text{qu}}}{Q_{\text{HO}}^{\text{cl}}} Q_{\text{RR}}^{\text{cl}} \quad (\text{C.40})$$

directly leads to

$$E_{\text{RR}}^{\text{PG}} = E_{\text{HO}}^{\text{qu}} - E_{\text{HO}}^{\text{cl}} + E_{\text{RR}}^{\text{cl}}$$

and so forth for A , S , etc. Zero point energy contributions are implicitly contained within the quantum HO reference partition function, and therefore the zero point energy is that of the quantum HO.

2. Sampling Grid Size Effects

Table C.1: Effect of Lebedev Grid Size on Predicted Rotational Thermodynamics of Methane in H-CHA

Grid size	Sample size	Resolution [deg.]	ΔE_{rot}^a [kJ mol ⁻¹]	ΔS_{rot}^a [J mol ⁻¹ K ⁻¹]
50	550	28.7	6.87	-7.74
74	1036	23.6	6.87	-7.43
110	1870	19.3	6.75	-7.61

^aStandard rotational contributions derived from the classical partition function at $T = 298$ K.

3. Model Parameters

Table C.2: Parameters for Relevant Molecules

Molecule	r_k [Å]
Water	1.40 ^a
Methane	1.90 ^b
Ethane	2.22 ^c
Propane	2.15 ^b
Methanol	1.80 ^d
Ethanol	2.15 ^e

^aRef [5]; ^bRef [6, 7]; ^cRef [8, 9]; ^dRef [10, 11]; ^eRef [12, 13].

Table C.3: Zeolite Parameters

Framework	Cell Volume [\AA^3]	% Occupiable
H-CHA	2391.6 ^a	17.27 ^b
H-MFI	5211.3 ^a	9.81 ^b

^aRef [14]; ^bRef [5].

References

- [1] Attila Szabo and Neil S. Ostlund. *Modern Quantum Chemistry : Introduction to Advanced Electronic Structure Theory*. Mineola, N.Y: Dover Publications, 1996. ISBN: 0486691861
- [2] William W Kennerly. *Molecules Rotating in Electric Fields by Quantum and Semi-Quantum Mechanics*. Cornell University, 2005
- [3] H. W. Kroto. *Molecular Rotation Spectra*. London, New York: Wiley, 1975. ISBN: 9780486495408
- [4] Morris Rose. *Elementary Theory of Angular Momentum*. New York: Dover, 1995. ISBN: 9780486684802
- [5] M.D. Foster et al. “A Geometric Solution to the Largest-Free-Sphere Problem in Zeolite Frameworks”. *Microporous Mesoporous Mater.* **90.1** (2006). Dedicated to the late Denise Barthomeuf, George Kokotailo and Sergey P. Zhdanov in appreciation of their outstanding contributions to zeolite science, 32–38. ISSN: 1387-1811. DOI: <https://doi.org/10.1016/j.micromeso.2005.08.025>. URL: <https://www.sciencedirect.com/science/article/pii/S1387181105003914>
- [6] Chen Zhang et al. “Unexpected Molecular Sieving Properties of Zeolitic Imidazolate Framework-8”. *J. Phys. Chem. Lett.* **3.16** (2012), 2130–2134
- [7] Donald Breck. *Zeolite Molecular Sieves: Structure, Chemistry, and Use*. New York: Wiley, 1973. ISBN: 9780471099857
- [8] Sonia Aguado et al. “Absolute Molecular Sieve Separation of Ethylene/Ethane Mixtures with Silver Zeolite A”. *J. Am. Chem. Soc.* **134.36** (2012), 14635–14637
- [9] Siddharth Gautam, Tingting Liu, and David Cole. “Sorption, Structure and Dynamics of CO₂ and Ethane in Silicalite at High Pressure: A Combined Monte Carlo and Molecular Dynamics Simulation Study”. *Molecules* **24.1** (2019), 99
- [10] E Johan et al. “Transport Mechanisms of Water and Organic Solvents through Microporous Silica in the Pervaporation of Binary Liquids”. *Microporous Mesoporous Mater.* **65.2-3** (2003), 197–208
- [11] Han Hu et al. “A Robust etb-type Metal–organic Framework Showing Polarity-exclusive Adsorption of Acetone over Methanol for their Azeotropic Mixture”. *Chem. Commun.* **55.46** (2019), 6495–6498
- [12] A Nalaparaju, XS Zhao, and a JW Jiang. “Molecular Understanding for the Adsorption of Water and Alcohols in Hydrophilic and Hydrophobic Zeolitic Metal–organic Frameworks”. *J. Phys. Chem. C* **114.26** (2010), 11542–11550

- [13] Yiwen Tang, David Dubbeldam, and Stefania Tanase. “Water–Ethanol and Methanol–Ethanol Separations Using in Situ Confined Polymer Chains in a Metal–Organic Framework”. *ACS applied materials & interfaces* **11.44** (2019), 41383–41393
- [14] Ch. Baerlocher and L.B. McCusker. *Database of Zeolite Structures*. URL: <http://www.iza-structure.org/databases/>