

Automated Reaction Kinetics and Network Exploration (Arkane): A Statistical Mechanics, Thermodynamics, Transition State Theory, and Master Equation Software

Alon Grinberg Dana,^{†,‡} Matthew S. Johnson,[†] Joshua W. Allen,[†] Sandeep
Sharma,[†] Sumathy Raman,[†] Mengjie Liu,[†] Connie W. Gao,[†] Colin A. Grambow,[†]
Mark J. Goldman,[†] Duminda S. Ranasinghe,[†] Ryan J. Gillis,[†] A. Mark Payne,[†]
Yi-Pei Li,[¶] Enoch E. Dames,[†] Zachary J. Buras,[†] Nick M. Vandewiele,[†] Nathan
W. Yee,[†] Shamel S. Merchant,[†] Beat Buesser,[†] Caleb A. Class,[†] C. Franklin
Goldsmith,[§] Richard H. West,^{||} and William H. Green^{*,†}

[†]*Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge,
MA 02139, United States*

[‡]*The Wolfson Department of Chemical Engineering, Grand Technion Energy Program
(GTEP), Technion – Israel Institute of Technology, Haifa 3200003, Israel*

[¶]*Department of Chemical Engineering, National Taiwan University, Taipei 10617, Taiwan*

[§]*School of Engineering, Brown University, Providence, RI 02912, United States*

^{||}*Department of Chemical Engineering, Northeastern University, Boston, MA 02115,
United States*

E-mail: whgreen@mit.edu

Phone: +1-617-253-4580

Abstract

The open-source statistical mechanics software described here, Arkane – Automated Reaction Kinetics and Network Exploration, facilitates computations of thermodynamic properties of chemical species, high-pressure limit reaction rate coefficients, and pressure-dependent rate coefficient over multi-well molecular potential energy surfaces (PES) including the effects of collisional energy transfer on phenomenological kinetics. Arkane is especially flexible, computing a variety of quantities using different methods, as determined by the users. For example, it can use estimates to fill in information for molecules or reactions where quantum chemistry information is missing. The software solves the internal energy master equation for complex unimolecular reactions systems. Inputs to the software include converged electronic structure computations performed by the user using a variety of supported software packages (Gaussian, Molpro, Orca, TeraChem, Q-Chem, Psi4). The software outputs high-pressure limit rate coefficients and pressure-dependent phenomenological rate coefficients, as well as computed thermodynamic properties (enthalpy, entropy, and constant pressure heat capacity) with added energy corrections. Some of the key features of Arkane include treatment of 1D, 2D or ND hindered internal rotation modes, treatment of free internal rotation modes, quantum tunneling effect consideration (implementing the Eckart or Wigner approach), transition state theory (TST) and Rice-Ramsperger-Kassel-Marcus (RRKM) rate coefficient computations, master equation solution with three implemented methods, inverse-Laplace transform of high-pressure limit rate coefficients into the energy domain, energy corrections based on bond-additivity or isodesmic reactions, automated and efficient PES exploration, and PES sensitivity analysis. The present work describes the design of Arkane, how it should be used, and refers to the theory that it employs. Arkane is distributed via the RMG-Py software suite (<https://github.com/ReactionMechanismGenerator/RMG-Py>).

1. Introduction

Over the last few decades, rapid advances in computing power and high-accuracy quantum chemistry methods have led to the widespread use of *ab initio* calculations for investigating chemical mechanisms. Such chemical kinetic mechanisms are of interest in a wide range of fields, including combustion, atmospheric chemistry, and astrochemistry. Quantum chemistry calculations provide a straightforward way to directly interrogate specific molecules and reaction pathways, which is especially valuable in systems which may be too complex to study experimentally.

In constructing these mechanisms, species thermochemistry and reaction kinetics are the primary quantities of interest. To obtain them, one can perform electronic structure calculations to determine geometric conformations, energies, and vibrational frequencies for a given set of species and transition states, then use established methods to convert that information to macroscopic thermochemistry values and rate coefficients. The partition function is used to compute thermochemistry properties, and partition functions of reactive species along with the transition state are foundational for rate coefficient calculations. High-pressure-limit rate coefficients can be calculated using conventional transition state theory (TST).²³ In addition, chemical activation may be important in some systems, which requires calculation of pressure-dependent rate coefficients using Rice-Ramsperger-Kassel-Marcus (RRKM) theory²⁴⁻²⁶ and master equation (ME) solvers. Finally, quantum tunneling effects must also be taken into consideration, via methods such as Eckart²⁷ or Wigner²⁸ corrections, or more sophisticated approaches like zero-curvature, small-curvature, or large-curvature tunneling (ZVT, SVT, LVT, respectively).²⁹

Our group also developed software to implement some of the rate and thermochemistry calculations mentioned above as part of the development of the Reaction Mechanism Generator (RMG) software package.^{30,31} Portions of this capability were historically exposed to users as the CanTherm and MEASURE modules in RMG, but they were specialized, hard to use, and never well-documented in the archival literature. Here we present the more general-

Table 1: Technical aspect comparison of software comparable with Arkane

| Name | Language | License ^a | Platforms ^b | GUI ^c | Supported ESS ^d | Citation |
|-----------|---------------------|----------------------|------------------------|------------------|---|------------------|
| Arkane | Python ^e | MIT | L/M | – | Gaussian, Molpro, Orca, TeraChem, Q-Chem, Psi4 | The present work |
| APUAMA | C++ | – | L/W | ✓ | – | Ref. 1 |
| Chemdis | Fortran | – | L/M/W | – | – | Refs. 2,3 |
| ChemRate | C++ | – | W | ✓ | Gaussian | Ref. 4 |
| CHIMERA | C++ | – | L/M/W | ✓ | GAMESS, Gaussian | Ref. 5 |
| Eyringpy | Python | – | L/M/W | – | Gaussian | Ref. 6 |
| Genesys | Java | proprietary software | L/M/W | – | Gaussian | Ref. 7 |
| KiSTheIP | Java | Custom ⁸ | L/M/W | ✓ | GAMESS, Gaussian, Molpro, NWChem, Orca | Ref. 9 |
| MESMER | C++ | GPLv2 | L/M/W | – | – | Ref. 10 |
| MESS | Fortran | – | L | – | – | Ref. 11 |
| MultiWell | Fortran | GPLv3 | L/M/W | – | Gaussian | Ref. 12 |
| PAC | Fortran | – | L | – | – | Ref. 13 |
| pMuTT | Python | MIT | L/M/W | – | Gaussian | Ref. 14 |
| Polyrate | Fortran | Custom ¹⁵ | L/M/W | – | CHARMM, GAMESS, Gaussian, Jaguar, MC-TINKER, MORATE, MULTILEVEL, NWChem | Ref. 16 |
| SSUMES | C++ | GPLv3 | L/W | – | Gaussian | Ref. 17 |
| TAMkin | Python | GPLv3 | L | – | ADF, CHARMM, CPMD, CP2K, Gaussian, Q-Chem, VASP | Ref. 18 |
| TheRate | Java | – | L/M/W | ✓ | Gaussian | Ref. 19 |
| TUMME | Python, and C++ | Apache-2.0 | L | ✓ | Gaussian | Ref. 20 |
| Unimol | Fortran | – | L/M/W | – | – | Ref. 21 |
| VariFlex | Fortran | – | L/M/W | – | – | Ref. 22 |

^a No specified licence means that the software is available from the respective authors upon request, usually with no fee.

^b L - Linux, M - MacOS, W - Windows

^c Available graphical user interface

^d Support for parsing of Electronic Structure Software (ESS) output files. If none is specified, input data must be entered in a pre-defined format and is not automatically read from electronic structure software output files.

^e Also uses Cython to convert performance critical code to C

purpose Automated Reaction Kinetics and Network Exploration (Arkane) software package, freely available under the open-source MIT license,³² which incorporates the capabilities of both CanTherm and MEASURE, along with several new features, as discussed in detail below. In brief, Arkane accepts the output of various quantum chemistry software packages (it currently supports Gaussian,³³ Molpro,^{34,35} Orca,³⁶ TeraChem,,^{37,38} Q-Chem,³⁹ and Psi4⁴⁰) and computes thermochemistry, high-pressure-limit rate coefficients, and pressure-dependent rate coefficients. For unimolecular reaction networks, it can use either RRKM or the inverse Laplace transform (ILT) method^{41,42} to calculate microcanonical rate coefficients, $k(E)$, and then calculate the phenomenological rates, $k(T, P)$, by solving the ME using several different approaches/approximations: modified strong collision (MSC),² reservoir state (RS),⁴³ chemically significant eigenvalues (CSE),^{44,45} and simulation least-squares (SLS).⁴⁶ Arkane was designed to be especially flexible, computing a variety of quantities using different methods, as determined by the users. Arkane can use estimates to fill in information for molecules or reactions where quantum chemistry information is missing. Finally, it is an open-source project, and anyone can extend and improve it.

Arkane can also run sensitivity analysis of the calculated pressure-dependent rate coefficients to the species and transition state energies in the unimolecular reaction network. Additionally, it can take advantage of the RMG-Py functionality and database to automatically explore unimolecular reaction networks using estimated energies and barriers.

Arkane and its predecessors have been successfully used to calculate thermochemistry and kinetics for mechanistic investigations of a wide range of chemistries, including fuel combustion and pyrolysis,^{48–54} sulfur chemistry,^{55–59} silicon chemistry,⁶⁰ nitrogen chemistry,^{61,62} alkylaromatic pyrolysis,^{63–66} and polycyclic aromatic hydrocarbon formation.^{67–71}

Aside from Arkane, there are numerous other software programs which support similar capabilities. Comparisons of a number of technical aspects and capabilities of such software are presented in Tables 1 and 2, respectively. Some of these programs can parse output files of third party electronic structure software (ESS), while others only allow manual parameter

Table 2: Theories and methods comparison of software related to Arkane

| Name | Thermo ^a | Reaction Rate Theories ^b | Master Equation Methods | Tunneling Methods ^c | Internal rotation Support ^d |
|--|---------------------|---|-------------------------|--------------------------------|--|
| Software packages that support multiwell reactions: | | | | | |
| Arkane | ✓ | TST, ILT, RRKM | CSE, MSC, RS, SLS | Eckart, Wigner | 1DHR, 2DHR, NDHR, Free rotors |
| Chemdis | – | RRKM | MSC | – | 1DHR |
| ChemRate | ✓ | RRKM | CSE | Eckart | 1DHR |
| CHIMERA | ✓ | TST, RRKM | MSC | – | Free rotors |
| KiSTheP | ✓ | TST, VTST, RRKM | – | Eckart, Wigner | HRDS |
| MESMER | – | ILT, RRKM | RS, CSE | Eckart | 1DHR |
| MESS | – | VTST, RRKM | CSE | Parabolic, Eckart, Quartic | 1DHR, Free rotors |
| MultiWell | ✓ | ILT, RRKM | Stochastic | Eckart, semi-classical | 1DHR, 2DHR, Free rotors |
| Polyrate | – | TST, VTST, QRRK | MSC | ZCT, SCT, LCT, OMT | MS |
| SSUMES | – | TST, VTST (using Unimol ²¹), RRKM | BRW ^e | Eckart, Wigner | 1DHR |
| TUMME | – | TST, VTST, RRKM | CSE | SCT | MS |
| Unimol | ✓ | RRKM | BRW ^e | – | – |
| Software packages that do not support multiwell reactions: | | | | | |
| APUAMA | ✓ | TST | – | Eckart, Wigner, SCT | – |
| Eyringpy | ✓ | TST | – | Eckart, Wigner | – |
| Genesys | ✓ | TST | – | Eckart, Wigner | 1DHR |
| PAC | ✓ | – | – | – | 1DHR, Free rotors |
| pMuTT | ✓ | TST | – | – | quasi-RRHO |
| TAMkin | ✓ | TST | – | Eckart, Wigner | 1DHR |
| TheRate | – | TST, CV-TST | – | Eckart, Wigner, ZCT, SCT | – |
| VariFlex | – | VRC-TST, PST, ILT | – | Eckart | 1DHR |

^a Capable of calculating thermochemical properties

^b TST – transition-state theory, ILT - inverse Laplace transform, RRKM – the Rice-Ramsperger-Kassel-Marcus theory, VTST – variational transition-state theory, QRRK – quantum Rice-Ramsperger-Kassel theory, CV-TST – canonical variational transition-state theory, VRC-TST – variational reaction coordinate transition-state theory, PST – phase-space theory.

^c ZCT – zero-curvature tunneling, SCT – small-curvature tunneling, LCT – large-curvature tunneling, OMT – optimized multidimensional tunneling

^d 1DHR – 1D hindered rotor, 2DHR – 2D hindered rotor, NDHR – ND hindered rotor, HRDS – hindered rotor density-of-states, MS – Multi-structure

^e BRW – biased-random-walk⁴⁷

input.

2. Arkane Design

The primary purpose of Arkane is to provide an open-source code that facilitates calculations of species thermochemical properties and high-pressure-limit as well as pressure-dependent reaction rate coefficients. Arkane is written in Python, which provides a stable, robust programming architecture for developing an extensible and modular codebase with a large suite of unit tests. Selected performance-critical modules of the code were written using Cython,^{72,73} which compiles Python code to C code through static typing, leading to speed enhancement for numerically-intensive tasks.

Large-scale detailed PES modeling involves many isomer wells and inter-connecting reaction paths. The required parameters for all of these wells and paths are often unavailable. Since Arkane is intimately connected with RMG-Py and its database, it is able to use estimation methods in RMG-Py to determine missing well energies and rate coefficient parameters. Combining these capabilities with a pressure-dependent sensitivity analysis feature, described below, provides users with a flexible, convenient tool for computing the pressure-dependent rate coefficients for systems involving PES's with many isomers with only moderate effort.

Here, we give general guidelines of how to compose Arkane input files and what to expect as an output. Users are strongly encouraged to visit Arkane's online documentation,⁷⁴ where an up-to-date and more detailed version of these guidelines can be found. At its core, Arkane takes one or more input files detailing the system being studied along with log files of the relevant quantum chemical computations, and produces output files containing the relevant computed thermodynamic and kinetic information. There are essentially three types of information that Arkane produces: species thermochemistry, high-pressure-limit reaction rate coefficients, and pressure-dependent reaction rate coefficients.

Table 3: Sections of an Arkane input file

| Information Type | Keyword | Format | Required |
|---|------------------------------------|------------|----------------------|
| Level of theory | <code>modelChemistry</code> | String | Yes |
| Author | <code>author</code> | String | No |
| Frequency scale factor | <code>frequencyScaleFactor</code> | Float | No ^a |
| Use atom corrections | <code>useAtomCorrections</code> | Boolean | No (default: True) |
| Atomic energies | <code>atomEnergies</code> | Dictionary | No ^a |
| Use bond additivity corrections | <code>useBondCorrections</code> | Boolean | No (default: True) |
| Bond additivity correction type ^b | <code>bondCorrectionType</code> | String | No (default: 'p') |
| Use isodesmic reaction corrections ^c | <code>useIsodesmicReactions</code> | Boolean | No (default: False) |
| Isodesmic reaction corrections reference set ^c | <code>referenceSets</code> | String | No (default: 'main') |
| Use Hindered Rotor | <code>useHinderedRotors</code> | Boolean | No (default: True) |
| Database ^d | <code>database</code> | – | No |
| Species ^d | <code>species</code> | – | Yes ^e |
| Transition state ^d | <code>transitionState</code> | – | No |
| Reaction ^d | <code>reaction</code> | – | No |
| P-dependent rate coefficient computation ^d | <code>network</code> | – | No |
| Thermodynamic computation instructions ^d | <code>thermo</code> | – | No |
| High-P limit rate coefficient computation ^d | <code>kinetics</code> | – | No |
| Parameters for solving the ME ^d | <code>pressureDependence</code> | – | No |
| Explore a PES ^d | <code>explorer</code> | – | No |
| Generate bond additivity corrections ^d | <code>bac</code> | – | No |

^a This information is not required if it is already available in Arkane for the specific levels of theory, see the documentation⁷⁴ for an availability list.

^b Arkane supports both Petersson and Melius bond additivity types. Pass either 'p' or 'm' to `bondCorrectionType`, respectively, to specify which type will be used.

^c Using Isodesmic Reaction corrections will also force the usage of Atom Corrections and will set the Use Bond Corrections argument to False. This ensures consistency with the non-target species of the isodesmic reaction.

^d Additional information is given in the Sections below.

^e Specifying a structure in addition to the species file path is required if parameter auto-completion is desired.

Inputs

Arkane can be executed either using its Python API (application programming interface), or by specifying a path to an Arkane input file through a command-line interface. Arkane input files are comprised of a series of segments, each relates to a component of the studied system (e.g., species or reaction) and additional computational instructions. The format follows Python syntax. Note that a distinct definition is required for each stationary point relevant for an Arkane computation (i.e., all species and transition states, as applicable). The Arkane input file would either include the species molecular properties or reference individual Arkane species files, as further discussed below. Code Excerpt 1 gives an example for an Arkane input file with instructions to compute thermodynamic properties of ethylene and a high-pressure-limit reaction rate coefficient for the reaction $\text{C}_2\text{H}_4 + \text{H}\cdot \rightleftharpoons \text{C}_2\text{H}_5\cdot$. Note that separate Arkane species input files referencing the relevant quantum calculation log files are required, as further explained below. Table 3 describes possible segments in an Arkane input file, many of which are described in more detail in the Sections below.

Code Excerpt 1: Example for an Arkane input file

```
modelChemistry = "CBS-QB3"

species('H', 'path/to/file/H.py', structure=SMILES('[H]'))
species('C2H4', 'path/to/file/ethene.py', structure=SMILES('C=C'))
species('C2H5', 'path/to/file/ethyl.py', structure=SMILES('[CH2]C'))
transitionState('TS', 'path/to/file/TS.py')

thermo('C2H4', 'NASA')

reaction(label='C2H4 + H <=> C2H5',
         reactants=['C2H4', 'H'],
         products=['C2H5'],
         transitionState='TS',
         tunneling='Eckart')
```

```
kinetics(label='H + C2H4 <=> C2H5',
         Tmin=(300, 'K'), Tmax=(2500, 'K'), Tcount=15)
```

Arkane Species

Species definitions are at the core of thermodynamic properties, and high-pressure-limit and pressure-dependent rate coefficient computations. Table 4 describes possible arguments of a *species* section. This section must be included for each relevant species.

Table 4: Arguments of an Arkane *species* segment

| Information type | Keyword | Format | Comments |
|------------------------------------|----------------------------------|----------------------|---|
| Label | <code>label</code> | String | Required. This is a unique identifier, usually the species name. |
| Arkane species file ^a | – | String | A path to an Arkane species file |
| Structure ^b | <code>structure</code> | – | Examples are given in Code Excerpt 2. |
| Zero point energy | <code>E0</code> | Tuple[Float, String] | Value and units, e.g., (-16.77, 'kJ/mol') |
| Modes | <code>modes</code> | List | Entries are RMG's <code>Translation</code> , <code>Rotation</code> , <code>Vibration</code> , and <code>Torsion</code> child classes. |
| Spin multiplicity | <code>spinMultiplicity</code> | Integer | |
| Number of optical isomers | <code>opticalIsomers</code> | Integer | Allowed values are 1 or 2. |
| Molecular weight | <code>molecularWeight</code> | Tuple[Float, String] | Value and units, e.g., (46.0254, 'amu'). |
| Collision model ^c | <code>collisionModel</code> | – | The Lennard-Jones ϵ and σ parameters, see discussion under Section 6. |
| Energy transfer model ^c | <code>energyTransferModel</code> | – | See discussion under Section 6. |

^a Could either be a Python or a YAML file, see text. If this argument is given, no other argument other than *label* is required, and vice versa - if no other argument is given, a file path is required. Note: this is a positional (non-keyword) argument.

^b A species structure is not required, yet it is highly recommended to provide it as it assists Arkane in auto-parameter assignments and is used for plotting outputs. A structure definition may be given using either SMILES, RMG adjacency list, or InChI.

^c Required for pressure-dependent calculations.

Species definitions can be accomplished via three different methods: 1. Using an Arkane Python-format species file (in addition to the main Arkane input file), pointing to log files of completed quantum chemistry calculations which Arkane will parse for the necessary molecular properties. 2. Entering the molecular properties in a *species* block directly in the main Arkane input file. 3. Pointing to a YAML-format species file generated by a previous Arkane run (originally using one of the former methods), containing all the relevant molecular properties. Each species must be defined by only one of the above methods, while

within a single input file any of the above methods may be used for different species. Table 5 describes possible sections in an Arkane species file (method 1), and Code Excerpt 3 shows an example of such file defining ethyl radical, $C_2H_5\cdot$. Both examples relate to method 1.

Code Excerpt 2: Various ways to define species structure

```
structure=SMILES('C[CH2]')

structure=adjacencyList("""multiplicity 2
                            1 C u0 p0 c0 {2,S} {5,S} {6,S} {7,S}
                            2 C u1 p0 c0 {1,S} {3,S} {4,S}
                            3 H u0 p0 c0 {2,S}
                            4 H u0 p0 c0 {2,S}
                            5 H u0 p0 c0 {1,S}
                            6 H u0 p0 c0 {1,S}
                            7 H u0 p0 c0 {1,S}""")

structure=InChI('InChI=1S/C2H5/c1-2/h1H2,2H3')
```

An Arkane species input file (e.g., Code Excerpt 3) refers to log file paths of the quantum calculations. Other than internal torsional information, all non-required arguments as indicated in Table 5 can be auto-assigned by Arkane if left unspecified. Bond enumeration is used to apply respective Petersson-type⁷⁵ bond additivity corrections⁷⁶ that correspond to the specific single point energy level of theory. The rotors section could define any combination of free and/or hindered torsional modes; atom indices are 1-indexed and correspond to the atom order in the electronic structure log files.

Table 5: Sections of an Arkane *species* input file

| Information type | Keyword | Format | Required |
|---------------------|-------------------------------|--------------------------------------|----------|
| Linearity | <code>linear</code> | Boolean | No |
| External symmetry | <code>externalSymmetry</code> | Integer | No |
| Optical isomers | <code>opticalIsomers</code> | Integer | No |
| Spin multiplicity | <code>spinMultiplicity</code> | Integer | No |
| Bond enumeration | <code>bonds</code> | Dictionary[String, Integer] | No |
| Geometry | <code>geometry</code> | Log ^a | Yes |
| Frequencies | <code>frequencies</code> | Log ^a | Yes |
| Energy | <code>energy</code> | Dictionary[String, Log] [†] | Yes |
| Rotors ^b | <code>rotors</code> | List | No |

^a The *Log* class is used to specify an Electronic Structure Software (ESS) output file path.

^a The *rotors* argument may contain any combination of *HinderedRotor* or *FreeRotor* class entries. The *HinderedRotor* class accepts three mandatory arguments (a *Log* class argument, *pivots* which is a list of the indices of the torsion pivotal atoms (e.g. the atoms that define the bond that is being rotated), and *top* which is a list of the indices of all the atoms on one side of the bond and two optional arguments (*symmetry* which is the internal torsion symmetry number and *fit* which directs Arkane whether to use a ‘cosine’, a ‘fourier’, or ‘best’ of both for the fit). Note that the internal torsion symmetry number is estimated by Arkane if left unspecified by the user.

Code Excerpt 3: Contents of an Arkane species file

```
linear = False
externalSymmetry = 1
opticalIsomers = 1
spinMultiplicity = 2
bonds = {'C-C': 1, 'C-H': 5}
geometry = Log('c2h5.log')
frequencies = Log('path/to/file/c2h5_freq.log')
energy = {'CBS-QB3': Log('path/to/file/c2h5.log')}
rotors = [HinderedRotor(scanLog=Log('path/to/file/c2h5_scan.log'), pivots=[1, 2], top=[1, 3, 4])]
```

Table 6: Arguments of the *thermo* segment

| Information Type | Keyword | Format | Required | Comments |
|-------------------------------------|--------------------|--------|----------|---|
| Label | <i>label</i> | String | Yes | A label corresponding to a pre-defined <i>species</i> |
| Output functional form ^a | <i>thermoClass</i> | String | Yes | Allowed values: ‘NASA’ or ‘Wilhoit’ |

^a When requesting a thermodynamic job, the user is required to specify the form of the output: either NASA polynomials⁷⁷ or the Wilhoit form.⁷⁸

Thermochemistry Calculations

The specification of a thermochemistry calculation is done within the main Arkane input file, and is specified separately for each previously-defined species for which thermodynamic parameters are requested. Table 6 describes possible arguments in an Arkane thermochemistry calculation, and Code Excerpt 4 shows an example of a main Arkane input file used to calculate thermodynamic parameters for ethyl radical, $C_2H_5\cdot$.

Code Excerpt 4: Contents of an Arkane thermodynamic calculation job

```
modelChemistry = 'CBS-QB3'  
useHinderedRotors = True  
species('C2H5', 'path/to/file/c2h5.py', structure=SMILES('C[CH2]'))  
thermo('C2H5', 'NASA')
```

The level of theory specification is used to apply appropriate atom, bond, and spin-orbital energy corrections. An up-to-date list of supported levels of theory, the supported chemical elements, and available energy corrections at each level is provided online in Arkane’s documentation.⁷⁴ A *thermo* job specification is specified separately for each defined species for which thermodynamic data is requested.

Table 7: Arguments for a *reaction* segment

| Information Type | Keyword | Format | Required |
|-----------------------|------------------------------|--------------|----------|
| Label | <code>label</code> | String | Yes |
| Reactants | <code>reactants</code> | List[String] | Yes |
| Products | <code>products</code> | List[String] | Yes |
| Transition state | <code>transitionState</code> | String | Yes |
| Tunneling | <code>tunneling</code> | String | No |
| Kinetics ^a | <code>kinetics</code> | – | No |

^a An optional argument used for computing the microcanonical rate coefficient of a reaction using the ILT method for solving the ME.

High-Pressure-Limit Reaction Rate Coefficients

For high-pressure-limit reaction rate coefficient $k_{\infty}(T)$ calculations, all relevant stationary points (i.e., species and transition state), as well as a *reaction* block, referencing the relevant stationary points, must be defined. A *kinetics* job specification, pointing to the label of the relevant reaction and providing the temperature range of interest, is specified separately for each defined *reaction* for which a high-pressure-limit rate coefficient computation is requested. Tables 7-8 describe the arguments of the *reaction* and *kinetics* blocks in a high-pressure-limit rate coefficient calculation, and Code Excerpt 5 shows an example of a main Arkane input file used to calculate high-pressure-limit rate coefficient parameters for the $\text{C}_2\text{H}_4 + \text{H}\cdot \rightleftharpoons \text{C}_2\text{H}_5\cdot$ reaction.

Table 8: Arguments for a *kinetics* segment

| Information Type | Keyword | Format | Required | Comments |
|--|-------------------------------------|----------------------------|----------|--|
| Label | <code>label</code> | String | Yes | A label corresponding to a pre-defined <code>reaction</code> |
| Minimum temperature ^a | <code>Tmin</code> | Tuple[Float, String] | No | Value and units, e.g., (300, 'K') |
| Maximum temperature ^a | <code>Tmax</code> | Tuple[Float, String] | No | Value and units, e.g., (2500, 'K') |
| Number of temperature points ^a | <code>Tcount</code> | Integer | No | The number of temperature points between the minimum and maximum, spaced according to the reciprocal temperature |
| Discrete temperatures ^a | <code>Tlist</code> | Tuple[List[Float], String] | No | Values and units, e.g., ([300, 700, 1200, 2500], 'K') |
| Sensitivity analysis conditions ^b | <code>sensitivity_conditions</code> | List[Tuple[Float, String]] | No | E.g., [(300, 'K'), (1000, 'K')] |

^a Either `Tlist` or `Tmin`, `Tmax`, and `Tcount` must be specified.

^b See discussion under Section 7.

Code Excerpt 5: Contents of an Arkane high-pressure-limit rate coefficient calculation job

```
modelChemistry = 'CBS-QB3'  
useHinderedRotors = True  
species('H', 'path/to/file/h.py', structure=SMILES('[H]'))  
species('C2H4', 'path/to/file/c2h4.py', structure=SMILES('CC'))  
species('C2H5', 'path/to/file/c2h5.py', structure=SMILES('C[CH2]'))
```

```

transitionState('TS1', 'path/to/file/TS1.py')
reaction(label='H + C2H4 <=> C2H5',
         reactants=['H', 'C2H4'],
         products=['C2H5'],
         transitionState='TS1',
         tunneling='Eckart')
kinetics(label='H + C2H4 <=> C2H5',
         Tmin=(300, 'K'), Tmax=(2500, 'K'), Tcount=25)

```

Table 9: Arguments for a pressure-dependent network segment

| Information Type | Keyword | Format | Required |
|-----------------------------------|----------------------------|---------------------------|----------|
| Label | <code>label</code> | String | Yes |
| Isomers ^a | <code>isomers</code> | List[String] | Yes |
| Reactants ^a | <code>reactants</code> | List[Tuple[String]] | No |
| Products ^a | <code>products</code> | List[Tuple[String]] | No |
| Bath gas mixture ^a | <code>bathGas</code> | Dictionary[String, Float] | Yes |
| Elementary reactions ^b | <code>pathReactions</code> | List[String] | No |

^a String entries must correspond to previously defined `species`.

^b String entries must correspond to previously defined `reactions`.

Pressure-Dependent Reaction Rate Coefficients

Determination of pressure-dependent reaction rate coefficients is somewhat more involved than either thermochemistry or high-pressure-limit kinetics calculations. To execute the former, each species participating in a pressure-dependent network must be defined along with respective collision parameters and an energy transfer model (see detailed discussion and comparison of the latter under Section 6). Moreover, in addition to appropriate species, transition states, and reaction input segments, these calculations require a *network* block which specifies the species participating in the pressure-dependent system and their respective roles as reactants in bimolecular entry or exit channels or as isomers (single-species wells). Tables 9-10 describe possible arguments in an Arkane pressure-dependent network and ME parameters, respectively, while Code Excerpt 6 shows a partial example of an Arkane

input file used to calculate pressure-dependent rate coefficient parameters for the C₂H₅OH PES (note that species, transition states, and reaction segments are not shown).

Code Excerpt 6: Partial contents of an Arkane pressure-dependent rate coefficient calculation job

```
network(  
    label='acetyl + O2',  
    isomers=['acetylperoxy', 'hydroperoxylvinoxy'],  
    reactants=[('acetyl', 'oxygen')],  
    bathGas={'N2': 1.0},  
)  
  
pressureDependence(  
    label='acetyl + O2',  
    Tmin=(300.0, 'K'), Tmax=(2000.0, 'K'), Tcount=25,  
    Pmin=(0.01, 'bar'), Pmax=(100.0, 'bar'), Pcount=10,  
    maximumGrainSize=(1.0, 'kcal/mol'),  
    minimumGrainCount=250,  
    method='modified strong collision',  
    interpolationModel=('chebyshev', 6, 4),  
)
```

Outputs

There are several files generated during an Arkane run. First, a log file, *arkane.log*, is generated, logging the calculation steps and any warnings or errors raised during the operations. The central output file, *output.py*, attempts to condense the relevant information from the calculations into a single Python-syntax document, which can be useful for copying and reusing this data, but is not intended to be run independently. In addition, a Chemkin readable file with the relevant thermodynamic and rate coefficient parameters is generated, allows for convenient analysis of the output in common chemical kinetics packages. A YAML format file bearing all necessary molecular properties for future calculations

Table 10: Arguments for a pressure-dependent ME parameters segment^a

| Information Type | Keyword | Format | Required | Comments |
|--|-------------------------------------|----------------------------------|----------|--|
| Label | <code>label</code> | String | Yes | A label corresponding to a pre-defined network |
| ME solution method | <code>method</code> | String | Yes | Allowed values: 'modified strong collision', 'reservoir state', 'chemically-significant eigenvalues', 'simulation least-squares' |
| Interpolation model | <code>interpolationModel</code> | Tuple | Yes | Allowed values: ('chebyshev', T_basis, P_basis), ('pdeparrhenius') ^b |
| Upper bound grain spacing | <code>maximumGrainSize</code> | Tuple[Float, String] | No | Defines the upper bound on grain spacing in master equation calculation. E.g.: (1.0, 'kcal/mol') |
| Minimum grain count | <code>minimumGrainCount</code> | Integer | No | Defines the minimum number of grains in the master equation calculation |
| Active K rotor | <code>activeKRotor</code> | Boolean | No | Whether to treat the K-rotor as active (default) or adiabatic |
| Active J rotor | <code>activeJRotor</code> | Boolean | No | Whether to treat the J-rotor as active (default) or adiabatic |
| Minimum temperature ^c | <code>Tmin</code> | Tuple[Float, String] | No | |
| Maximum temperature ^c | <code>Tmax</code> | Tuple[Float, String] | No | |
| Number of temperature points ^c | <code>Tcount</code> | Integer | No | The number of temperature points between the minimum and maximum ^c |
| Discrete temperatures ^c | <code>Tlist</code> | Tuple[List[Float], String] | No | Values and units, e.g., ([300, 700, 1200, 2500], 'K') |
| Minimum pressure ^c | <code>Pmin</code> | Tuple[Float, String] | No | |
| Maximum pressure ^c | <code>Pmax</code> | Tuple[Float, String] | No | |
| Number of pressure points ^c | <code>Pcount</code> | Integer | No | The number of pressure points between the minimum and maximum ^c |
| Discrete pressures ^c | <code>Plist</code> | Tuple[List[Float], String] | No | Values and units, e.g., ([0.01, 0.05, 0.1, 1], 'bar') |
| Sensitivity analysis conditions ^d | <code>sensitivity_conditions</code> | List[List[Tuple[Float, String]]] | No | E.g.: [[(1000, 'K'), (1, 'bar')], [(1500, 'K'), (10, 'bar')]] |

^a This segment is specified in the input file using the `pressureDependence` keyword.

^b The 'chebyshev' interpolation method is followed by the number of basis sets in temperature and pressure (both are integers). For Chebyshev polynomials a Gauss-Chebyshev distribution is used. Alternatively, if the 'pdeparrhenius' interpolation is requested, a linear distribution on an inverse temperature domain or a logarithmic pressure domain is used. Note that the Gauss-Chebyshev grid does not place `Tmin`, `Tmax`, `Pmin`, and `Pmax` at the endpoints, yet the interpolation is still valid up to these values.

^c Either `Tlist` or `Tmin`, `Tmax`, and `Tcount` must be specified. Likewise, either `Plist` or `Pmin`, `Pmax`, and `Pcount` must be specified.

^d See discussion under Section 7.

is given in a *species* subfolder per species for which a structure was given and per transition state for which structures of all reactants and products of the corresponding reaction are given. Also, an RMG dictionary file, *species_dictionary.txt* with 2D graph adjacency list representations of all species for which a structure was specified is given. Results of sensitivity analysis are given both as a human-readable text file (formatted as a table) as well as a more machine-friendly YAML file. Various parameters which are useful for scientific publications are listed in a comma-separated file titled *supporting_information.csv*, including: symmetry number, number of optical isomers, symmetry point group, rotational constants, computed (unscaled and unprojected) frequencies, electronic energy, zero point energy, 3D coordinate, T1 diagnostic, D1 diagnostic. Finally, the computed parameters are also saved in an RMG thermodynamic and/or kinetic library format, as appropriate, to facilitate adding the calculated data to the RMG database. Plots are automatically generated for thermodynamic parameters vs. temperature (heat capacities, enthalpies, entropies and free energies) torsion scans, PES of pressure-dependent networks and high pressure-limit reactions, rate coefficients vs. temperature and vs. pressure where relevant, and sensitivity analysis coefficients.

3. Statistical Mechanics

The Molecular Partition Function

Statistical mechanics can be used for understanding and predicting macroscopic phenomena as well as calculating macroscopic properties from the properties of the individual molecules that make up the system. The partition function, Q , is the sum over all states of the respective Boltzmann factor. It plays a prominent role in statistical mechanics and forms a connection between the accessible quantum mechanical energy states available to an N -body system and its thermodynamic functions. The calculation of Q for an N -body system of identical, indistinguishable ideal gas particles, in which the number of available states is far

greater than the number of particles, can be reduced to calculating the molecular partition function, q , for a one-body system. See the Supporting Information for more details.

Accurate molecular energies are obtained by solving the Schrödinger equation, yet practical computations are greatly simplified by invoking the Born-Oppenheimer (separation between nucleic and electronic motions), ideal gas (negligible intermolecular interactions), and rigid rotor-harmonic oscillator (RRHO, rotational and vibrational motions are separable; neglecting vibrational anharmonicities)^{79,80} approximations. These assumptions allow one to separate the partition function into the translational, rotational, torsional (internal rotation), vibrational, electronic, and nuclear components (Eqs. S3–S8).^{79,81} See the Supporting Information for more details.

Note that in the absence of an external magnetic field, energy levels associated with nuclear spins are degenerate, and consequently only contribute a multiplicative constant to Q . This results in a constant additive factor to the entropy, which cancels out when computing entropies of formation, free energies of formation, rate coefficients, and equilibrium constants. Due to this reason, and following common practice in this subfield, we therefore omit this nuclear spin factor in all that follows. By default, Arkane assumes a fixed mass for each element. For problems where one is interested in isotope effects, additional corrections are needed, see, e.g., Goldman et al.⁸²

The isobaric heat capacity, $C_{p,T}^{\circ}(M)$, is related to the partition function,⁸³ and for a polyatomic nonlinear molecule it is assumed to be the conventional RRHO ideal gas expression with corrections for internal torsional modes:

$$C_{p,T}^{\circ}(M) = 4R + N_{vib}R + RT^2 \left(\frac{\partial^2}{\partial T^2} \ln q_{torT}^{\circ} \right)_{p^{\circ}} + 2RT \left(\frac{\partial}{\partial T} \ln q_{torT}^{\circ} \right)_{p^{\circ}} \quad (1)$$

where N_{vib} is the number of vibrational degrees of freedom, and q_{torT}° is the torsional partition function (Eq. S6). For small-amplitude vibrations, Arkane uses the harmonic oscillator approximation:

$$C_v^{HO}(T, \nu_i) = R \cdot e^{\frac{\nu_i}{k_B T}} \left(\frac{\frac{\nu_i}{k_B T}}{e^{\frac{\nu_i}{k_B T}} - 1} \right)^2 \quad (2)$$

where ν_i is the corresponding harmonic oscillator frequency and k_B is the Boltzmann constant. The vibrational frequencies and the information needed to compute the torsional contribution of $C_{p,T}^o(M)$ can come either from quantum chemical calculations or be estimated. Here, we consider the case where the frequencies are coming from quantum calculations; below we explain the estimation method if the quantum data is unavailable. Internal rotations are automatically identified, assuming that any non-terminal single bond not part of a ring is rotatable.

The effect of torsions on the harmonic frequencies is eliminated by projection. For this end, the Hessian (matrix of partial second derivatives of the potential energy surface) must be made available to Arkane from the ESS log file. Since the Hessian is often not logged by default by many ESS, the user should add an appropriate directive to do so before executing the frequencies calculation. For example, the user must specify the “*iop(7/33=1)*” keyword in Gaussian or the “*print,HESSIAN*” keyword in Molpro. The frequencies are scaled either using a user-defined scaling factor, or by drawing values from the RMG-database when available for the corresponding level of theory. An updated list of levels of theory for which the frequency scaling factor is available in Arkane is given in Arkane’s documentation.⁷⁴ The vibrational partition function is calculated quantum-mechanically subsequent to the torsional frequencies projections accounting for the scaling factor.

The heat capacity of a semi-classical hindered rotor with a cosine potential can be expressed as approximately:

$$\frac{C_v^{HR}(T, V_0, \nu)}{R} = \frac{d}{dT} \left(T^2 \frac{d \ln(Q_{HRsemi})}{dT} \right) \approx -\frac{1}{2} + \eta \left[\eta - \frac{I_1(\eta)}{I_0(\eta)} - \eta \left(\frac{I_1(\eta)}{I_0(\eta)} \right)^2 \right] + \left(\frac{2\eta \frac{\nu}{V_0} e^{-\frac{\nu}{V_0} \eta}}{1 - e^{-2\frac{\nu}{V_0} \eta}} \right)^2 \quad (3)$$

with

$$\eta = \frac{V_0}{2k_B T} \quad (4)$$

where I_n denotes the corresponding Bessel function,⁸⁴ and the unknowns V_0 and ν denote the barrier height and frequency corresponding to the rotor respectively.⁸⁵ The heat capacity of a single harmonic oscillator frequency can be given by the sum of the two last factors in Eq. (2). N_{vib} vibrations and N_{rot} hindered rotors are then fit to 13 residual heat capacity points (Pseudocode Excerpt 7), giving a full set of frequencies and rotors for each molecule. Along with the assumption that each molecule is a symmetric top modeled as a 1-D K-rotor, this enables Arkane to evaluate the density of states of a given molecule.

Code Excerpt 7: Pseudo code description for setting N_{vib} and N_{rot}

```

if N_vib + 2 * N_rot < 13:
    number of vibrations = N_vib
    number of rotors = N_rot
else if N_vib + 2 < 13:
    number of vibrations = N_vib
    number of rotors = 1
else:
    number of vibrations = 3
    number of rotors = 1

```

During chemical kinetic model generation,³¹ large quantities of data are required, and often access to electronic structure calculations is limited or non-existing. Arkane also supports

such cases, in which statistical mechanical frequencies can be estimated using the RMG-database⁸⁶ using a group frequency model.⁸⁵ Internal rotations are automatically identified, assuming that any single bond in localized structures is rotatable. Each heavy atom and the atoms directly bonded to it descend down the group frequency tree and are assigned $3N_{atoms} - N_{rotors} - 6$ frequencies associated with the most specific match. Each group is associated with a set of frequency ranges with numbers of frequencies for each range. Frequencies associated with the same range are spaced linearly across the range. In the entire molecule, if $N_{groupfreq} > 3N_{atoms} - 6 - N_{rotors}$ then the matches with the lowest symmetry are removed until $N_{groupfreq} \leq 3N_{atoms} - 6 - N_{rotors}$. A residual heat capacity is then calculated by subtracting the quantitatively known modes for vibration, translation, and external rotation. For a nonlinear molecule this gives Eq. (5). The residual heat capacity is then fit to an appropriate number of vibrations and rotations.

$$C_v^{Res}(T) = C_p(T) - 4R - \sum_{i=1}^{N_{groupvib}} C_i^{GroupVib}(T) \quad (5)$$

Multidimensional Torsion Treatment

In many molecules, molecular torsions interact with each other and cannot be treated independently. For these cases Arkane provides two methods to handle the coupling. The first is an interface to the 2D-NS method within the Q2DTor software.⁸⁷ This method treats the rotations of two rotors quantum mechanically as independent of the motions of the rest of the molecule. The second is an implementation of ND classical and semi-classical rotors. Multidimensional potential data and the reduced moment of inertia at each point are fit to piecewise linear interpolants. This allows the ND classical partition function to be computed as:

$$Q_{classical} = \left(\frac{2\pi k_B T}{h^2} \right)^{\frac{N}{2\prod \sigma_i}} \int \cdots \int_N \sqrt{I(\phi_1, \dots, \phi_N)} \exp \left(\frac{-V(\phi_1, \dots, \phi_N)}{RT} \right) d\phi_1 \dots d\phi_N \quad (6)$$

where N is the number of rotors, V is the interpolated potential, I is the interpolated reduced moment of inertia, and σ_i is the rotor symmetry number of the i^{th} rotor. The integration is carried out using multi-dimensional Simpson’s rule. We can construct the semi-classical partition function by first finding the harmonic oscillator partition functions for the torsions:

$$Q_{HO_i} = \frac{h\nu_i}{k_B T} \quad (7)$$

The semi-classical partition function can then be computed from the semi-classical correction:⁸⁸

$$\frac{Q_{semiclassical}}{Q_{classical}} = \prod_i \frac{Q_{HO_i}}{1 - \exp(-Q_{HO_i})} \quad (8)$$

The harmonic oscillator frequencies are obtained by fitting an N-dimensional polynomial to the potential around the equilibrium geometry. The Hessian is then computed at the equilibrium geometry from this function and the frequencies calculated from its eigenvalues. The semi-classical method agrees well with quantum mechanical methods in 1D and provides a convenient and much faster alternative to the more accurate, but more expensive 2D-NS method from Q2DTor within Arkane.

4. Calculation of Thermochemical Properties

In kinetic mechanism simulations, the most important thermochemical property is usually the Gibbs free energy of a reaction, $\Delta_r G_T^\circ$, which provides a means for calculating the equilibrium constant of the reaction, K_{eq}° .⁸³

$$K_{\text{eq}}^{\circ} = \exp \left[-\frac{\Delta_{\text{r}}G_T^{\circ}}{RT} \right] \quad (9)$$

The T subscript denotes that the Gibbs free energy depends on temperature itself and the circle superscript indicates the standard pressure of 1 bar. The Gibbs free energy of reaction is obtained from the corresponding enthalpy, $\Delta_{\text{r}}H_T^{\circ}$, and entropy, $\Delta_{\text{r}}S_T^{\circ}$, of reaction.

While the enthalpy of formation cannot be obtained from the molecular partition functions and has to be derived from additional quantum chemical energy calculations (described below), its temperature-dependence, the molecular entropy, and molecular isobaric heat capacity are easily obtained from the partition functions (Eqs. S3–S8).⁸³ Note that the partition function calculation itself assumes ideal gas behaviour of the phase, and therefore inter-molecular interactions are not considered.

Enthalpy of Formation from Atomization Energies

Given the electronic energy of a molecule, $E_0(M)$, from a quantum chemical calculation including zero-point energy, the atomization energy is:

$$\Delta_{\text{at}}H_0^{\circ}(M) = \left[\sum_{a \in A(M)} E_0(a) \right] - E_0(M) \quad (10)$$

Here, the set of all atoms in the molecule M is denoted by $A(M)$ and a is an atom in that set. The electronic energy of an atom, $E_0(a)$, typically incorporates spin-orbit corrections as well.⁸³ Next, the enthalpy of formation at 0 K is given by:

$$\Delta_{\text{f}}H_0^{\circ}(M) = \left[\sum_{a \in A(M)} \Delta_{\text{f}}H_0^{\circ}(a) \right] - \Delta_{\text{at}}H_0^{\circ}(M) \quad (11)$$

$\Delta_{\text{f}}H_0^{\circ}(a)$ are known 0 K enthalpies of formation of the atoms. Very accurate values are available for some atoms in the Active Thermochemical Tables.^{89,90} Lastly, calculating the enthalpy of formation at nonzero temperatures requires enthalpy increments for the molecule

and all elements in their reference states:^{83,91}

$$\Delta_f H_T^\circ(M) = \Delta_f H_0^\circ(M) + \Delta H_{0 \rightarrow T}^\circ(M) - \sum_j \Delta H_{0 \rightarrow T}^\circ(\text{element}_j) \quad (12)$$

$\Delta H_{0 \rightarrow T}^\circ(M)$ is given by Eq. S12 and $\Delta H_{0 \rightarrow T}^\circ(\text{element})$ are also known and tabulated values, usually for $T = 298 \text{ K}$, are available from CODATA.⁹² These values are for elements in their reference states renormalized per constituent atom. For example, the value for carbon corresponds to graphite and the value for hydrogen corresponds to one half of $\text{H}_2(\text{g})$.

Bond Additivity Corrections

Calculated *ab initio* atomization energies are frequently not very accurate because the electronic states of molecules have a significantly different character from the atoms that make up the molecules. Empirically correcting systematic errors in atomization energies and enthalpies of formation associated with bond or atom types using bond additivity corrections (BACs) is a common approach to improve the accuracy of *ab initio* calculations.

Arkane implements two different types of BACs. The first implementation, Peterson-type, corrects for specific bond types (C–H, C–C, C–O, C=C, etc.) by using a fitted correction for each type of bond and summing the parameters for all bonds in a molecule.⁷⁵ The corrected enthalpy of formation is thus given by:

$$\Delta_f^{\text{BAC}} H_{298}^\circ(M) = \Delta_f H_{298}^\circ(M) + \sum_{b \in B(M)} C(b) \quad (13)$$

Here, $B(M)$ is the set of bonds in molecule M and $C(b)$ is the correction parameter for bond b .

The BACs defined in Eq. (13) require a parameter for every bond type that might be encountered, which means that they cannot be applied to molecules with bond types that were not present in the training data. Furthermore, the corrections could be problematic for molecules with significant resonance character or any molecule with bonds of non-integer

bond order. Alternative BACs that use parameters for atom types instead of bond types, Melius-type, are given by:⁹³

$$\Delta_f^{\text{BAC}} H_{298}^\circ(M) = \Delta_f H_{298}^\circ(M) - E_{\text{BAC}}(M) - \sum_{a \in A(M)} \alpha(a) - \sum_{b \in B(M)} E_{\text{BAC}}(b) \quad (14)$$

$E_{\text{BAC}}(M)$ is a molecular correction term due to the spin of the molecule and the individual atoms comprising the molecule, given by:

$$E_{\text{BAC}}(M) = K \cdot \left(S(M) - \sum_{a \in A(M)} S(a) \right) \quad (15)$$

where K is an empirical parameter, and $S(M)$ and $S(a)$ are spin quantum numbers for the molecule M and the atom a , respectively. The first sum in Eq. (14) captures corrections due to all atoms in the molecule where α is a fitted parameter for each atom type. The second sum in Eq. (14) contains bond corrections given by:

$$E_{\text{BAC}}((x, y)) = \sqrt{\beta(x)\beta(y)}e^{-\xi R_{xy}} + \sum_{w \in N(x) \setminus y} [\gamma(w) + \gamma(x)] + \sum_{z \in N(y) \setminus x} [\gamma(z) + \gamma(y)] \quad (16)$$

where x and y are the atoms in the bond, R_{xy} is the bond distance between the two atoms, β and γ are additional parameters for each atom type, the set of atoms immediately adjacent to atom a is denoted by $N(a)$, and ξ is set to 3 \AA^{-1} , the value recommended by the developers of this method.⁹³

Database of Reference Species

When fitting either type of BACs, it is imperative to utilize a large collection of molecules with precisely known heats of formation. To this end, a database of reference species with precisely known heats of formation has been compiled from various sources for use in Arkane, and stored in the accompanying RMG-database repository. This reference database is a

collection of just over 400 species with heats of formation known to within 1 kJ/mol, taken from the Active Thermochemical Tables,⁸⁹ Cioslowski et al.,⁹⁴ NIST,⁹⁵ Pedley et al.,⁹⁶ and the Computational Chemistry Comparison and Benchmark Database (CCCBDB).⁹⁷ The elements present in this set of species include H, C, N, O, F, S, Cl, and Br. The set contains both stable species and radicals, as well as some cationic and anionic species.

The database stores the values (plus uncertainty) for the heat of formation of each molecule from the literature, but also the heats of formation calculated by a subset of the levels of theory that Arkane has BAC support for. The heat of formation calculation for each level of theory includes atom energy corrections (that are fitted separately), but does not include BACs. This way, Arkane can re-fit BACs for any level of theory by loading in this database, allowing for transparency as to how BACs stored in Arkane were obtained. This also allows a user to re-fit BACs on a subset of the reference species or by using a different weighting scheme if desired. Adding in new levels of theory with BAC support in Arkane involves calculating all of the species in the reference database at the desired level of theory and updating the database with these values. Arkane has scripts and code to help with this process, including fitting atom energy corrections and updating the database.

Isodesmic Reactions

Alternatively, instead of applying AECs and BACs, Arkane can calculate heats of formation for species using an isodesmic reaction approach. In this approach, a hypothetical reaction is constructed where the number of each element and type of bond (e.g. C-H, C=C) is conserved. The species involved include the target species in question, as well as any number of additional species that have precisely known heats of formation. An example isodesmic reaction involving benzene as a target is shown in Figure 1.

The key insight behind using isodesmic reactions is that calculating the heat of reaction is much less error prone than the energies of any individual species in the reaction for a given level of theory. This is due to a convenient cancellation of errors, i.e., a systematic error

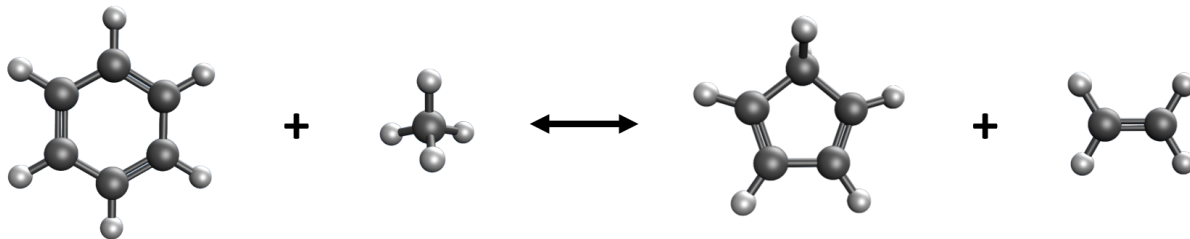


Figure 1: An example isodesmic reaction involving benzene as the target molecule. Notice that while this reaction is likely unphysical, the number of carbon and hydrogen atoms is conserved, in addition to the number of each bond type. A key feature is that all of the non-target species have heats of formation that are precisely known in the literature.

present for a given bond feature in a reactant of an isodesmic reaction will tend to cancel out for the same feature in a respective product.

Using the species found in the aforementioned reference species database, Arkane can automatically find isodesmic reactions using the algorithmic approach of Buerger et al.⁹⁸ Following this algorithmic approach, Arkane supports the reaction classes RC2 (conserve the number of each type of bond on either side of the reaction), RC3 (extends RC2 by also conserving the identity and the total number of additional atoms bonded to any atom on either side of a chemical bond), and RC4 (extends RC3 by also conserving the bond type and identity of each neighbouring atom, including the total number of additional atoms bonded to the neighbouring atoms) as defined in Buerger et al. Arkane also implements additional constraining possibilities, such as conserving the number of rings containing n -atoms (e.g., conserve the number of 6 member rings), conserving charge across the reactions, and limiting the allowable charges for the participating species. Especially when higher reaction classes like RC4 are used, this method can yield highly accurate heats of formation when the target species is similar in nature to those in the reference database.

5. Calculation of High-Pressure-Limit Rate Coefficients

High-pressure-limit rate coefficients can be calculated using traditional transition state theory (TST, Equation (17)),^{99,100} where σ is the reaction path symmetry number,¹⁰¹ ΔV^\ddagger is

the energy barrier (without zero-point energy correction) energy barrier, and $Q_{TS}(T)$, $Q_R(T)$ are the transition state and the reactant well partition functions, respectively.

$$k_{\text{TST}}(T) = \sigma \frac{k_b T}{h} \frac{Q_{\text{TS}}(T)}{Q_R(T)} e^{-\frac{\Delta V^\ddagger}{k_b T}}, \quad (17)$$

Because TST treats the nuclear motion in the reaction coordinate classically, quantum effects such as tunneling are not taken into account. For more accurate rate calculations, one should apply a tunneling correction to the TST rate coefficient (Equation (18)), where $\kappa(T)$ is the tunneling correction factor (also called the transmission coefficient). Since the transmission probability is affected by the shape of the reaction barrier, $\kappa(T)$ needs to be determined based on the potential energy surface along the reaction coordinate. As a first approximation, Hirschfelder and Wigner²⁸ showed that $\kappa(T)$ can be derived as a function of the curvature of the energy surface (Equation (19)), where ν_{TS} is the imaginary frequency of the transition state. This approximation, which only relies on the imaginary TS frequency, is satisfactory when the quantum effect is small; however, this is not always the case in practice.

$$k_{\text{TST}}^{\text{corrected}}(T) = \kappa(T) k_{\text{TST}}(T), \quad (18)$$

$$\kappa(T) = 1 + \frac{1}{24} \left(\frac{h |\nu_{\text{TS}}|}{k_b T} \right)^2, \quad (19)$$

An improved estimate for $\kappa(T)$ can be derived by approximating the potential energy surface using the Eckart function.²⁷ This estimate is particularly useful because it has the proper asymptotic properties of the one-dimensional PES. With a Boltzmann distribution of incident particles, one can solve the tunneling correction $\kappa(T)$ for the Eckart potential using Eqs. S12–S19.^{102,103}

Currently, Arkane supports both the Wigner (Equation (19)) and Eckart (Equation S12) corrections, which can be invoked by using the keywords *Wigner* and *Eckart*, respectively,

when executing a kinetics calculation. The Eckart tunneling approximation is widely used for theoretical kinetics since it is rather inexpensive computationally and reasonably accurate above 500 K for hydrogen transfer reactions, in comparison to more sophisticated multidimensional approximation methods such as small¹⁰⁴ or large¹⁰⁵ curvature tunneling corrections.¹⁰⁶

6. Calculation of Pressure-Dependent Rate Coefficients

Many key gas-phase chemical reactions that take place either within the atmosphere or in industrial processes commonly involve multiple inter-connected potential wells and exhibit strong pressure dependence.⁴⁵ More often than not, these reactions involve radical species whose lifetimes are short and whose rates of unimolecular decay are difficult to measure. While experimental tools such as shock tubes have been invaluable in determining many phenomenological rate coefficients, $k(T, P)$'s, used in detailed kinetic models,^{107,108} they cannot practically provide all necessary rate coefficients over the full range of temperature and pressure conditions encountered in all systems of interest. Thus, rate theory is frequently employed to evaluate pressure-dependent rate coefficients.

Pressure-dependent reactions could be categorized as "chemically activated", where there are several colliding reactants (typically two, but rarely also three¹⁰⁹), or "thermally activated", where there is only one reactant that has to reach an energetically excited state to react, commonly via collisions with the bath gas. Examples include $A + B \rightarrow C$ and $C \rightarrow A + B$, respectively, both forming initially an energetically excited C^* species (Fig. 2). The latter could react via dissociation into $A + B$ or isomerization into D , depending on its energy, and could also be collisionally excited or stabilized. Accounting for a unimolecular reaction network could become important once isomerization and dissociation reactions become competitive with one another as well as with collisional stabilization. Often, an excited species will undergo multiple reactions before being collisionally stabilized, appearing

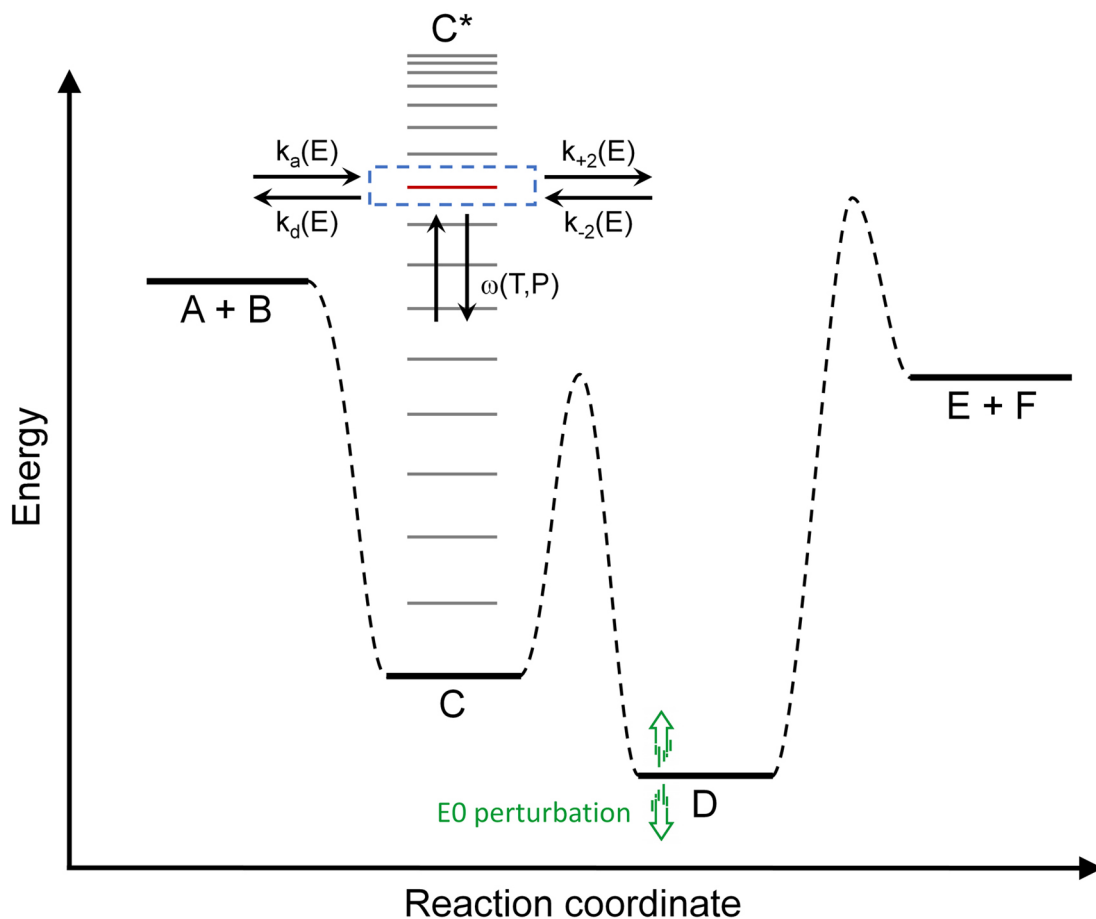


Figure 2: A generalized potential energy surface. Phenomenological rate coefficients, $k(T, P)$, are computed from a population balance on all energy states using the master equation. The gray horizontal lines represent quantized energy levels for isomer C. Arkane computes both the phenomenological rate coefficients $k(T, P)$ and their sensitivities to perturbations in the energies of the stationary points on the PES.

to "skip" over intermediate species due to collisional energy transfer processes being slower than reactive processes (e.g., $A + B \rightarrow D$ in Fig. 2). These processes can be represented as net reaction rates from each potential well to every other potential well on the network, not just those directly connected, where the rate depends on the bath gas concentration (system pressure).

Such pressure-dependent networks can be modelled probabilistically, with the time-evolution of the system governed by a master equation (ME) describing the molecular processes on a collisional timescale.^{44,110,111} A full treatment of the energy states of

large species is challenging, as there are simply too many states to be considered. The Rice–Ramsperger–Kassel–Marcus (RRKM)^{112–114} approximation, on the other hand, simplifies the system and results in a two-dimensional ME where the state of a species is a function of two quantities only: the total energy E and total angular momentum quantum number J , indicating that energy transfers involve both rotation and vibration. The projection of J along the main symmetry axis of the species is termed K , and it is considered active if rotation modes are strongly coupled to vibration modes; otherwise K is conserved (inactive). A two-dimensional ME is relatively difficult to solve, and often an enormous further simplification is made where the dependencies of the microcanonical rate coefficients, $k(E, J)$, and other quantities on J are neglected, and a one-dimensional ME is solved instead, in which E is the only independent variable, i.e., obtaining $k(E)$.^{44,110,115} The user can direct Arkane to treat J and K as either active or adiabatic using the `activeJRotor` and `activeKRotor` keywords, respectively (default is to treat both as active).

Arkane also has an option to approximate an asymmetric top molecule (where none of the principal moments of inertia are equal in magnitude) as a symmetric top, producing a degenerate two-dimensional external rotation (assuming that the rotation having the smallest moment of inertia is separable from the other two).¹¹⁶ The axis having the unique moment of inertia is referred to as the symmetric top K due to the quantum number notation commonly used for this term of the rotational partition function equation. This K -rotor is designated as active and allowed to exchange energy with internal vibrational degrees of freedom.

Bimolecular collisions with an inert species M are the primary means by which a species changes its energy in the present context. A reasonable estimate for the total rate of collisions of species i , $k_{coll,i}(T)$, comes from the Lennard-Jones collision theory (Eq. (20)).¹¹⁷ Here, μ_i is the reduced mass, d_i is the collision diameter, and k_B is the Boltzmann constant. The collision diameter is generally taken as $d \approx \frac{1}{2}(\sigma_i + \sigma_M)$, the arithmetic average of the Lennard-Jones σ parameter for the species and the bath gas. The parameter $\Omega^{(2,2)*}$ represents a configurational integral, which is well-approximated by Eqs. (21)–(22) with ϵ

the Lennard-Jones well energy. Assuming the total gas concentration to be constant and that the gas is ideal, we obtain an expression for the collision frequency $\omega_i(T, P)$ (Fig. 2), which makes explicit the pressure dependence (Eq. (23)). The Lennard-Jones σ_i and ϵ_i parameters can be obtained from literature,^{118,119} estimated by the group additivity method, or be directly calculated.^{117,120} Some literature values along with group additivity estimates can be obtained from the RMG website.¹²¹

$$k_{coll,i}(T) = \sqrt{\frac{8k_B T}{\pi \mu_i}} \pi d_i^2 \Omega^{(2,2)*} \quad (20)$$

$$\Omega^{(2,2)*} = 1.16145 \tilde{T}^{-0.14874} + 0.52487 e^{-0.7732 \tilde{T}} + 2.16178 e^{-2.437887 \tilde{T}} \quad (21)$$

$$\tilde{T} \equiv \frac{k_B T}{\sqrt{\epsilon_i \epsilon_M}} \quad (22)$$

$$\omega_i(T, P) = k_{coll,i}(T) \frac{P}{k_B T} \quad (23)$$

The Lennard-Jones parameters assist in estimating the total rate of collisions, yet the effect these collisions have on the energy state of a species must also be accounted for. While there are a variety of models to estimate the probability of a collision resulting in a transfer of a species from state E' to state E , $P_i(E, E')$, the most common one is the single exponential down model (Eqs. (24)–(25)). The $C(E')$ parameter is determined from normalization constraints, while $\langle \Delta E_d \rangle$ depends on the specific potential energy surface (PES) and the nature of the bath gas. A compilation of literature values for $\langle \Delta E_d \rangle$ are given in Fig 3. As seen in this figure, literature models of relatively large systems tend to have larger $\langle \Delta E_d \rangle$ values. Also, nitrogen-based systems exhibit a somewhat different trend of $\langle \Delta E_d \rangle$ vs. temperature than carbon-based systems.

$$P(E, E') = C(E') e^{-\frac{E'-E}{\langle \Delta E_d \rangle}}, E < E' \quad (24)$$

$$\langle \Delta E_d \rangle = \alpha \left(\frac{T}{T_0} \right)^\gamma \quad (25)$$

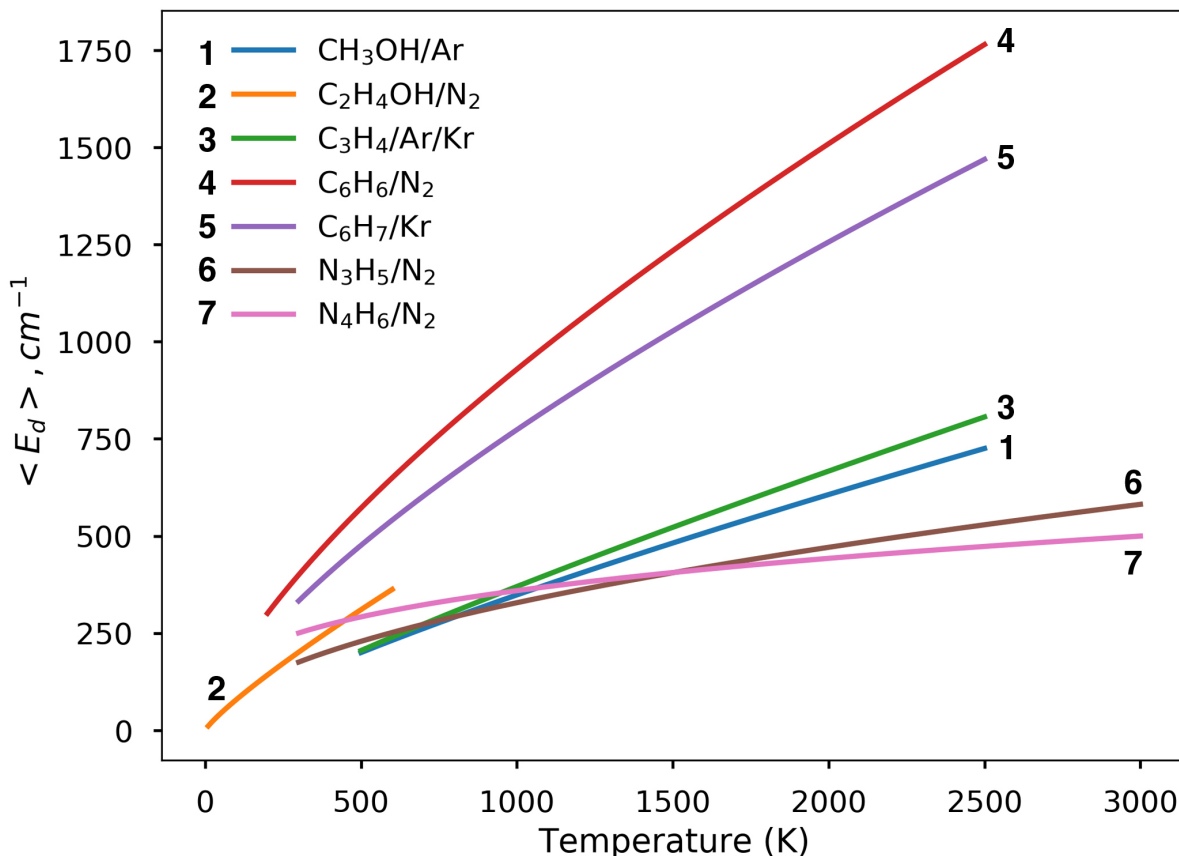


Figure 3: A comparison of literature values for the single exponential down energy transfer model parameter, $\langle \Delta E_d \rangle$. The legend represents the PES and the bath gas used in each study. References: CH₃OH/Ar,¹²² C₂H₄OH/N₂,¹²³ C₃H₄/Ar/Kr,¹²⁴ C₆H₆/N₂,¹²⁵ C₆H₇/Kr,¹²⁶ N₃H₅/N₂,¹²⁷ N₄H₆/N₂.¹²⁷

A transformation of the ME solution into phenomenological rate coefficients is required for predicting the emergent properties of a process, i.e., $k(E, J)$ or $k(E)$ into $k(T, P)$. Three methods of varying rigor, computational cost, and robustness for solving the ME and transforming it into $k(T, P)$'s are implemented in Arkane: The modified strong collision (MSC)² method is the fastest and most robust, but utilizes the least realistic approximations. The

reservoir state (RS)⁴³ method uses better approximations, leading to increased accuracy, but is more computationally demanding. The chemically-significant eigenvalues (CSE)^{44,45} method is the most theoretically sound, but is much more computationally expensive than the other methods and not very robust. Finally, the simulation least-squares (SLS)⁴⁶ method is often about as accurate as CSE and seems to be more robust. Some of these methods were previously extensively discussed and compared elsewhere.¹¹⁵ Briefly, the MSC method treats collisional stabilizations of a reactive isomer as a single-step process rather than collisional energy redistribution within the reactive energy space, the RS method assumes Boltzmann distribution of the low energy grains, while the CSE method simplifies the system by eigendecomposing it and neglecting the faster relaxing modes associated with energy transfer leaving only the chemically significant eigenmodes associated with reaction. The SLS method uses fitting of the ME solution to specific conditions and fluxes.

Arkane can also perform inverse Laplace transformations (ILT) of phenomenological rate coefficients (Eq. (26)),¹²⁸ which can be useful in generating a pressure-dependent network where quantum calculations of transition states are not available (e.g., using experimental or estimated rate coefficients).

$$k(T, P, E_i, J_j) = A \times \frac{\rho(E_i - E_o)}{\rho(E_i)} \quad (26)$$

Here, ρ is the density of states at the specified energy level, E_i is the internal energy being solved, E_o is the internal energy of the transition state complex, E_j is the external rotational energy, and A is the Arrhenius pre-exponential factor which includes the T^n term of the modified Arrhenius form as well as any contribution of a negative activation energy. Arkane then uses this value when solving for the pressure-dependent rates in the entire network at the specified temperature and pressure, and repeats this process for each condition specified.

The output format of the calculated pressure-dependent rate coefficients is either Arrhenius coefficients at selected pressures (called PLOG), or as a Chebyshev polynomial¹²⁹ that can be interpreted by standard kinetic calculators such as Cantera,¹³⁰ Chemkin,¹³¹

7. Additional Features

PES Exploration

One of the challenges in determining pressure-dependent rate coefficients is attaining an initial comprehensive picture of the PES of interest, identifying the significant wells and pathways to be further calculated using electronic structure methods. Arkane takes advantage of its intimate connection to RMG, its database, and its estimation capabilities, to explore unimolecular networks.¹³³ Wells energies and barrier heights are either determined by RMG or parsed from electronic structure calculations (optional). The data generated by RMG originates from its libraries and estimation methods,^{30,31,134} and barrier heights are determined by applying the ILP method.

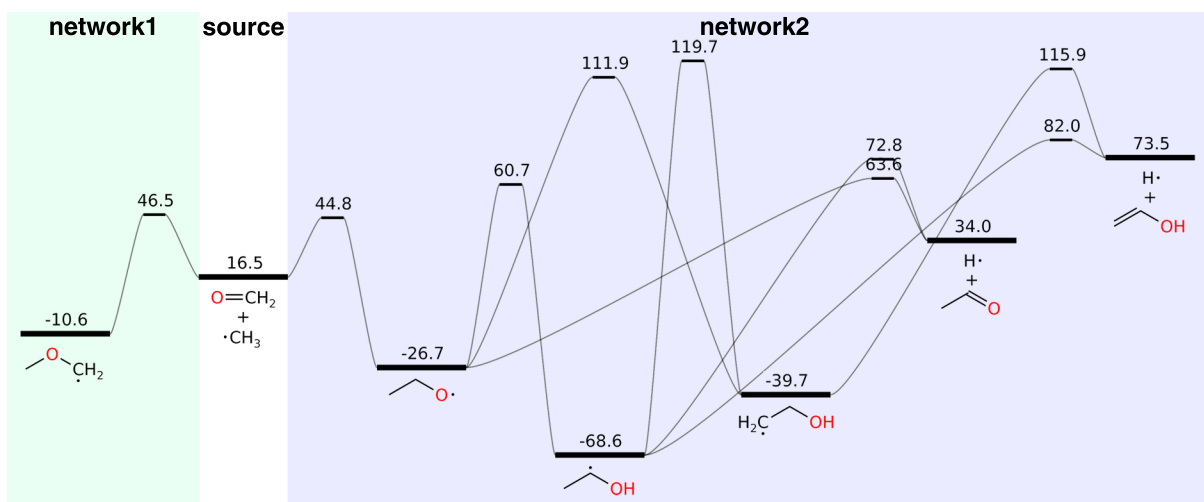


Figure 4: Two pressure-dependent networks explored by Arkane starting from the same source, formaldehyde and methyl radical, generated for 300–2500 K and 0.1–100 bar. Values represent relative energy (E_0) in kJ mol^{-1} . The `explore_tol`, `energy_tol`, and `flux_tol` parameters used were 0.01, 10, and 10^{-8} , respectively. The execution time was about three minutes on a conventional personal computer.

The PES exploration process begins from a specified "source" – either an isomer or bimolecular reactants of interest. Several optional parameters can be set for an explo-

ration job, controlling the size of the reported PES. The primary exploration tolerance (`explore_tol`) controls the network expansion process: whenever a well has a net flux higher than `explore_tol` times the characteristic network rate at any T/P conditions within the desired range, that well will be considered as part of the network. The characteristic network rate is defined as the total flux from all net reactions away from the source. The maximum number of radicals per species could also be specified to limit the species search space. Often, the exploration process may result in relatively large networks that include unimportant reactions, and a systematic reduction scheme is desired. A reduced network is obtained by only keeping reactions that answer at least one of two criteria, pertaining to energy and flux. The energy criterion, `energy_tol`, considers the energy difference between a TS and the source. If this difference is lesser than or equal to a temperature-dependent threshold energy determined by `energy_tol` (Eq. 27), this pathway will be kept. The flux criterion, `flux_tol`, considers the flux fraction a reaction has relative to an arbitrary network input flux, assuming steady state. A reaction with a flux fraction higher than or equal to `flux_tol` will be kept. Several networks could be generated when the source is bimolecular. For example, a reduced C₂H₅O PES generated from formaldehyde and methyl radical as a bimolecular source consisting of two pressure dependent networks is given in Fig. 4.

$$E_{0_{TS_i}} - E_{0_{source}} = < \text{energy_tol}RT \quad (27)$$

PES Sensitivity Analysis

Arkane can compute sensitivities of rate coefficients to E_0 energies of stationary points (wells and TSs) on the studied PES. These sensitivity data allow one to assess the effect of E_0 uncertainties on the calculated rate coefficients. This feature can be applied to both high pressure limit and pressure-dependent rate coefficient calculations simply by adding one of the following `sensitivity_conditions` list attribute to the kinetics or pressure-dependent block of the input file, respectively (Code Excerpt 8).

Code Excerpt 8: PES Sensitivity Analysis conditions

```
sensitivity_conditions = [(1000, 'K'), (2000, 'K')]

sensitivity_conditions = [(1000, 'K'), (1, 'bar')],
                        [(2000, 'K'), (10, 'bar')]]
```

The user may add as many T or T and P conditions as desired in the above format. Arkane will separately perturb each well and TS by 0.1 kcal mol⁻¹ for each condition and will recalculate forward and reverse rates for each case. Semi-normalized sensitivity coefficients are calculated according to Eq. (28), where $S_{i,j,l}$ is the sensitivity coefficient for k_i and the j^{th} perturbed stationary point at the l^{th} condition, k_i is the i^{th} rate coefficient, and E_{0j} is the E_0 value of the j^{th} stationary point. The output of a sensitivity analysis (SA) is saved in a designated `sensitivity` folder, and consists of a text file per reaction / network delineating the sensitivity coefficients in mol J⁻¹ (for both reaction directions, if reversible), along with a horizontal bar figure per desired condition.

$$S_{i,j,l} = \frac{\partial \ln k_i}{\partial E_{0j}} \quad (28)$$

For example, this analysis was recently reported in a theoretical study of the N₃H₅ PES¹³⁵ (Fig. 5). While the entire N₃H₅ PES was calculated using the MSC² method in that study, the sensitivities were calculated using Arkane for a sub-section of the surface implementing the CSE approach,^{44,45} which is more accurate yet less robust (hence, appropriate for a sub-network). The analysis highlights pressure-dependent "fall-off" regions, e.g., as in the case of the N₂H₃ + NH₂ \rightleftharpoons n-N₃H₅ reaction at 1 bar and 1500 K, where increasing E_0 of the reactants well or the product isomer changes the density of states around the entry well, increasing or decreasing the rate, respectively (Fig. 5 a). It also shows a complex dependency of the thermally activated reaction n-N₃H₅ \rightleftharpoons H₂NN(S) + NH₃ rate coefficient on E_0 's of the adduct and the exit channel transition state (Fig. 5 b). Finally, the rate coefficient of the well-skipping reaction N₂H₃ + NH₂ \rightleftharpoons H₂NN(S) + NH₃ is sensitive to the energies of

the entry channel and transition state, and not to the energy of the skipped well (Fig. 5 c).

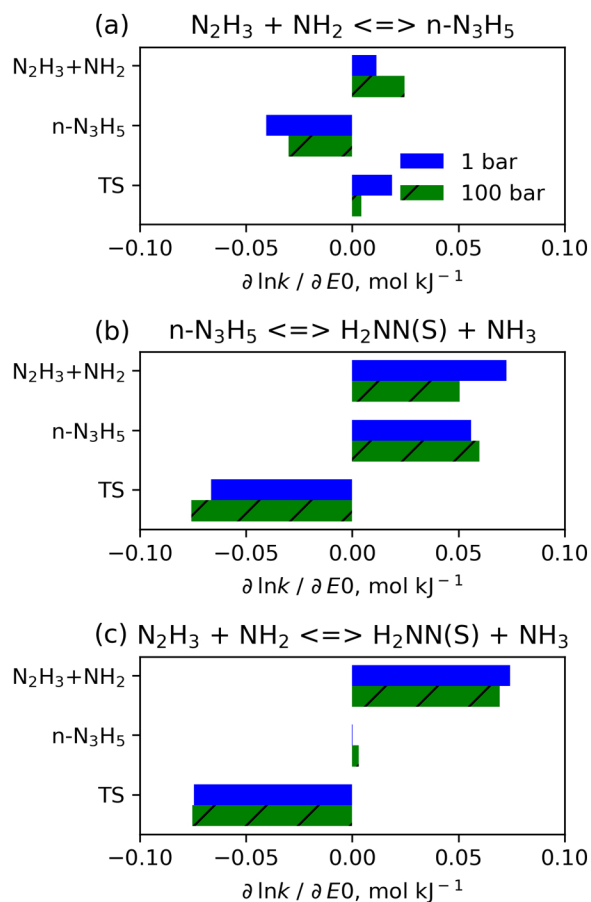


Figure 5: Semi-normalized sensitivity coefficients to E_0 of $\text{N}_2\text{H}_3 + \text{NH}_2$, $n\text{-N}_3\text{H}_5$, and the transition state connecting the $n\text{-N}_3\text{H}_5$ isomer to the $\text{H}_2\text{NN(S)} + \text{NH}_3$ product well for selected pressure-dependent rate coefficients at 1 and 100 bar and at 1500 K. Reprinted (adapted) with permission from J. Phys. Chem. A 2019, 123, 4679–4692.¹³⁵ Copyright 2019 American Chemical Society.

Molecular Properties Files

The statistical mechanics properties of a stationary point, used for the various calculations discussed in this work, are typically available from separate files, i.e., output files of harmonic frequencies, single point, and zero or several torsion scans calculations. An Arkane species file will commonly point to the respective output file paths to be parsed. However, often a more convenient workflow is required, e.g., when the same stationary points is used for

different projects, particularly when executing on different machines, as well as for archiving the information. To facilitate this workflow, Arkane saves a molecular properties file for both stable species and transition states in YAML format.¹³⁶ This file stores all relevant statistical mechanics properties parsed by Arkane and used for internal computations, as well as calculated thermodynamic properties (for stable species) and chemical identifiers. Once generated, the molecular properties file can be used in any subsequent Arkane run by referring to its path.

Automated Parameter Determination

To facilitate automatic calculations of thermodynamic properties and reaction rate coefficients and to assist users in conducting calculations in Arkane with minimal human errors, many parameters are automatically determined by Arkane. Note that parameters specified by the user in the input file will override their respective determined values. Some parameters are trivial to determine from the ESS output file, such as molecular weight, spin multiplicity, or whether the species is linear. For BAC, bond orders or bond lengths are determined from the 2D molecular structure or from the 3D coordinates, respectively. Determining the different types of symmetries, however, is less trivial, and is detailed below.

The rotational (external) symmetry number (the number of different ways the molecule can be rigidly rotated into a configuration indistinguishable from the original) as well as the optical isomers parameter (a value of 1 if no chiral centers exist, and 2 if one or more exist), are determined in Arkane using the Brute Force Symmetry Analyzer program by S. Patchkovskii.¹³⁷ This approach was adopted instead of summing the number of rotational symmetry operations,¹⁰¹ since in our experience the point groups reported by the various ESS are frequently inaccurate.

Torsional (internal) symmetry numbers for 1D rotors compensate for indistinguishable geometries due to a torsional rotation (Eq. S6), and are determined based on the individual PES torsion scan points. A torsional scan is considered symmetric by the algorithm iff

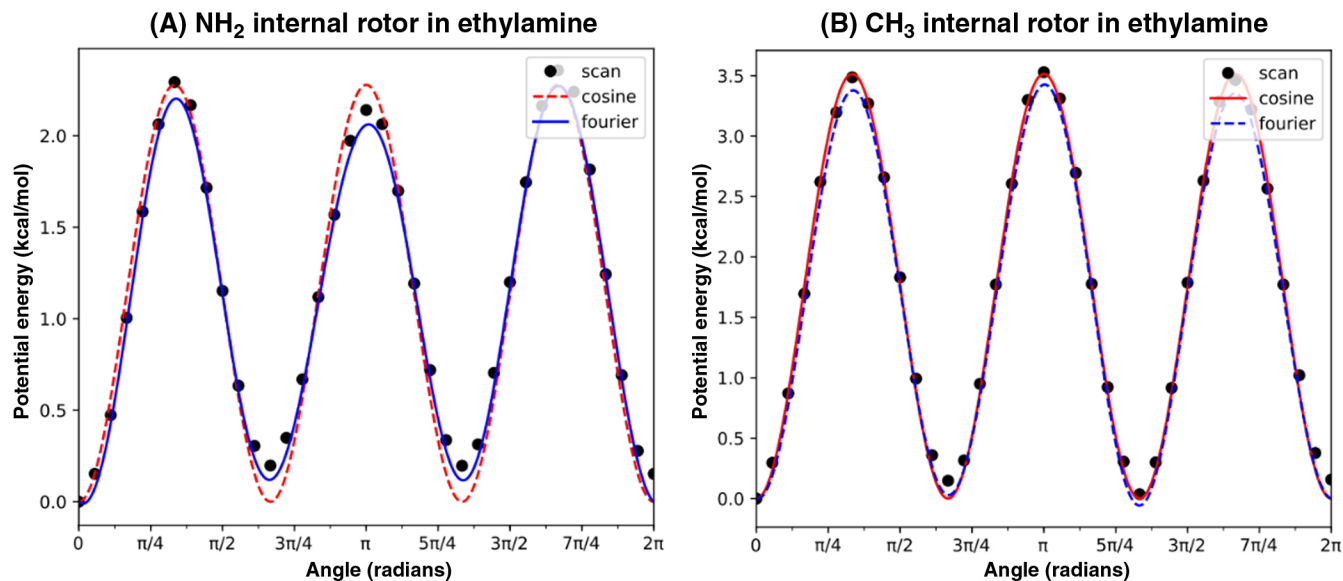


Figure 6: Two internal rotation scans for ethylamine ($\text{CH}_3\text{CH}_2\text{NH}_2$) of (A) the NH_2 rotor and (B) the CH_3 rotor taken at 10 degree increments. Computations were done at the $\omega\text{B97X-D/Def2-TZVP}$ level of theory^{138,139} using Gaussian 16.³³ Symbols represent the discrete computations, solid lines represent the Fourier series fit, dashed lines represent a cosine series fit. Plotted automatically by Arkane.

all its torsional barriers in a 2π scan are equal in height (energy) and are equally spaced. This criterion mandates identical minimal energy values of all wells. In such cases, the symmetry number is the number of identical peaks (and identical wells/valleys). Arkane first identifies the peaks and valleys in the scan, and determines the “worst energy resolution“ of the scan, R_{scan} . This variable is defined as the largest absolute energy difference between an extremum’s energy and the energy of a point adjacent to an extremum. Two points in the scan with energy difference larger than R_{scan} are determined to have different energies. Arkane uses this variable to determine whether all peaks and all valleys have the same energies, respectively. Figure 6 shows 1D scans of the two internal rotation modes in ethylamine ($\text{CH}_3\text{CH}_2\text{NH}_2$), both scans arguably seem qualitatively similar at first glance. Using the discrete calculations, Arkane automatically determined correct symmetry numbers for these modes of 1 (no symmetry) for the NH_2 rotor and a three-fold symmetry for the CH_3 rotor.

8. Summary

This work presents and documents the Arkane software, which automates the journey one has to undertake from electronic structure computations of stationary points to obtain macroscopic thermochemical parameters, high-pressure limit rate coefficients, and pressure-dependent rate coefficients. Arkane is an open-source software that is available as a root module in the RMG software suite (<https://github.com/ReactionMechanismGenerator/RMG-Py/>). Tools like Arkane advance our understanding of nonequilibrium kinetics and assist in the progress towards reliable and routine kinetics prediction in arbitrarily complex chemical systems.

Acknowledgement

This overall project was funded primarily by the Gas Phase Chemical Physics Program of the US Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences (under award number DESC0014901). A.G.D was supported by The George J. Elbaum Scholarship in Engineering, The Ed Satell Foundation, The Zuckerman STEM Leadership Program, and the Stephen and Nancy Grand Technion Energy Program (GTEP).

References

- (1) Euclides, H. O.; P. Barreto, P. R. APUAMA: A software tool for reaction rate calculations. *Journal of Molecular Modeling* **2017**, *23*, 176.
- (2) Chang, A. Y.; Bozzelli, J. W.; Dean, A. M. Kinetic Analysis of Complex Chemical Activation and Unimolecular Dissociation Reactions using QRRK Theory and the Modified Strong Collision Approximation. *Zeitschrift für Physikalische Chemie* **2000**, *214*, 1533–1568.

- (3) Dean, A. M.; Bozzelli, J. W.; Ritter, E. R. CHEMACT: A Computer Code to Estimate Rate Constants for Chemically-Activated Reactions. *Combustion Science and Technology* **1991**, *80*, 63–85.
- (4) Mokrushin, V.; Bedanov, V.; Tsang, W.; Zachariah, M. R.; Knyazev, V. D.; McGivern, W. S. ChemRate, version 1.5.10. 2011; <https://kinetics.nist.gov/ChemRate/>.
- (5) Novoselov, K. P.; Shirabaikin, D. B.; Umanskii, S. Y.; Vladimirov, A. S.; Minushev, A. K.; Korkin, A. A. CHIMERA: A software tool for reaction rate calculations and kinetics and thermodynamics analysis. *Journal of Computational Chemistry* **2002**, *23*, 1375–1389.
- (6) Dzib, E.; Cabellos, J. L.; Ortíz-Chi, F.; Pan, S.; Galano, A.; Merino, G. Eyringpy: A program for computing rate constants in the gas phase and in solution. *International Journal of Quantum Chemistry* **2019**, *119*, e25686.
- (7) Vandewiele, N. M.; Van Geem, K. M.; Reyniers, M.-F.; Marin, G. B. Genesys: Kinetic model construction using chemo-informatics. *Chem. Eng. J.* **2012**, 526–538.
- (8) KiSThelP website: Copyright. <http://kisthelp.univ-reims.fr/copyright.php>.
- (9) Canneaux, S.; Bohr, F.; Henon, E. KiSThelP: A program to predict thermodynamic properties and rate constants from quantum chemistry results. *Journal of Computational Chemistry* **2014**, *35*, 82–93.
- (10) Glowacki, D. R.; Liang, C.-H.; Morley, C.; Pilling, M. J.; Robertson, S. H. MESMER: An Open-Source Master Equation Solver for Multi-Energy Well Reactions. *The Journal of Physical Chemistry A* **2012**, *116*, 9545–9560.
- (11) Georgievskii, Y.; Miller, J. A.; Burke, M. P.; Klippenstein, S. J. Reformulation and

- Solution of the Master Equation for Multiple-Well Chemical Reactions. *The Journal of Physical Chemistry A* **2013**, *117*, 12146–12154.
- (12) Barker, J. R. Multiple-Well, multiple-path unimolecular reaction systems. I. MultiWell computer program suite. *International Journal of Chemical Kinetics* **2001**, *33*, 232–245.
- (13) McBride, B. J.; Gordon, S. Computer Program for Calculating and Fitting Thermodynamic Functions. 1992; <https://ntrs.nasa.gov/citations/19930003779>.
- (14) Lym, J.; Wittreich, G. R.; Vlachos, D. G. A Python Multiscale Thermochemistry Toolbox (pMuTT) for thermochemical and kinetic parameter estimation. *Comput. Phys. Commun.* **2019**, 106864.
- (15) The Truhlar Group Software License Form. <https://comp.chem.umn.edu/license/form-user.html>.
- (16) Zheng, J.; Bao, J. L.; Meana-Pañeda, R.; Zhang, S.; Lynch, B. J.; Corchado, J. C.; Chuang, Y.-Y.; Fast, P. L.; Hu, W.-P.; Liu, Y.-P. et al. Polyrate. 2017; <https://comp.chem.umn.edu/polyrate/>.
- (17) Miyoshi, A. SSUMES: Steady-State Unimolecular Master-Equation Solver. 2018; <http://akrmys.com/ssumes/>.
- (18) Ghysels, A.; Verstraelen, T.; Hemelsoet, K.; Waroquier, M.; Van Speybroeck, V. TAMkin: A Versatile Package for Vibrational Analysis and Chemical Kinetics. *Journal of Chemical Information and Modeling* **2010**, *50*, 1736–1750.
- (19) Duncan, W. T.; Bell, R. L.; Truong, T. N. TheRate: Program for ab initio direct dynamics calculations of thermal and vibrational-state-selected rate constants. *Journal of Computational Chemistry* **1998**, *19*, 1039–1052.

- (20) Zhang, R. M.; Xu, X.; Truhlar, D. G. TUMME: Tsinghua University Minnesota Master Equation program. *Computer Physics Communications* **2022**, *270*, 108140.
- (21) Gilbert, R. G.; Smith, S. C.; Jordan, M. J. T. UNIMOL Program Suite. 1993; <http://www.ccl.net/cca/software/SOURCES/FORTRAN/unimol/index.shtml>.
- (22) Klippenstein, S. J.; Wagner, A. F.; Dunbar, R. C.; Wardlaw, D. M.; Robertson, S. H.; Miller, J. A. VariFlex, version 1.00. 1999.
- (23) Eyring, H. The Activated Complex in Chemical Reactions. *The Journal of Chemical Physics* **1935**, *3*, 107–115.
- (24) Marcus, R. A.; Rice, O. K. The Kinetics of the Recombination of Methyl Radicals and Iodine Atoms. *Journal of Physical Chemistry* **1951**, *55*, 897–908.
- (25) Marcus, R. A. Unimolecular Dissociations and Free Radical Recombination Reactions. *The Journal of Chemical Physics* **1952**, *20*, 359–364.
- (26) Holbrook, K. A.; Pilling, M. J.; Robertson, S. H.; Robinson, P. J. *Unimolecular reactions.*, 2nd ed.; Wiley: Chichester; New York, 1996.
- (27) Eckart, C. The Penetration of a Potential Barrier by Electrons. *Physical Review* *35*, 1303–1309.
- (28) Hirschfelder, J. O.; Wigner, E. Some Quantum-Mechanical Considerations in the Theory of Reactions Involving an Activation Energy. *The Journal of Chemical Physics* *7*, 616–628.
- (29) Gonzalez-Lafont, A.; Truong, T. N.; Truhlar, D. G. Interpolated variational transition-state theory: Practical methods for estimating variational transition-state properties and tunneling contributions to chemical reaction rates from electronic structure calculations. *The Journal of Chemical Physics* **1991**, *95*, 8875–8894.

- (30) Gao, C. W.; Allen, J. W.; Green, W. H.; West, R. H. Reaction Mechanism Generator: Automatic construction of chemical kinetic mechanisms. *Computer Physics Communications* **2016**, *203*, 212–225.
- (31) Liu, M.; Grinberg Dana, A.; Johnson, M.; M.J., G.; A., J.; A.M., P.; K., G. C. H.; YeeN.W.; K., M. E. B.; R.H., W. et al. Reaction Mechanism Generator v3.0: Advances in Automatic Mechanism Generation. *Journal of Chemical Information and Modeling* **2021**,
- (32) MIT License. <https://spdx.org/licenses/MIT.html>.
- (33) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H. et al. Gaussian 16, Revision B.01. 2016.
- (34) Werner, H.; Knowles, P.; Knizia, G.; Manby, F.; Schütz, M. Molpro: A General-Purpose Quantum Chemistry Program Package. *WIREs Comput. Mol. Sci.* **2012**, *2*, 242–253.
- (35) Werner, H.; Knowles, P.; Knizia, G.; Manby, F.; Schütz, M.; Celani, P.; Györffy, W.; Kats, D.; Korona, T.; Lindh, R. et al. MOLPRO, version 2015.1, a package of ab initio programs. *University of Cardiff Chemistry Consultants (UC3): Cardiff, Wales, UK* **2015**,
- (36) ORCA. <https://orcaforum.kofo.mpg.de/>, Last accessed 22 September 2021.
- (37) Ufimtsev, I. S.; Martínez, T. J. Quantum Chemistry on Graphical Processing Units. 1. Strategies for Two-Electron Integral Evaluation. *J. Chem. Theory Comput.* **2008**, *4*, 222–231.
- (38) Ufimtsev, I. S.; Martínez, T. J. Quantum Chemistry on Graphical Processing Units.

2. Direct Self-Consistent-Field Implementation. *J. Chem. Theory Comput.* **2009**, *5*, 1004–1015.
- (39) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X. et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *An International Journal at the Interface Between Chemistry and Physics* **2015**, *113*, 184–215.
- (40) Parrish, R. M.; Burns, L. A.; Smith, D. G. A.; Simmonett, A. C.; A. Eugene De-Prince, I.; Hohenstein, E. G.; Bozkaya, U.; Sokolov, A. Y.; Remigio, R. D.; Richard, R. M. et al. Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *J. Chem. Theory Comput.* **2017**, *13*, 3185–3197.
- (41) Forst, W. Unimolecular rate theory test in thermal reactions. *The Journal of Physical Chemistry* **1972**, *76*, 342–348.
- (42) Davies, J. W.; Green, N. J. B.; Pilling, M. J. The testing of models for unimolecular decomposition via inverse laplace transformation of experimental recombination rate data. *Chemical Physics Letters* **1986**, *126*, 373–379.
- (43) Green, N. J. B.; Bhatti, Z. A. Steady-state master equation methods. *Physical Chemistry Chemical Physics* **2007**, *9*, 4275–4290.
- (44) Miller, J. A.; Klippenstein, S. J. Master Equation Methods in Gas Phase Chemical Kinetics. *J. Phys. Chem. A* **2006**, *110*, 10528–10544.
- (45) Miller, J. A.; Klippenstein, S. J. Determining phenomenological rate coefficients from a time-dependent, multiple-well master equation: “species reduction” at high temperatures. *Physical Chemistry Chemical Physics* **2013**, *15*, 4744–4753.

- (46) Johnson, M. S.; Green, W. H. Examining the Accuracy of Methods for Obtaining Pressure Dependent Rate Coefficients. *Faraday Discussions* **2022**,
- (47) Lim, K. F.; Gilbert, R. G. Modeling collisional energy transfer in highly excited molecules. *The Journal of Chemical Physics* **1990**, *92*, 1819–1830.
- (48) Van Geem, K. M.; Pyl, S. P.; Marin, G. B.; Harper, M. R.; Green, W. H. Accurate High-Temperature Reaction Networks for Alternative Fuels: Butanol Isomers. *Industrial & Engineering Chemistry Research* **2010**, *49*, 10399–10420.
- (49) Harper, M. R.; Van Geem, K. M.; Pyl, S. P.; Marin, G. B.; Green, W. H. Comprehensive reaction mechanism for n-butanol pyrolysis and combustion. *Combustion and Flame* **2011**, *158*, 16–41.
- (50) Magoon, G. R.; Aguilera-Iparraguirre, J.; Green, W. H.; Lutz, J. J.; Piecuch, P.; Wong, H.; Oluwole, O. O. Detailed Chemical Kinetic Modeling of {JP-10} (exo-Tetrahydrodicyclopentadiene) High-Temperature Oxidation: Exploring the Role of Biradical Species in Initial Decomposition steps. *International Journal of Chemical Kinetics* **2012**, *44*, 179–193.
- (51) Rosado-Reyes, C. M.; Tsang, W.; Alecu, I. M.; Merchant, S. S.; Green, W. H. Dehydration of Isobutanol and the Elimination of Water from Fuel Alcohols. *The Journal of Physical Chemistry A* **2013**, *117*, 6724–6736.
- (52) Allen, J. W.; Scheer, A. M.; Gao, C. W.; Merchant, S. S.; Vasu, S. S.; Welz, O.; Savee, J. D.; Osborn, D. L.; Lee, C.; Vranckx, S. et al. A coordinated investigation of the combustion chemistry of diisopropyl ketone, a prototype for biofuels produced by endophytic fungi. *Combustion and Flame* **2014**, *161*, 711–724.
- (53) Zhang, K.; Lokachari, N.; Ninnemann, E.; Khanniche, S.; Green, W. H.; Curran, H. J.; Vasu, S. S.; Pitz, W. J. An experimental, theoretical, and modeling study of the

- ignition behavior of cyclopentanone. *Proceedings of the Combustion Institute* **2019**, *37*, 657–665.
- (54) Jacob Goldman, M.; Yee, N.; Kroll, J.; Green, W. Pressure-dependent kinetics of peroxy radicals formed in isobutanol combustion. *Phys. Chem. Chem. Phys.* **2020**, *22*, 19802–19815.
- (55) Patwardhan, P. R.; Timko, M. T.; Class, C. a.; Bonomi, R. E.; Kida, Y.; Hernandez, H. H.; Tester, J. W.; Green, W. H. Supercritical water desulfurization of organic sulfides is consistent with free-radical kinetics. *Energy and Fuels* **2013**, *27*, 6108–6117.
- (56) Kida, Y.; Class, C. A.; Concepcion, A. J.; Timko, M. T.; Green, W. H. Combining experiment and theory to elucidate the role of supercritical water in sulfide decomposition. *Physical Chemistry Chemical Physics* **2014**, *16*, 9220–8.
- (57) Class, C. A.; Aguilera-Iparraguirre, J.; Green, W. H. A kinetic and thermochemical database for organic sulfur and oxygen compounds. *Phys. Chem. Chem. Phys.* **2015**, *17*, 13625–13639.
- (58) Class, C. A.; Liu, M.; Vandeputte, A. G.; Green, W. H. Automatic mechanism generation for pyrolysis of di-tert-butyl sulfide. *Phys. Chem. Chem. Phys.* **2016**, *18*, 21651–21658.
- (59) Class, C. A.; Vasiliou, A. K.; Kida, Y.; Timko, M. T.; Green, W. H. Detailed kinetic model for hexyl sulfide pyrolysis and its desulfurization by supercritical water. *Physical Chemistry Chemical Physics* **2019**, *21*, 10311–10324.
- (60) Slakman, B. L.; Simka, H.; Reddy, H.; West, R. H. Extending Reaction Mechanism Generator to Silicon Hydride Chemistry. *Industrial & Engineering Chemistry Research* **2016**, *55*, 12507–12515.

- (61) Grinberg Dana, A.; Buesser, B. A.; Merchant, S. S.; Green, W. H. Automated Reaction Mechanism Generation Including Nitrogen as a Heteroatom. *International Journal of Chemical Kinetics* **2018**, *50*, 243–258.
- (62) Grinberg Dana, A.; Moore, K. B.; Jasper, A. W.; Green, W. H. Large Intermediates in Hydrazine Decomposition: A Theoretical Study of the N₃H₅ and N₄H₆ Potential Energy Surfaces. *The Journal of Physical Chemistry A* **2019**, *123*, 4679–4692.
- (63) Carr, A. G.; Class, C. A.; Lai, L.; Kida, Y.; Monroe, T.; Green, W. H. Supercritical Water Treatment of Crude Oil and Hexylbenzene: An Experimental and Mechanistic Study on Alkylbenzene Decomposition. *Energy & Fuels* **2015**, *29*, 5290–5302.
- (64) Lai, L.; Gudiyella, S.; Liu, M.; Green, W. H. Chemistry of Alkylaromatics Reconsidered. *Energy & Fuels* **2018**, *32*, 5489–5500.
- (65) Lai, L.; Green, W. H. Thermochemistry and Kinetics of Intermolecular Addition of Radicals to Toluene and Alkylaromatics. *The Journal of Physical Chemistry A* **2019**, *123*, 3176–3184.
- (66) Lai, L.; Khanniche, S.; Green, W. H. Thermochemistry and Group Additivity Values for Fused Two-Ring Species and Radicals. *The Journal of Physical Chemistry A* **2019**, *123*, 3418–3428.
- (67) Long, A. E.; Merchant, S. S.; Vandeputte, A. G.; Carstensen, H.-H.; Vervust, A. J.; Marin, G. B.; Van Geem, K. M.; Green, W. H. Pressure dependent kinetic analysis of pathways to naphthalene from cyclopentadienyl recombination. *Combustion and Flame* **2018**, *187*, 247–256.
- (68) Buras, Z. J.; Dames, E. E.; Merchant, S. S.; Liu, G.; Elsamra, R. M. I.; Green, W. H. Kinetics and Products of Vinyl + 1,3-Butadiene, a Potential Route to Benzene. *The Journal of Physical Chemistry A* **2015**, *119*, 7325–38.

- (69) Buras, Z. J.; Chu, T.-C.; Jamal, A.; Yee, N. W.; Middaugh, J. E.; Green, W. H. Phenyl Radical + Propene: A Prototypical Reaction Surface for Aromatic-Catalyzed 1,2-Hydrogen-Migration and Subsequent Resonance-Stabilized Radical Formation. *Physical Chemistry Chemical Physics* **2018**,
- (70) Chu, T.-C.; Buras, Z. J.; Ołwald, P.; Liu, M.; Goldman, M. J.; Green, W. H. Modeling of aromatics formation in fuel-rich methane oxy-combustion with an automatically generated pressure-dependent mechanism. *Physical Chemistry Chemical Physics* **2019**, *21*, 813–832.
- (71) Liu, M.; Chu, T.-C.; Jocher, A.; Smith, M. C.; Lengyel, I.; Green, W. H. Predicting polycyclic aromatic hydrocarbon formation with an automatically generated mechanism for acetylene pyrolysis. *International Journal of Chemical Kinetics* **2021**, *53*, 27–42.
- (72) Behnel, S.; Bradshaw, R.; Citro, C.; Dalcin, L.; Seljebotn, D.; Smith, K. Cython: The Best of Both Worlds. *Comput. Sci. Eng.* **2011**, *13*, 31–39.
- (73) Abramowitz, A. e. a. Cython C-Extensions for Python. <https://cython.org/>, Last accessed 22 September 2021.
- (74) RMG Developers, Arkane User’s Guide. 2019; <http://reactionmechanismgenerator.github.io/RMG-Py/users/arkane/index.html>.
- (75) Petersson, G. A.; Malick, D. K.; Wilson, W. G.; Ochterski, J. W.; Montgomery, J. A.; Frisch, M. J. Calibration and Comparison of the Gaussian-2, Complete Basis Set, and Density Functional Methods for Computational Thermochemistry. *J. Chem. Phys.* **1998**, *109*, 10570–10579.
- (76) Benson, S. W.; Cruickshank, F. R.; Golden, D. M.; Haugen, G. R.; O’Neal, H. E.; Rodgers, A. S.; Shaw, R.; Walsh, R. Additivity rules for the estimation of thermochemical properties. *Chemical Reviews* **1969**, *69*, 279–324.

- (77) McBride, B. J.; Gordon, S.; Reno, M. A. Coefficients for Calculating Thermodynamic and Transport Properties of Individual Species. *NASA Technical Memorandum 4513* **1993**,
- (78) Wilhoit, R. C. Ideal Gas Thermodynamic Functions. *Thermodynamics Research Center (TRC) Current Data News* **1975**, *3*.
- (79) McQuarrie, D. A. *Statistical Mechanics*; University Science Books, 2000.
- (80) Jensen, F. *Essentials of Computational Chemistry: Theories and Models*; John Wiley & Sons, Ltd, 2007.
- (81) Cramer, C. J. *Introduction to Computational Chemistry*; John Wiley & Sons, Ltd, 2004.
- (82) Goldman, M. J.; Vandewiele, N. M.; Ono, S.; Green, W. H. Computer-generated isotope model achieves experimental accuracy of filiation for position-specific isotope analysis. *Chemical Geology* **2019**, *514*, 1–9.
- (83) Ruscic, B.; Bross, D. H. Thermochemistry. In *Mathematical Modelling of Gas-Phase Complex Reaction Systems: Pyrolysis and Combustion*; Faravelli, T., Manenti, F., Ranzi, E., Eds.; Computer-Aided Chemical Engineering; Elsevier: Cambridge, MA, 2019; Vol. 45.
- (84) Geelen, B. Accurate solution for the modified Bessel function of the first kind. *Advances in Engineering Software* **1995**, *23*, 105–109.
- (85) Allen, J. W. Dissertation: Predictive Chemical Kinetics: Enabling Automatic Mechanism Generation and Evaluation. *Massachusetts Institute of Technology* **2013**, *02139*, 59–64.
- (86) Matthew S., J.; Xiaorui, D.; Alon, G. D.; Yunsie, C.; Ryan, G.; Max, L.; Nathan, Y.;

- Katrin, B.; Emily, M.; Colin, G. et al. The RMG Database for Chemical Property Prediction. **2022**,
- (87) Ferro-Costas, D.; Cordeiro, M. N. D.; Truhlar, D. G.; Fernández-Ramos, A. Q2DTor: A program to treat torsional anharmonicity through coupled pair torsions in flexible molecules. *Computer Physics Communications* **2018**, *232*, 190–205.
- (88) Pitzer, K. S.; Gwinn, W. D. Energy Levels and Thermodynamic Functions for Molecules with Internal Rotation I. Rigid Frame with Attached Tops. *J. Chem. Phys.* **1942**, *10*, 428–440.
- (89) Ruscic, B. Uncertainty Quantification in Thermochemistry, Benchmarking Electronic Structure Computations, and Active Thermochemical Tables. *Int. J. Quantum Chem.* *114*, 1097–1101.
- (90) Ruscic, B.; Bross, D. H. Active Thermochemical Tables (ATcT) Values Based on Ver. 1.122e of the Thermochemical Network. 2019; <https://atct.anl.gov>, (accessed July 3, 2019).
- (91) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. Assessment of Gaussian-2 and Density Functional Theories for the Computation of Enthalpies of Formation. *J. Chem. Phys.* **1997**, *106*, 1063–1079.
- (92) Cox, J. D.; Wagman, D. D.; Medvedev, V. A. *CODATA Key Values for Thermodynamics*; Hemisphere: New York, 1989.
- (93) Anantharaman, B.; Melius, C. F. Bond Additivity Corrections for G3B3 and G3MP2B3 Quantum Chemistry Methods. *J. Phys. Chem. A* **2005**, *109*, 1734–1747.
- (94) Cioslowski, J.; Schimeczek, M.; Liu, G.; Stoyanov, V. A set of standard enthalpies of formation for benchmarking, calibration, and parametrization of electronic structure methods. *The Journal of Chemical Physics* **2000**, *113*, 9377–9389.

- (95) NIST Chemistry WebBook - NIST Standard Reference Database Number 69. <https://webbook.nist.gov/chemistry/>.
- (96) Pedley, J. B.; Naylor, R. D.; Kirby, S. P. Prediction of Standard Enthalpies of Formation. In *Thermochemical Data of Organic Compounds*; Springer Netherlands: Dordrecht, 1986; pp 7–51.
- (97) NIST Computational Chemistry Comparison and Benchmark Database NIST Standard Reference Database Number 101 Release 21, August 2020, Editor: Russell D. Johnson III . <https://cccbdb.nist.gov/>.
- (98) Buerger, P.; Akroyd, J.; Mosbach, S.; Kraft, M. A systematic method to estimate and validate enthalpies of formation using error-cancelling balanced reactions. *Combustion and Flame* **2018**, *187*, 105–121.
- (99) Truhlar, D. G.; Hase, W. L.; Hynes, J. T. Current Status of Transition-State Theory. *The Journal of Physical Chemistry* **1983**, *87*, 2664–2682.
- (100) Truhlar, D. G.; Garrett, B. C.; Klippenstein, S. J. Current Status of Transition-State Theory. *The Journal of Physical Chemistry* **1996**, *100*, 12771–12800.
- (101) Fernández-Ramos, A.; Ellingson, B. A.; Meana-Pañeda, R.; Marques, J. M. C.; Truhlar, D. G. Symmetry numbers and chemical reaction rates. *Theoretical Chemistry Accounts* *118*, 813–826.
- (102) Johnston, H. S.; Heicklen, J. Tunnelling Corrections for Unsymmetrical Eckart Potential Energy Barriers. *The Journal of Physical Chemistry* *66*, 532–533.
- (103) Garrett, B. C.; Truhlar, D. G. Semiclassical tunneling calculations. *The Journal of Physical Chemistry* *83*, 2921–2926.
- (104) Liu, Y. P.; Lu, D. H.; Gonzalez-Lafont, A.; Truhlar, D. G.; Garrett, B. C. Direct dynamics calculation of the kinetic isotope effect for an organic hydrogen-transfer

- reaction, including corner-cutting tunneling in 21 dimensions. *Journal of the American Chemical Society* **1993**, *115*, 7806–7817.
- (105) Fernandez-Ramos, A.; Truhlar, D. G. Improved algorithm for corner-cutting tunneling calculations. *The Journal of Chemical Physics* **2001**, *114*, 1491–1496.
- (106) Sirjean, B.; Dames, E.; Wang, H.; Tsang, W. Tunneling in Hydrogen-Transfer Isomerization of n-Alkyl Radicals. *J. Phys. Chem. A* **2011**, *116*, 319–332.
- (107) Bowman, C. T.; Hanson, R. K. Shock tube measurements of rate coefficients of elementary gas reactions. *J. Chem. Phys.* **1979**, *83*, 757–763.
- (108) Hanson, R. K.; Davidson, D. F. Recent advances in laser absorption and shock tube methods for studies of combustion chemistry. *Progress in Energy and Combustion Science* **2014**, *44*, 103–114.
- (109) Burke, M. P.; Klippenstein, S. J. Ephemeral collision complexes mediate chemically termolecular transformations that affect system chemistry. *Nature Chemistry* **2017**, *9*, 1078–1082.
- (110) Pilling, M. J.; Robertson, S. H. Master Equation Models for Chemical Reactions of Importance in Combustion. *Annu. Rev. Phys. Chem.* **2003**, *54*, 245–275.
- (111) Miller, J. A.; Klippenstein, J. K.; Robertson, S. H.; Pilling, M. J.; Green, N. J. B. Detailed balance in multiple-well chemical reactions. *Physical Chemistry Chemical Physics* **2009**, *11*, 1128–1138.
- (112) Marcus, R. A. Lifetimes of Active Molecules. II. *The Journal of Chemical Physics* **1952**, *20*, 355–359.
- (113) Wieder, M.; Marcus, R. A. Dissociation and Isomerization of Vibrationally Excited Species. II. Unimolecular Reaction Rate Theory and Its Application. *The Journal of Chemical Physics* **1962**, *37*, 1835–1852.

- (114) Marcus, R. A. Dissociation and Isomerization of Vibrationally Excited Species. III. *The Journal of Chemical Physics* **1965**, *43*, 2658–2661.
- (115) Allen, J. W.; Goldsmith, C. F.; Green, W. H. Automatic estimation of pressure-dependent rate coefficients. *Phys. Chem. Chem. Phys.* **2012**, *14*, 1131–1155.
- (116) Forst, W. *Theory of Unimolecular Reactions*; Academic Press, Inc., 1973.
- (117) Jasper, A. W.; Miller, J. A. Lennard–Jones parameters for combustion and chemical kinetics modeling from full-dimensional intermolecular potentials. *Combustion and Flame* **2014**, *161*, 101–110.
- (118) Smith, G. P.; Golden, D. M.; Frenklach, M.; Moriarty, N. W.; Eiteneer, B.; Goldenberg, M.; Bowman, C. T.; Hanson, R. K.; Song, S.; Jr., W. C. G. et al. GRI-Mech 3.0. <http://combustion.berkeley.edu/gri-mech/>, Last accessed 22 September 2021.
- (119) Glarborg, P.; Miller, J. A.; Ruscic, B.; Klippenstein, S. J. Modeling nitrogen chemistry in combustion. *Progress in Energy and Combustion Science* **2018**, *67*, 31–68.
- (120) Jasper, A. W.; Miller, J. A. OneDMin. <https://tcg.cse.anl.gov/papr/codes/onedmin.html>, 2014; Last accessed 22 September 2021.
- (121) RMG - Reaction Mechanism Generator website. <https://rmg.mit.edu/>, Last accessed 22 September 2021.
- (122) Jasper, A. W.; Klippenstein, S. J.; Harding, L. B.; Ruscic, B. Kinetics of the Reaction of Methyl Radical with Hydroxyl Radical and Methanol Decomposition. *J. Phys. Chem. A* **2007**, *111*, 3932–3950.
- (123) Greenwald, E. E.; North, S. W.; Georgievskii, Y.; Klippenstein, S. J. A Two Transition State Model for Radical-Molecule Reactions: A Case Study of the Addition of OH to C₂H₄. *J. Phys. Chem. A* **2005**, *109*, 6031–6044.

- (124) Miller, J. A.; Klippenstein, S. J. From the Multiple-Well Master Equation to Phenomenological Rate Coefficients: Reactions on a C₃H₄ Potential Energy Surface. *J. Phys. Chem. A* **2003**, *107*, 2680–2692.
- (125) Miller, J. A.; Klippenstein, S. J. The Recombination of Propargyl Radicals and Other Reactions on a C₆H₆ Potential. *J. Phys. Chem. A* **2003**, *107*, 7783–7799.
- (126) Jasper, A. W.; Hansen, N. Hydrogen-assisted isomerizations of fulvene to benzene and of larger cyclic aromatic hydrocarbons. *Proceedings of the Combustion Institute* **2013**, *34*, 279–287.
- (127) Grinberg Dana, A.; Moor, K. B.; Jasper, A. W.; Green, W. H. Large Intermediates in Hydrazine Decomposition: A Theoretical Study of the N₃H₅ and N₄H₆ Potential Energy Surfaces. *J. Phys. Chem. A* **2019**, *123*, 4679–4692.
- (128) Venkatesh, P. K.; Carr, R. W.; Cohen, M. H.; Dean, A. M. Microcanonical Transition State Theory Rate Coefficients from Thermal Rate Constants via Inverse Laplace Transformation. *J. Phys. Chem. A* **1998**, *102*, 8104–8115.
- (129) Heal, G. R. Evaluation of the integral of the Arrhenius function by a series of Chebyshev polynomials — use in the analysis of non-isothermal kinetics. *Thermochimica Acta* **1999**, *340-341*, 69–76.
- (130) Goodwin, D. G.; Speth, R. L.; Moffat, H. K.; Weber, B. W. Cantera: An object-oriented software toolkit for chemical kinetics, thermodynamics, and transport processes. <https://www.cantera.org>, Version 2.4.0. <https://www.cantera.org>, 2018; Last accessed 22 September 2021.
- (131) CHEMKIN-PRO 15131, Reaction Design, San Diego, CA. <https://www.ansys.com/>, 2013; Last accessed 22 September 2021.

- (132) Johnson, M. J. Reaction Mechanism Simulator. <https://github.com/ReactionMechanismGenerator/ReactionMechanismSimulator.jl>, 2019; Last accessed 22 September 2021.
- (133) Johnson, M. S.; Grinberg Dana, A.; Green, W. H. A Workflow for Automatic Generation and Efficient Refinement of Pressure Dependent Networks. **2022**,
- (134) Johnson, M. S.; Green, W. H. A Machine Learning Based Approach to Reaction Rate Estimation. *ChemRxiv* **2022**,
- (135) Grinberg Dana, A.; Moore, K. B.; Jasper, A. W.; Green, W. H. Large Intermediates in Hydrazine Decomposition: A Theoretical Study of the N₃H₅ and N₄H₆ Potential Energy Surfaces. *J. Phys. Chem. A* **2019**, *123*, 4679–4692.
- (136) Ben-Kiki, O.; Evans, C.; dot Net, I. YAML Ain't Markup Language (YAML™) Version 1.2, 3rd Edition. <https://yaml.org/spec/1.2/spec.html>, 2009; Last accessed 22 September 2021.
- (137) Patchkovskii, S. Brute Force Symmetry Analyzer. <https://github.com/alongd/symmetry>, 2003; Last accessed 22 September 2021.
- (138) Chai, J. D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Physical Chemistry Chemical Physics* **2008**, *10*, 6615–6620.
- (139) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics* **2005**, *7*, 3297–3305.