Computational Infrared Spectroscopy of 958 Phosphorus-bearing Molecules

Supplementary Material

Juan Camilo Zapata Trujillo[1], Anna-Maree Syme[1], Keiran N. Rowell[2], Brendan P. Burns[3,4], Ebubekir S. Clark[1], Maire N. Gorman[5], Lorrie S. D. Jacob[1], Panayioti Kapodistrias[1], David J. Kedziora[6], Felix A. R. Lempriere[1], Chris Medcraft[1], Jensen O'Sullivan[1], Evan G. Robertson[7], Georgia G. Soares[4,8], Luke Steller[4,8], Bronwyn L. Teece[4,8], Chenoa D. Tremblay[9], Clara Sousa-Silva[10], Laura K. McKemmish[1,*]

[1] *School of Chemistry, University of New South Wales, Sydney, NSW, 2052, Australia* [2] *School of Chemistry, University of Sydney, Sydney, NSW, 2052, Australia* [3] *School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, NSW, 2052, Australia* [4] *Australian Centre for Astrobiology, University of New South Wales, Sydney, NSW, 2052, Australia* [5] *Department of Physics, Aberystwyth University, Ceredigion, UK, SY23 3BZ, UK* [6] *Complex Adaptive Systems Lab, Data Science Institute, University of Technology Sydney, NSW, 2007, Australia* [7] *Department of Chemistry and Physics, La Trobe Institute for Molecular Science, La Trobe University, Victoria, 3086, Australia* [8] *School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW, 2052, Australia* [9] *CSIRO Astronomy and Space Science, Bentley WA 6102, Australia* [10] *Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, 02138, United States of America.*

[*] l.mckemmish@unsw.edu.au

Supplementary Files

- A read.me file explaining the full supplementary information contents;

- A csv file listing all molecules considered with relevant information (e.g. SMILES code, boiling point);

- A csv file with tabulated frequencies (cm^{-1}) and intensities (km $mole^{-1}$) including the empirical formula and SMILES code for each molecule, the mode to which the frequency and intensity belongs to, and the mode kind (i.e. fundamental, scaled fundamental, overtone or combination band);

- A csv file containing the force constants for fundamental frequencies for the 250 molecules with GVPT2 anharmonic data available;

- A zip file with individual folders for each molecules named by molecular formulae and SMILES codes. Within each folder there is all RASCALL, CQC-H1 and where available CQC-A1 quantum chemistry data for the molecule, along with the raw Gaussian output files and links to all other known spectral data sources.

Appendix A: Anharmonic Calculations

A1: Theory Background

Vibrational frequency and intensity calculations in quantum chemistry are mainly performed within the so-called double harmonic approximation, where the potential energy and the dipole moment are assumed to be quadratic and linear in the normal mode coordinates, respectively [54]. This

approximation is particularly appealing in quantum chemistry due to its computational ease and affordable scaling with larger systems, thus it has become a valuable tool in the interpretation of experimental vibrational spectra. However, there are two major drawbacks to the double harmonic approximation that limit its performance. Firstly, harmonic frequency calculations tend to systematically overestimate experimental frequencies and secondly, the selection rules that govern the harmonic approximation allow only the prediction of fundamental frequencies, neglecting overtones and combination bands. Multiplicative scaling factors can be applied to the calculated frequencies to match experimental values [51, 32, 40, 1, 38, ?, 29]. However, this approach only provides improvements in the frequencies' positions without considering their respective intensities, and it has no effect on the neglected overtones and combination bands.

To obtain computationally-derived vibrational spectra in better agreement with experimental data, anharmonicity must be explicitly considered in the calculations. Variational approaches like the Vibrational Self-consistent Field (VSCF) [17, 34, 20, 16, 18, 49, 43] or Vibrational Configuration Interaction (VCI) [21, 18, 52] theories can be used to this end, yet they are often limited by the molecular size and memory requirements. Instead, Vibrational Second-order Perturbation Theory (VPT2; [41]), has gained popularity in this matter as it represents a good compromise between accuracy and computational cost for medium-to-large sized molecules [10, 30, 28, 35, 9]. In the context of VPT2, the vibrational Hamiltonian is divided into unperturbed (H^0) and perturbative terms (H^1 and H^2), where the former corresponds to the common harmonic Hamiltonian and the perturbative terms incorporate third and semi-diagonal fourth derivatives to the potential energy, respectively (the second perturbative term H^2 also includes a kinetic contribution from the vibrational angular momentum) [46]. The perturbative processing of this Hamiltonian results in a handful of simple and general formulas that can be used to calculate the vibrational frequencies for fundamentals, overtones and combination bands [5, 13]. In a similar fashion, equations for the calculation of vibrational intensities are also derived under the VPT2 framework by considering both mechanical and electrical (higher-order derivatives of the dipole moment function) anharmonicity [55, 56, 6, 12, 13].

The analytical expressions for calculating both frequencies and intensities allow the simple and straightforward computation of more realistic vibrational spectra. However, a critical limitation arises when resonances or near-degenerate states appear [42, 2]. These states in most cases lead to nearly vanishing denominators in the VPT2 working equations, thus resulting in nonphysical values for the calculated frequencies and intensities. Vibrational energies are more commonly affected by type I ($\omega_i \approx 2\omega_j$) and II ($\omega_i \approx \omega_j + \omega_k$) Fermi resonances (FR), whereas vibrational intensities are plagued with both Fermi-type and the so-called Darling-Denninson resonances (DDR) ($\omega_i \approx \omega j$) [22, 12, 15]. Several approaches have been proposed to deal with resonance states and here we aim to provide a general description of those most commonly used.

The first and most common approach is the so-called deperturbed VPT2 (DVPT2) method which consists on the identification and removal of resonance states from the perturbative formulation. As the resonance terms are completely disregarded from the calculations, the DVPT2 method is unable to provide a complete picture of the vibrational nature underlying systems plagued with resonance states. This drawback is overcome by the Generalised VPT2 (GVPT2) [5, 45] method where the identified resonance states are treated separately through variational calculations and latter reintroduced as off-diagonal terms in the computations. In both cases (DVPT2 and GVPT2) the resonant terms are identified via two consecutive tests: a frequency difference threshold (Δ_ω) followed by the Martin test [39] to evaluate the deviation between the VPT2 result and a model

variational calculation (K). Taking a different approach, the calculations can also be performed under the Degeneracy-corrected Second-order perturbation theory (DCPT2) method where all possible resonant terms are replaced by non-divergent expressions [37]. This method allows to compute vibrational frequencies without further concerns for resonant terms in the perturbative formulation, but struggles when strong couplings between low- and high-frequency vibrations occur [14, 15]. As an alternative, Blonio and co-workers developed the Hybrid-degeneracy Corrected VPT2 (HDCPT2) method that mixes both standard VPT2 and DCPT2 frameworks to calculate anharmonic vibrational frequencies [14]. Using a transition function, the method is able to identify those states that would be better treated under a VPT2 formulation and, likewise, those with the DCPT2 method. These VPT2 variants (currently coded in the Gaussian package) are mostly based on the conventional Rayleigh-Schrödiner perturbation theory allowing simple algebraic equations for the anharmonic frequencies and intensities. However, alternative significant work dealing with resonance states in VPT2 has also been performed recently, considering Van Vleck perturbation theory instead [36, 48].

Despite the advantages provided by the aforementioned VPT2 flavours, there are some considerations that are worth mentioning. Though default values have been defined for the appropriate identification of resonant states in the DVPT2 and GVTP2 methods, in most cases, it is recommended to assign these values based on the specific molecular system under study. This limitation clearly hinders any high-throughput calculation of vibrational spectra as defining appropriate thresholds becomes impractical when assessing hundreds of molecules at once. Instead, one could make use of the HDCPT2 method that performs similarly to GVPT2, with the advantage of a threshold-free formulation. However, the current version of HDCPT2 only allows the calculation of vibrational frequencies and the extension to vibrational intensities is still under study [14]. We thus use GVPT2 in our calculations.

Finally, it is important to note that, due to their perturbative nature, though all VPT2 approaches generally perform well for semi-rigid molecules, there are substantial errors when dealing with large-amplitude vibrations, torsion and inversion modes, in the presence of double-well potentials and when considering floppy molecules [7, 11, 28, 46, 45].

A2: Further Computational Details

For the CQC-A1 approach, the GVPT2 method was used as it allows a general treatment of resonance states affecting both frequencies and intensities. These resonant states were identified using the following default thresholds:

$$\Delta_\omega^{1-2} = 200 \text{ cm}^{-1}; \ K^{1-2} = 1 \text{ cm}^{-1}$$
$$\Delta_\omega^{2-2} = 100 \text{ cm}^{-1}; \ K^{2-2} = 10 \text{ cm}^{-1}$$
$$\Delta_\omega^{1-1} = 100 \text{ cm}^{-1}; \ K^{1-1} = 10 \text{ cm}^{-1}$$

Where the $1-2$ superscript corresponds to Fermi-type resonances and both $2-2$ and $1-1$ represent Darling-Denninson resonances. The cubic and semidiagonal quartic derivatives of the potential were obtained by numerical differentiation of the analytic second derivatives, with the default 0.01 Å step.
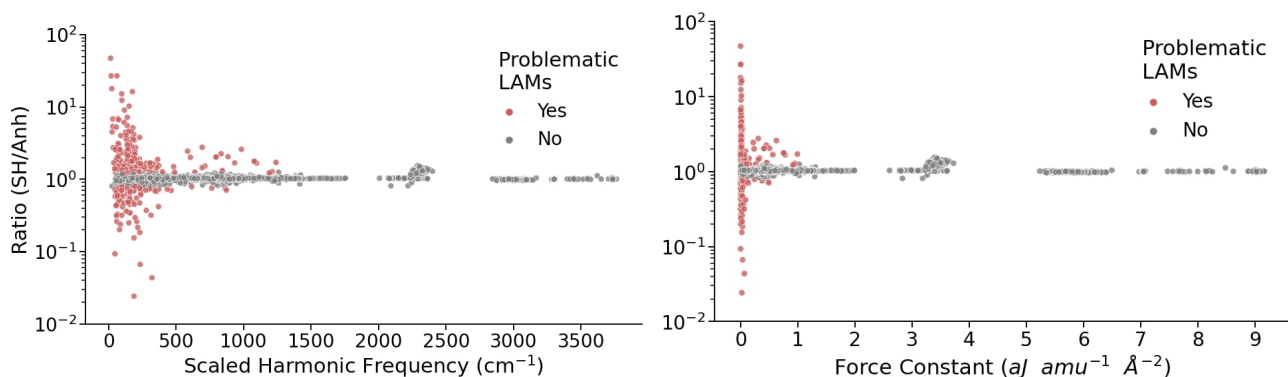
Figure S1: Ratio between the scaled harmonic (SH) and GVPT2 anharmonic (Anh) frequencies as a function of the scaled harmonic frequencies (left panel) and force constants (right panel) for 250 molecules, highlighting problematic modes due to large amplitude vibrations (LAMs) in our data. Both y axes are given in logarithmic scale.

A3: Discussion of Anomolous Anharmonic Results

Figure S1 shows the ratio between the scaled harmonic (SH) and anharmonic fundamental frequencies (Anh) as a function of frequency (left panel) and force constant (right panel). Generally, this ratio is very close to 1, i.e. the harmonic and anharmonic calculations have very similar predictions. However, within our data, there are two cases where the calculations differ substantially: (1) many low-frequency transitions with small force constants have very large differences (up to a factor of 50) between the harmonic and anharmonic frequency predictions and (2) the P-H stretch frequencies are in many molecules 200-700 cm^{-1} or more higher in the anharmonic compared to the harmonic calculations. In summary, the first issue is well-known for large amplitude vibrations (LAMs) and is an inherent limitation of perturbative approaches, while the second issue is far more concerning and can be traced back to deficiencies in the density functional.

The first issue, i.e. large scaled harmonic-anharmonic frequency (SH/Anh) differences for transitions with small force constants, is well known. The small force constants are characteristic of large amplitude vibration (LAMs) which are motions that occur along very flat areas of the potential energy surface. GVPT2 anharmonic frequencies for large amplitude vibrations (LAMs) can greatly differ from the scaled harmonic ones, due to the unsatisfactory performance of perturbation theory in these cases [7, 11, 28, 46, 45]. Under the assumption that perturbation theory fails when the correction is too large, we have chosen to flag anharmonic calculations as problematic due to LAM all frequencies with SH/Anh ratio < 0.8 and SH/Anh ratio > 1.2 that also have a small force constant < 1 aJ amu^{-1} $\mathring{A}^{-2}$.

The second issue, large SH/Anh ratios for P-H stretches, is more unusual. The prototypical example is PH$_3$, with results shown in Figure S2 for the harmonic, anharmonic, RASCALL and reference ExoMol spectra for PH$_3$. For the P-H stretch signal (near 2,360 cm^{-1}), the scaled harmonic and anharmonic calculations differ by over 1,000 cm^{-1}, with the scaled harmonic calculations seen as reliable by comparison to the ExoMol and RASCALL data sources. The highly symmetric nature of PH$_3$ causes intrinsic degeneracies that results in unreliable vibrational frequencies and intensities. Some improvement can be obtained by formally lowering the symmetry in the calculations (NoSymm option in Gaussian), as seen in Figure S2 comparing the results from the ωB97X-D - GVPT2 (black) and ωB97X-D - GVPT2 - NoSymm (green) calculations. However, unacceptably
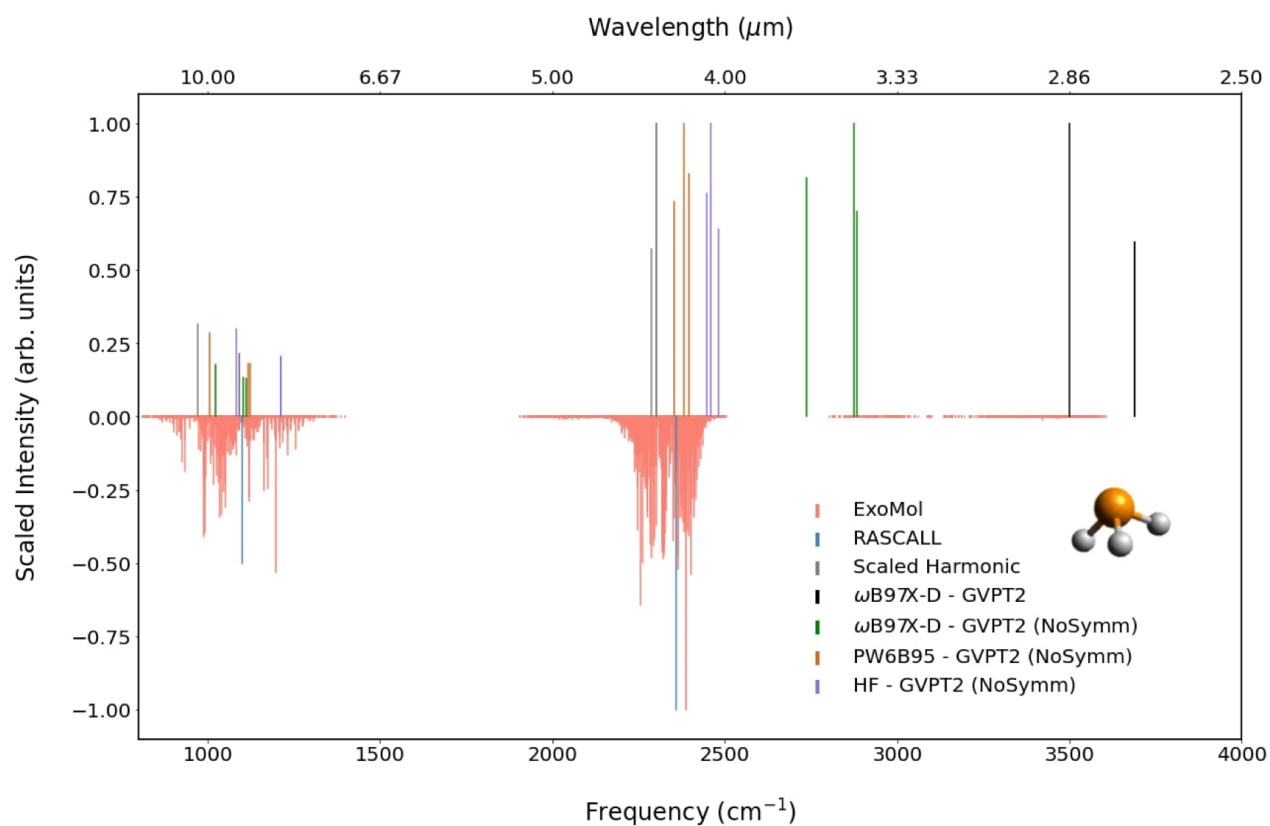
Figure S2: Comparison of the ExoMol, RASCALL, scaled harmonic and GVPT2 anharmonic frequencies for phosphine. A demonstration of the limitations present in the GVPT2 procedure for highly symmetric systems.

large errors still occur of 100s of cm^{-1}, and the harmonic calculations are far superior (compared to experimental data). Further, similar very large P-H anharmonic stretch frequencies were found in many unsymmetric molecules; for example, the CH$_5$OP isomer with SMILES code OCP, exhibited differences of 405 and 467 cm^{-1} between the scaled harmonic and anharmonic P-H stretch frequencies as calculated by ωB97X-D/def2-SVPD. Therefore, the use of NoSymm is helpful but not sufficient for resolving the issue of large SH/Anh ratios for P-H stretches.

We considered neglected resonances as a possible cause for large SH/Anh ratios for P-H stretches, but our tests showed that deuterating one of the hydrogens in the CH$_5$OP isomer (which should alter the resonance patterns) did not remove these errors.

Recent results from [8] (published subsequent to our calculations) suggested that the density functional might be the issue; this benchmark study covered ten small molecules and found that the ωB97X-D functional can yield unstable anharmonic vibrational frequencies, potentially due to an inappropriate calculation of higher-order derivatives of the potential energy. Our CH$_5$OP isomer output files showed warnings for large cubic and quartic force constants. We confirmed the culpability of the functional through tests of HF/def2-SVPD and PW6B95/def2-SVPD calculations for the PH$_3$ and CH$_5$OP isomer that did not show the very large scaled harmonic-anharmonic shifts observed with ωB97X-D. Therefore, we conclude that a different choice of functional is required for reliable anharmonic calculations, despite the very good performance of ωB97X-D for general chemistry (e.g. [27, 61]). We defer detailed consideration of alternative functionals with anharmonic methods to a future publications in order to enable detailed comparison to experiment for a large number of molecules.

However, new density functionals are unlikely to enable the accurate anharmonic treatment of large amplitude modes, for which the harmonic approximation is not a sufficiently accurate approximation for perturbative treatments to be reliable; more expensive variational approaches like VCI are likely to be required. In automated high-throughput approach, the most practical solution is probably the identification of problematic transitions through flags, as described above. We will undertake future investigations that explicitly compare predicted harmonic and anharmonic computed frequencies against experimental data for a large number of molecules. This will enable us to improve this flagging process as well as develop methods to automatically identify likely problematic molecules.

Appendix B: Reaction Network Modelling

Significant work has gone into detailed kinetic models of the chemical reaction networks in the Earth's atmosphere. GEOS-Chem [53] and the Master Chemical Mechanism (MCM) [47] are two large scale chemistry focused models, with the MCM aiming to be a 'near-explicit' mechanism: modelling 6,500+ chemical species and 17,000+ reactions. This reaction network is developed from the explicit representation of the degradation of 143 volatile organic compounds (VOCs), with pre-defined rules of chemical reactivity following an initiation reaction (e.g. oxidation, ozonolysis, photolysis), generating the vast number of reactions that need to be represented [50]. Not all species represented in the atmospheric models have well-known behaviour, and these less well understood species are propagated through the reaction network until they form end-products whose chemistry is known (e.g. CO, ROH species).

The aim of making atmospheric models comprehensive is complicated by a combinatorial explosion of possible species and minor reaction channels. To remedy the overabundance of VOCs

to be modelled, a process known as 'lumping' is used [59, 60], where only species with branching fractions >5% are explicitly represented [33], and a generic intermediate is used to represent all molecules that ultimately form similar products. A fundamental challenge of modelling the Earth's atmospheric chemistry is therefore the reduction of model complexity, so the models are computationally tractable and can be used for simulation.

On Earth, field campaigns and lab-based measurements can more easily acquire relevant experimental data, in contrast to the data needed for exoplanet atmospheres. Nevertheless, due to the diversity of VOCs known and predicted, Earth's atmospheric models are increasingly relying on theoretical methods to fill out missing experimental data. A review of how theory, and its interplay with experiment, has advanced our understanding of Earth's atmospheric reactions networks is provided in [58], with practical examples given in [57]. In Appendix C, various theoretical kinetics methods for predicting reaction rates are outlined briefly.

Exoplanetary atmospheres will be far more diverse, with hot gas giants a particular current focus as these will be the first exoplanet atmospheres to be characterised as they are easiest to observe. The reaction networks are generally more limited, but are appropriate for a wider range of physical conditions and temperatures. [19, 31] provide good introductions to exoplanetary atmosphere modelling.

Appendix C: Kinetic Rates

Prediction of molecular reaction rates is usually more challenging both experimentally and computationally than thermochemistry, with accuracies within an order of magnitude for rates generally considered very good. Molecular reaction rates can be determined using transition state theory (TST) [44]. The foundations of TST was the insight by [23] that the transition state represents a dividing surface that minimises the reactive flux between reactants and products. The transition state is a first order saddle point: the point of maximum energy (i.e. bond strain) along the minimum energy reaction pathway. Standard TST rates can be calculated for bimolecular reactions from the activation energy, and partition functions of the TS and the reactants. Unimolecular reaction rates can be calculated by Rice-Ramsperger-Kassel-Marcus (RRKM) theory, where the reaction rate at a given energy is the ratio of the density of (vibrational) states at the minimum energy well and at the TS [24]. The vibrational states are usually treated as coupled harmonic oscillators with fast intramolecular vibrational energy redistribution between them.

Where a reaction is barrierless along the minimum energy pathway, i.e. does not have an activation energy, modifications to TST are required. An example of barrierless reactions are radical-radical recombination reactions, which combine with no barrier, or conversely, dissociation of a closed-shell molecule into two radicals. These barrierless reactions are often important to photochemistry and atmospheric chemistry. In barrierless reactions, a variational approach (VTST) is used to determine a dividing surface along the reaction coordinate that minimises the reactive flux by maximising the Gibbs free energy through consideration of the entropy changes of the reactants [3]. For reactions with a "loose" transition state, the variable reaction coordinate (VRC-TST) approach can be used to determine the dividing surface more flexibly, through use of optimised pivot points between the reacting fragments [25]. Several codes are available with different methodologies for calculating barrierless reaction rate constants, including: MultiWell (VTST) [4], Polyrate (VRC-TST) [62], and MESMER where the inverse Laplace transform approach is used when high pressure limit data are available [26].

Theoretical rate constants can often deviate by one or two orders of magnitude from experiment, but provide mechanistic insight and are generally more suitable to model a particular reaction than a "surrogate" experimental reaction rate from an analogous system. Theoretical kinetics therefore provides an important pathway to supplement reaction networks with estimates of reaction rates whenever experimental data are missing or difficult to obtain.

References

[1] I. M. Alecu, Jingjing Zheng, Yan Zhao, and Donald G. Truhlar. Computational thermochemistry: Scale factor databases and scale factors for vibrational frequencies obtained from electronic model chemistries. *J. Chem. Theory Comput.*, 6(9):2872–2887, 2010.

[2] R. D. Amos, N. C. Handy, W. H. Green, D. Jayatilaka, A. Willetts, and P. Palmieri. Anharmonic vibrational properties of CH2F2: A comparison of theory and experiment. *Journal of Chemical Physics*, 95(11):8323–8336, 1991.

[3] J. L. Bao and D. G. Truhlar. Variational transition state theory: theoretical framework and recent developments. *Chem. Soc. Rev.*, 46(24):7548â7596, 2017.

[4] J. R. Barker. Multiple-Well, multiple-path unimolecular reaction systems. I. MultiWell computer program suite, volume=33, issn=0538-8066. *Int. J. Chem. Kinet.*, (4):232â245, 2001.

[5] V. Barone. Anharmonic vibrational properties by a fully automated second-order perturbative approach. *J. Chem. Phys*, 122:1–10, 2005.

[6] V. Barone, J. Bloino, C. A. Guido, and F. Lipparini. A fully automated implementation of VPT2 Infrared intensities. *Chem. Phys. Lett*, 496(1-3):157–161, 2010.

[7] Vincenzo Barone, Malgorzata Biczysko, Julien Bloino, Monika Borkowska-Panek, Ivan Carnimeo, and Paweł Panek. Toward Anharmonic Computations of Vibrational Spectra for Large Molecular Systems. *Int. J. Quantum Chem.*, 112(9):2185–2200, 2012.

[8] Vincenzo Barone, Giorgia Ceselin, Marco Fusè, and Nicola Tasinato. Accuracy Meets Interpretability for Computational Spectroscopy by Means of Hybrid and Double-Hybrid Functionals. *Front. Chem.*, 8(October):1–14, 2020.

[9] Krzysztof B. Beć and Christian W. Huck. Breakthrough potential in near-infrared spectroscopy: Spectra simulation. A review of recent developments. *Front. Chem.*, 7(FEB):1–22, 2019.

[10] Malgorzata Biczysko, Julien Bloino, Ivan Carnimeo, Paweł Panek, and Vincenzo Barone. Fully ab initio IR spectra for complex molecular systems from perturbative vibrational approaches: Glycine as a test case. *J. Mol. Struct.*, 1009:74–82, 2012.

[11] J. Bloino, A. Baiardi, and M. Biczysko. Aiming at an accurate prediction of vibrational and electronic spectra for medium-to-large molecules: An overview. *Int. J. Quantum Chem.*, 116(21):1543–1574, 2016.

[12] J. Bloino and V. Barone. A second-order perturbation theory route to vibrational averages and transition properties of molecules: General formulation and application to infrared and vibrational circular dichroism spectroscopies. *J. Chem. Phys*, 136(12), 2012.

[13] Julien Bloino. A VPT2 route to near-infrared spectroscopy: The role of mechanical and electrical anharmonicity. *J. Phys. Chem. A.*, 119(21):5269–5287, 2015.

[14] Julien Bloino, Malgorzata Biczysko, and Vincenzo Barone. General perturbative approach for spectroscopy, thermodynamics, and kinetics: Methodological background and benchmark studies. *J. Chem. Theo. Comp.*, 8(3):1015–1036, 2012.

[15] Julien Bloino, Malgorzata Biczysko, and Vincenzo Barone. Anharmonic Effects on Vibrational Spectra Intensities: Infrared, Raman, Vibrational Circular Dichroism, and Raman Optical Activity. *J. Phys. Chem. A.*, 119(49):11862–11874, 2015.

[16] M. Bounouar and Ch Scheurer. The impact of approximate VSCF schemes and curvilinear coordinates on the anharmonic vibrational frequencies of formamide and thioformamide. *J. Chem. Phys.*, 347(1-3):194–207, 2008.

[17] J. M. Bowman. The Self-Consistent-Field Approach to Polyatomic Vibrations. *Acc. Chem. Res.*, 19(7):202–208, 1986.

[18] J. M. Bowman, T.r Carrington, and H. D. Meyer. Variational quantum approaches for computing vibrational energies of polyatomic molecules. *Mol. Phys.*, 106(16-18):2145–2182, 2008.

[19] David C Catling and James F Kasting. *Atmospheric evolution on inhabited and lifeless worlds.* Cambridge University Press, 2017.

[20] G. M. Chaban, J. O. Jung, and R. Benny Gerber. Ab initio calculation of anharmonic vibrational states of polyatomic systems: Electronic structure combined with vibrational self-consistent field. *J. Chem. Phys*, 111(5):1823–1829, 1999.

[21] O. Christiansen. Vibrational structure theory: New vibrational wave function methods for calculation of anharmonic vibrational energies and vibrational contributions to molecular properties, jun 2007.

[22] Byron T. Darling and David M. Dennison. The water vapor molecule. *Phys. Rev.*, 57(2):128–139, 1940.

[23] Henry Eyring. The activated complex in chemical reactions. *J. Chem. Phys.*, 3(2):107â115, 1935.

[24] Wendell Forst. *Theory of Unimolecular Reactions.* Academic Press, London, 1973.

[25] Y. Georgievskii and S. J. Klippenstein. Variable reaction coordinate transition state theory: Analytic results and application to the C_2H_3+H → C_2H_4 reaction. *J. Chem. Phys*, 118(12):5442â5455, 2003.

[26] David R. Glowacki, Chi-Hsiu Liang, Christopher Morley, Michael J. Pilling, and Struan H. Robertson. MESMER: An Open-Source Master Equation Solver for Multi-Energy Well Reactions. *J. Phys. Chem.*, 116(38):9545–9560, September 2012.

[27] Lars Goerigk and Nisha Mehta. A trip to the density functional theory zoo: warnings and recommendations for the user. *Australian Journal of Chemistry*, 72(8):563–573, 2019.

[28] Justyna Grabska, Mirosław A. Czarnecki, Krzysztof B. Beć, and Yukihiro Ozaki. Spectroscopic and Quantum Mechanical Calculation Study of the Effect of Isotopic Substitution on NIR Spectra of Methanol. *J. Phys. Chem. A.*, 121(41):7925–7936, 2017.

[29] Magnus W.D. Hanson-Heine. Benchmarking DFT-D Dispersion Corrections for Anharmonic Vibrational Frequencies and Harmonic Scaling Factors. *J. Phys. Chem. A.*, 123(45):9800–9808, 2019.

[30] Lawrence B. Harding, Yuri Georgievskii, and Stephen J. Klippenstein. Accurate Anharmonic Zero-Point Energies for Some Combustion-Related Species from Diffusion Monte Carlo. *J. Phys. Chem. A.*, 121(22):4334–4340, 2017.

[31] Kevin Heng. *Exoplanetary atmospheres: theoretical concepts and foundations*, volume 30. Princeton University Press, 2017.

[32] Karl K. Irikura, Russell D. Johnson, and Raghu N. Kacker. Uncertainties in scaling factors for ab initio vibrational frequencies. *J. Phys. Chem. A.*, 109(37):8430–8437, 2005.

[33] Michael E. Jenkin, Sandra M. Saunders, and Michael J. Pilling. The Tropospheric Degradation of Volatile Organic Compounds: A Protocol for Mechanism Development. *Atmos. Environ.*, 31(1):81–104, 1997.

[34] J. O. Jung and R. B. Gerber. Vibrational wave functions and spectroscopy of (H2O)n, n=2,3,4,5: Vibrational self-consistent field with correlation corrections. *J. Chem. Phys*, 105(23):10332–10348, 1996.

[35] Christian G. Kirchler, Cornelia K. Pezzei, Krzysztof B. Beć, Sophia Mayr, Mika Ishigaki, Yukihiro Ozaki, and Christian W. Huck. Critical evaluation of spectral information of benchtop vs. portable near-infrared spectrometers: Quantum chemistry and two-dimensional correlation spectroscopy for a better understanding of PLS regression models of the rosmarinic acid content in Rosmarin. *Analyst*, 142(3):455–464, 2017.

[36] Sergey V. Krasnoshchekov, Elena V. Isayeva, and Nikolay F. Stepanov. Criteria for first- and second-order vibrational resonances and correct evaluation of the Darling-Dennison resonance coefficients using the canonical Van Vleck perturbation theory. *Journal of Chemical Physics*, 141(23):1–17, 2014.

[37] Kathleen M. Kuhler, Donald G. Truhlar, and Alan D. Isaacson. General method for removing resonance singularities in quantum mechanical perturbation theory. *J. Chem. Phys.*, 104(12):4664–4671, 1996.

[38] Marie L. Laury, Matthew J. Carlson, and Angela K. Wilson. Vibrational frequency scale factors for density functional theory and the polarization consistent basis sets. *J. Comp. Chem.*, 33(30):2380–2387, nov 2012.

[39] Jan M.L. Martin, Timothy J. Lee, Peter R. Taylor, and Jean Pierre François. The anharmonic force field of ethylene, C2H4, by means of accurate ab initio calculations. *J. Chem. Phys.*, 103(7):2589–2602, 1995.

[40] Jeffrey P. Merrick, Damian Moran, and Leo Radom. An evaluation of harmonic vibrational frequency scale factors. *J. Phys. Chem. A.*, 111(45):11683–11700, 2007.

[41] H. H. Nielsen. The vibration-rotation energies of molecules. *Rev. Mod. Phys.*, 23(2):90–136, 1951.

[42] Harald H Nielsen. The vibration-rotation energies of polyatomic molecules part ii. accidental degeneracies. *Physical Review*, 68(7-8):181, 1945.

[43] Paweł T. Panek, Adrian A. Hoeske, and Christoph R. Jacob. On the choice of coordinates in anharmonic theoretical vibrational spectroscopy: Harmonic vs. anharmonic coupling in vibrational configuration interaction. *J. Chem. Phys.*, 150(5), 2019.

[44] George A. Petersson. Perspective on "the activated complex in chemical reactions". *Theor. Chem. Acc.*, 103(3-4):190â195, 2000.

[45] Cristina Puzzarini, Julien Bloino, Nicola Tasinato, and Vincenzo Barone. Accuracy and Interpretability: The Devil and the Holy Grail. New Routes across Old Boundaries in Computational Spectroscopy. *Chem. Rev.*, 119(13):8131–8191, 2019.

[46] Cristina Puzzarini, Nicola Tasinato, Julien Bloino, Lorenzo Spada, and Vincenzo Barone. State-of-the-art computation of the rotational and IR spectra of the methyl-cyclopropyl cation: Hints on its detection in space. *Phys. Chem. Chem. Phys.*, 21(7):3431 – 3439, 2019.

[47] A. Rickard and J. Young. The Master Chemical Mechanism (MCM) v3.2, 2005.

[48] Andreana M. Rosnik and William F. Polik. VPT2+K spectroscopic constants and matrix elements of the transformed vibrational Hamiltonian of a polyatomic molecule with resonances using Van Vleck perturbation theory. *Molecular Physics*, 112(2):261–300, 2014.

[49] Tapta Kanchan Roy and R. Benny Gerber. Vibrational self-consistent field calculations for spectroscopy of biological molecules: New algorithmic developments and applications. *Phys. Chem. Chem. Phys.*, 15(24):9468–9492, 2013.

[50] S. M. Saunders, M. E. Jenkin, R. G. Derwent, and M. J. Pilling. Protocol for the Development of the Master Chemical Mechanism, MCM v3 (Part A): Tropospheric Degradation of Non-Aromatic Volatile Organic Compounds. *Atmos. Chem. Phys.*, 3(1):161–180, 2003.

[51] Anthony P. Scott and Leo Radom. Harmonic vibrational frequencies: An evaluation of Hartree-Fock, Møller-Plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors. *J. Phys. Chem.*, 100(41):16502–16513, 1996.

[52] Yohann Scribano, David M. Lauvergnat, and David M. Benoit. Fast vibrational configuration interaction using generalized curvilinear coordinates and self-consistent basis. *J. Chem. Phys.*, 133(9):1–13, 2010.

[53] The International GEOS-Chem User Community. GEOS-Chem 12.6.1, 2019.

[54] G. Turrell. Theory of Infrared Spectroscopy. In *Encyclopedia of Analytical Chemistry*, pages 1–32. John Wiley & Sons, Ltd, Chichester, UK, sep 2006.

[55] Juana Vázquez and John F. Stanton. Simple(r) algebraic equation for transition moments of fundamental transitions in vibrational second-order perturbation theory. *Mol. Phys.*, 104(3):377–388, 2006.

[56] Juana Vázquez and John F. Stanton. Treatment of Fermi resonance effects on transition moments in vibrational perturbation theory. *Mol. Phys.*, 105(1):101–109, 2007.

[57] L. Vereecken, D. R. Glowacki, and M. J. Pilling. Theoretical Chemical Kinetics in Tropospheric Chemistry: Methodologies and Applications. *Chem. Rev.*, 115(10):4063â4114, 2015.

[58] Luc Vereecken and Joseph S. Francisco. Theoretical studies of atmospheric reaction mechanisms in the troposphere. *Chem. Soc. Rev.*, 41(19):6259, 2012.

[59] S. W. Wang, P. G. Georgopoulos, G. Li, and H. Rabitz. Condensing complex atmospheric chemistry mechanisms. 1. The direct constrained approximate lumping (DCAL) method applied to alkane photochemistry. *Environ. Sci. Technol.*, 32(13):2018–2024, 1998.

[60] L. E. Whitehouse, A. S. Tomlin, and M. J. Pilling. Systematic reduction of complex tropospheric chemical mechanisms, Part II: Lumping using a time-scale based approach. *Atmos. Chem. Phys.*, 4(7):2057–2081, 2004.

[61] Juan Camilo Zapata and Laura K. McKemmish. Computation of Dipole Moments: A Recommendation on the Choice of the Basis Set and the Level of Theory. *J. Phys. Chem. A.*, 124(37):7538–7548, 2020.

[62] J. Zheng, J. L. Bao, R. Meana-Pañeda, S. Zhang, J. C. Lynch, B. J.and Corchado, and et al. Polyrate-version 2017-C, 2017.