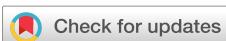


RESEARCH ARTICLE | JANUARY 05 2024

Determining internal coordinate sets for optimal representation of molecular vibration

Kemal Oenen  ; Dennis F. Dinu  ; Klaus R. Liedl 



J. Chem. Phys. 160, 014104 (2024)
<https://doi.org/10.1063/5.0180657>



View
Online



Export
Citation

CrossMark



The Journal of Chemical Physics
Special Topic: Time-resolved
Vibrational Spectroscopy

Submit Today



Determining internal coordinate sets for optimal representation of molecular vibration

Cite as: J. Chem. Phys. 160, 014104 (2024); doi: 10.1063/5.0180657

Submitted: 11 October 2023 • Accepted: 10 December 2023 •

Published Online: 5 January 2024



View Online



Export Citation



CrossMark

Kemal Oenen, Dennis F. Dinu, and Klaus R. Liedl^a

AFFILIATIONS

Department of General, Inorganic and Theoretical Chemistry, University of Innsbruck, Innrain 80, 6020 Innsbruck, Austria

^aAuthor to whom correspondence should be addressed: Klaus.Liedl@uibk.ac.at

ABSTRACT

Arising from the harmonic approximation in solving the vibrational Schrödinger equation, normal modes dissect molecular vibrations into distinct degrees of freedom. Normal modes are widely used as they give rise to descriptive vibrational notations and are convenient for expanding anharmonic potential energy surfaces as an alternative to higher-order Taylor series representations. Usually, normal modes are expressed in Cartesian coordinates, which bears drawbacks that can be overcome by switching to internal coordinates. Considering vibrational notations, normal modes with delocalized characters are difficult to denote, but internal coordinates offer a route to clearer notations. Based on the Hessian, normal mode decomposition schemes for a given set of internal coordinates can describe a normal mode by its contributions from internal coordinates. However, choosing a set of internal coordinates is not straightforward. While the Hessian provides unique sets of normal modes, various internal coordinate sets are possible for a given system. In the present work, we employ a normal mode decomposition scheme to choose an optimal set. Therefore, we screen reasonable sets based on topology and symmetry considerations and rely on a metric that minimizes coupling between internal coordinates. Ultimately, the NOMODECO toolkit presented here generates internal coordinate sets to find an optimal set for representing molecular vibrations. The resulting contribution tables can be used to clarify vibrational notations. We test our scheme on small to mid-sized molecules, showing how the space of definable internal coordinate sets can significantly be reduced.

© 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0180657>

I. INTRODUCTION

A. Coordinate systems

Considering the study of molecular vibration through quantum mechanics, the choice of coordinates determines the expressions for the kinetic energy operator (KEO) and potential energy surface (PES). Various convenient sets of coordinates exist.^{1–7} However, there is not *one* coordinate system that is optimal for describing the vibration of different molecules. An optimal coordinate system should provide a high separability of high-dimensional problems while maintaining mathematical simplicity.⁸ As the motion of nuclei is diverse and depends on the molecule size and topology, choosing such an optimal coordinate system is non-trivial.

The first coordinate systems are rooted in describing molecular vibrations based on classical mechanics, i.e., atoms vibrate as point particles connected by spring-like forces.^{2,9,10} For this *harmonic oscillator* framework, solutions are widely known,^{11–14} corresponding to the *normal modes of vibration*. This classical treatment is

usually done in mass-weighted Cartesian displacement coordinates (CDCs). In the quantum treatment, the coordinate system is transformed to normal coordinates (NCs).² Each NC is associated with a normal mode of vibration and computed from the mass-weighted CDC by a linear transformation. The nuclear Hamiltonian for a polyatomic system on an NC basis simplifies the $3N$ -dimensional Schrödinger equation to $3N$ one-dimensional quantum harmonic oscillator equations.¹⁵

Depending on whether the transformation from Cartesian coordinates (CCs) to a new set of coordinates is linear or non-linear, the resulting coordinates are rectilinear or curvilinear.¹⁶ The NCs, as mentioned above, are rectilinear. They are widely used in variational solutions of the Schrödinger equation using the Watson Hamiltonian^{3,4} in the Eckart frame,¹⁷ as implemented in many quantum-chemical algorithms.^{18–32} Nevertheless, rectilinear NCs are limited as they originate from the harmonic approximation, in which only small displacement amplitudes around the equilibrium geometry are considered. Consequently, large amplitude motions

occurring in floppy systems or systems with multiple minima pose a challenge for approaches based on NCs. Furthermore, the delocalized nature of normal modes and NCs leads to the individual vibrational motions becoming highly coupled. Hence, methods in which the PES is given, e.g., in an n -mode expansion,^{19,25,26,31,33–39} are less accurate if truncation happens at low coupling orders.¹⁶

The inherent shortcomings of conventional rectilinear NCs can be somewhat overcome by modifying them. To derive less coupled NCs, one can variationally optimize^{40–42} linear combinations of NCs or employ localization schemes.^{43–49} Localized coordinates (LCs) are used, e.g., to compute the anharmonic vibrational spectra of large polypeptides.^{50–54} Combining optimization and localization strategies provides hybrid optimized and localized coordinates (HOLCs).^{55,56}

Alternatively, one can switch to curvilinear coordinates. These inherently describe non-linear vibrational motions and, thus, can facilitate the description of the vibrational potential energy. However, the choice of curvilinear coordinates is not straightforward. Considering molecules, the most obvious choice are *internal coordinates* (ICs), which describe the internal degrees of freedom of a molecule, e.g., changes in bond lengths, bond angles, and dihedral angles.^{2,13,14} Other types of curvilinear coordinates for studying molecular vibration include, e.g., Jacobi,^{57–60} Radau,^{61–63} valence,^{64,65} hyperspherical,^{66–72} ellipsoidal,^{68,73} or polyspherical coordinates.^{16,74–85} It has further become increasingly important to combine different coordinates, e.g., in multi-layer multi-configuration time-dependent Hartree, as illustrated by combined Jacobi and Cartesian coordinates in the description of the Zundel cation.⁸⁶

Although curvilinear coordinates solve some issues rectilinear NCs pose, their applicability is limited as they complicate the expression of the nuclear KEO. The main issues arise from kinetic coupling terms, as the nuclear KEO is not separable in curvilinear coordinates, in contrast to the KEO in rectilinear NCs. These problems are overcome by numerical approaches or simplified analytical expressions for the nuclear KEO in curvilinear (internal) coordinates.^{87–90} Thus, further applications of curvilinear coordinates in computing molecular vibration can be expected. Facilitating the choice of curvilinear coordinates, hence, becomes increasingly important.

ICs are generally considered curvilinear, as they include angular coordinates for describing changes in in-plane, out-of-plane, and dihedral angles. However, ICs were initially defined as rectilinear coordinates, with the transformation between Cartesian and internal coordinates being linear.^{1,2,13,14,91} These rectilinear ICs can also be used in normal mode decomposition schemes (see Sec. III B). Depending on the application, ICs can be further redefined in multiple ways. Prominent application areas of ICs in quantum chemistry are geometry optimization and molecular vibration.

Let us first consider geometry optimizations. Here, the efficiency is influenced by the optimization algorithm,^{92–94} the approximation of the Hessian (if used),⁹⁵ and the coordinates used to describe the system.^{96,97} In the 1990s, much attention was given to choosing coordinate systems for geometry optimization. The simplest choice is Cartesian coordinates (CCs). Baker demonstrated that CCs work well if a good approximation to the Hessian is available,^{98,99} although CCs are highly coupled and include redundant translations and rotations. Various IC types have been proposed as they reduce the coupling. IC sets are non-redundant if the

number of coordinates equals the number of internal degrees of freedom, i.e., $3N - 6$ ($3N - 5$ for linear molecules) coordinates, where N is the number of atoms in the molecule.

Non-redundant ICs are, e.g., Z-matrix-type or primitive internal coordinates.¹⁰⁰ As Z-matrix coordinates describe bond lengths, bond angles, and dihedral angles, they are intuitive from a chemist's point of view. However, for geometry optimizations, Z-matrix coordinates can be problematic, especially when dealing with cyclic or larger systems, where choosing Z-matrix becomes increasingly arbitrary. A poorly constructed Z-matrix leads to strong coupling between the coordinates.^{96–98,101} Pulay and Fogarasi^{96,102} proposed the *natural internal coordinates* (NICs), which show much less coupling. These NICs comprise bonds and linear combinations of angles and torsions as deformational coordinates. NICs, however, are obtained from complicated algorithmic procedures and show problems of redundancy.¹⁰¹ Hence, Baker proposed delocalized ICs,¹⁰¹ which are very easily generated and avoid redundancy problems while still performing better than CCs. In contrast to the NICs, delocalized ICs are linear combinations of potentially all Z-matrix/primitive ICs. An extensive review of such coordinate types, including different combinations of approximated Hessian methods and optimization algorithms, was given, e.g., by Bakken and Helgaker.¹⁰³

In molecular vibration, limitations imposed by a rectilinear NC formulation can be overcome by curvilinear IC formulations. For example, in contrast to vibrational self-consistent field (VSCF) or vibrational configuration interaction (VCI) methods^{31,104–106} in NC formulation, IC formulated approaches inherently avoid high coupling between individual coordinates.^{107–109} In particular, soft torsional modes, which are highly coupled with other normal modes, lead to an incorrect wave function representation in the rectilinear NC formulation. In 2010, Suwan and Gerber⁸⁸ proposed VSCF equations in ICs, showing for a set of small molecules that torsional frequencies can be described more accurately compared to conventional VSCF in NCs. In the same year, Scribano *et al.*¹¹⁰ reported an improved description of the torsional frequency in methanol when using curvilinear NCs coupled with a numerical KEO in their VSCF/VCI scheme. Note that, in both studies, the choice of ICs is straightforward, as the molecules are comprised of a few atoms and the ICs are easily hand-picked from the (unique) Z-matrix coordinates.

In addition, Strobusch and Scheurer^{89,90} proposed a hierarchical expansion of the KEO in curvilinear coordinates and applied it to VSCF/VCI computations of H_2O_2 . Their scheme systematically analyzes which terms of the KEO in curvilinear coordinates contribute most and which degrees of approximation are needed for good accuracy. Note that, again, the choice of ICs is straightforward, as H_2O_2 has $3 \times 4 - 6 = 6$ internal degrees of freedom, and the choice of ICs naturally falls on the six primitive ICs (three bonds, two angles, and one dihedral angle). We argue that ill-defined coordinate sets in fully automated VSCF/VCI for curvilinear ICs will lead to similar issues encountered in the optimization problem. This problem has not occurred for the small systems investigated so far. The use of optimized but abstract curvilinear coordinates¹¹¹ could be one way of solving this issue.

In the present work, we propose a systematic sampling of IC sets based on a normal mode decomposition scheme, defining the best one to deliver the least coupling in the representation of normal

modes in internal coordinates. Therefore, we introduce the concept of normal mode decomposition and its application to quantifying molecular vibration.

B. Normal mode decomposition

In infrared and Raman spectroscopy, it is established to assign the experimentally observed frequencies to characteristic patterns of molecular vibration. One usually adopts notations comprising stretching, bending, and deformations. Although attempts to conventionalize these notations are quite old,^{112–114} there is no established convention to date. These notations predominantly rely on NCs from the harmonic oscillator approximation.² While it is straightforward to denote stretching and bending vibrations, notations for delocalized deformations in the fingerprint region are cumbersome as they involve the changes of multiple spatially distant atoms.

Initially, the classification of vibrational motions was achieved by determining the relative amplitudes of symmetry (RAS) coordinates. Based on earlier studies by Torkington¹¹⁵ and Thomas,¹¹⁶ the use of potential energy distribution (PED) as a more fundamental way of classifying molecular vibration was first discussed in 1952 by Morino and Kuchitsu.¹¹⁷ As PED turned out to be more reliable than RAS,¹¹⁷ it was adopted and modified later on in Refs. 118 and 119. Already in 1954, Taylor¹²⁰ pointed out that only considering the contribution of ICs to potential energy is incomplete. He proposed an analogous formalism for the kinetic energy distribution (KED) and suggested that combining PED and KED would determine the characteristic IC for a specific normal mode. Only in 1974 did Rytter¹²¹ describe the total energy distribution (TED) to determine the contributions of internal (or symmetry²) coordinates to the total energy of a given mode. The original TED definition was somewhat erroneous¹²² and, thus, extensively redefined by Alix and Müller.¹²³ Since the 1970s, multiple studies have considered the topic of vibrational energy distributions (VEDs), the inherent properties of the energy distribution matrices,^{124–126} as well as applications in assigning vibrational notations.^{127–129}

We here want to focus on some key aspects of VED.¹³⁰ One can define PED and KED matrices for each normal mode, which can be averaged to obtain a TED matrix. Furthermore, to avoid having two (or three) matrices per mode, simplification procedures enable the reduction to a single matrix for all modes together. Depending on the simplification approach, the PED/KED/TED matrices are summarized with or without the inclusion of the off-diagonal elements.^{123,124,130} We describe these approaches more precisely in Sec. III. At this point, it is noteworthy that when including the off-diagonal elements, the PED/KED/TED matrices become equivalent, leading to a single VED matrix, which is conveniently still called the PED matrix. In the alternative approach, this equivalence is not given. As Whitmer¹³⁰ pointed out, choosing which of the three matrices to use in classifying molecular vibration depends on the coordinate system. PED is preferable when using, e.g., ICs, and KED is preferable when using, e.g., CCs. In the former case, the force constant matrix \mathbf{F} becomes nearly diagonal, and in the latter case, the kinetic energy matrix \mathbf{G}^{-1} becomes nearly diagonal, both making the simplification methods, as mentioned above, nearly exact. In summary, although the distinction between PED, KED, and TED is important from a theoretical point of view, our and

other similar methods generally use PED, as it is either equivalent to the other matrices or is the more preferred one.

Related to VED is the concept of *intrinsic frequencies*, i.e., the frequency of a hypothetical vibrational motion solely described by a particular IC. While the first ideas for intrinsic frequencies were already proposed, e.g., in Taylor's work¹²⁰ in the 1950s, it was in 1989 that Boatz and Gordon¹³¹ proposed a scheme to compute intrinsic frequencies from PED matrix elements. They defined the intrinsic frequency as the sum of the contributions of all normal modes to a particular IC. They utilized these in characterizing bonds and illustrating their benefits in providing a different perspective on normal mode analysis compared to the PED matrix.¹³² The Boatz–Gordon approach was initially implemented in GAMESS. At present, there is, to the best of our knowledge, no other implementation available. In the present work, we adopt the Boatz–Gordon approach in a modified way (see Sec. III) and implement it as standalone software called NOMODECO, which is freely available on GitHub.¹³³

At present, the most versatile software tool for PED analysis is the so-called vibrational energy distribution analysis (VEDA).¹³⁴ Despite its limitation to GAUSSIAN outputs and the WINDOWS operating system, VEDA has the unique advantage that it finds an optimal set of ICs through an optimization procedure. In contrast, the PED analyses mentioned above are based on hand-picked IC sets. The results obtained from PED depend on the chosen IC set,^{134,135} and there are empirical selection rules for a good IC set, e.g., described in the original Boatz–Gordon approach.¹³¹ However, these rules are insufficient for selecting the best IC set in an automatic procedure. In this respect, we generate many sets of ICs to find “the best,” based on a geometry-based initial guess.¹³⁶ Thereby, we provide a set of ICs that best decomposes the normal modes, which facilitates the definition of reasonable notations even for delocalized vibrations. A similar strategy is followed in the VEDA software. However, we stick solely to primitive ICs (see Sec. III B) and avoid linear combinations or mixed coordinates to maintain the intuitive interpretation of PED matrices and intrinsic frequencies. Note that in notations of vibrational transitions from anharmonic approaches, normal modes heavily mix, and hence, the focus will shift from computing PED matrices to determining dominant harmonic modes.¹³⁷

II. METHODOLOGY

Figure 1 depicts the program structure of the NOMODECO toolkit to automatically perform a normal mode decomposition for a given molecular geometry and a corresponding Hessian. The program extracts the optimized geometry and Hessian from a MOLPRO output file¹³⁸ (interfaces to other software packages such as GAUSSIAN¹³⁹ or ORCA¹⁴⁰ are available). After determining the molecular point group and several geometric properties, e.g., the number of rings in a system and the identification of planar and linear units, the program generates all primitive ICs and all reasonable IC sets for the molecule. As the number of possible IC sets grows rapidly with the system size, the program performs a screening based on symmetry and topology criteria. The latter strategy is based on prescriptions given by Decius in 1949¹³⁶ (see Sec. III B), yielding which and how many ICs to choose for a complete and non-redundant set.

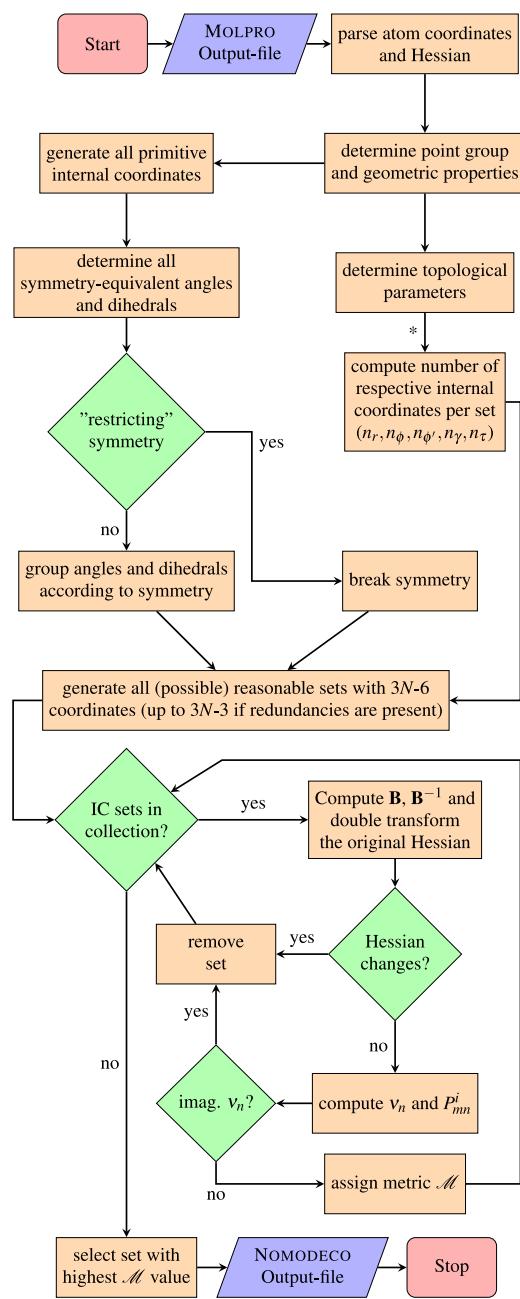


FIG. 1. Flowchart of NOMODECO. The procedure (*) for determining the number of ICs per set is explained in Sec. III B.

To obtain the same intrinsic frequency for symmetry-equivalent ICs,¹³¹ all ICs must be grouped to combine only symmetry-equivalent ones in an IC set. In some cases (such as CH₄), redundant ICs are included to satisfy this symmetry condition. The NOMODECO toolkit generates sets with redundant ICs by allowing more angle and dihedral coordinates than the required number given by the aforementioned topological prescription. In

some cases, the symmetry is somewhat “restricting” (such as SF₆), and the inclusion of all symmetry-equivalent ICs of a type leads to an over-determined and, hence, not complete set. Here, symmetry is either reduced or switched off, and only the topological criteria are considered. Finally, the program computes for all constructed IC sets the matrices and quantities introduced in Sec. III A. With this, it skips IC sets that are incomplete or yield unreasonable intrinsic frequencies. The other well-defined IC sets are assigned a metric value/score (see Sec. III C). The IC set with the highest score is considered the best set. With the results in Sec. IV, we elaborate on further details not covered in Fig. 1. Currently, the toolkit has been tested for molecules with up to 21 atoms.

Geometry optimization and harmonic frequency calculations to obtain the Hessian were performed using the MOLPRO software package.¹³⁸ All calculations were performed at the HF/6-31G(d,p) level of theory to keep computational costs low, as electron correlation is not important here. For the Hessian, we rely on analytical differentiation in the HF method to avoid errors one would have from numerical differentiation.

III. THEORY

Section III A introduces the most relevant definitions for PED and the modified equations in the presence of redundancies. The notation is largely taken from Whitmer¹³⁰ as well as Boatz and Gordon.¹³¹ Section III B illustrates the strategy for determining the number of ICs for a given molecular topology. Finally, the metric used in determining the best internal coordinate is outlined in Sec. III C.

A. Decomposition of normal modes

1. The **B**-matrix

A molecule with N atoms has $3N$ Cartesian coordinates \vec{X} or $S = 3N - 6$ ($S = 3N - 5$ for linear molecules) internal coordinates \vec{R} ,

$$\vec{X}^\top = (x_1, y_1, z_1, \dots, x_N, y_N, z_N), \quad (1)$$

$$\vec{R}^\top = (r_1, r_2, \dots, r_S). \quad (2)$$

These coordinate systems can be transformed into each other by the Wilson **B**-matrix ($S \times 3N$),²

$$\vec{R} = \mathbf{B} \cdot \vec{X}, \quad (3)$$

assuming a linear relationship between changes in Cartesian and internal coordinates.

2. Force constant matrices

The potential energy of the molecule can be approximated by quadratic force constants, which are second derivatives of the energy E with respect to Cartesian or internal coordinates,

$$f_{ij} = \frac{\partial^2 E}{\partial x_i \partial x_j}, \quad (4)$$

$$F_{ij} = \frac{\partial^2 E}{\partial r_i \partial r_j}. \quad (5)$$

These expressions are matrix elements of the Cartesian force constant matrix \mathbf{f} and the internal force constant matrix \mathbf{F} . These force constant matrices can be transformed into each other,

$$\mathbf{F} = (\tilde{\mathbf{B}}^{-1})^\top \mathbf{f} \tilde{\mathbf{B}}^{-1}, \quad (6)$$

$$\mathbf{f} = \tilde{\mathbf{B}}^\top \mathbf{F} \tilde{\mathbf{B}} \quad (7)$$

by the $\tilde{\mathbf{B}}$ -matrix, which is the augmented \mathbf{B} -matrix generated by appending additional rows to make it a quadratic matrix ($3N \times 3N$). The additional rows represent rotational and translational motions (six for non-linear molecules or five for linear molecules). Within Eckart conditions, these motions are invariant to the potential energy. Practically, these rows correspond to the normal modes (see Sec. III A 4) with low or zero eigenvalues.¹⁴¹ For clarity, we distinguish the notation of augmented or not-augmented only when the computation of the matrix inverse is explicitly needed.

3. The kinetic energy matrix and treatment of redundant internal coordinates

The inverse of $\tilde{\mathbf{B}}$ can be calculated from the inverse kinetic energy matrix \mathbf{G} .² They are related by

$$\mathbf{G} = \tilde{\mathbf{B}} \mathbf{m}^{-1} \tilde{\mathbf{B}}^\top, \quad (8)$$

where \mathbf{m}^{-1} ($3N \times 3N$) is the reciprocal mass matrix. For non-redundant IC sets with $S = 3N - 6$, the \mathbf{G} matrix is real and symmetric [$(S + 6) \times (S + 6)$]. From the diagonalization,

$$\mathbf{GK} = \mathbf{Ke}, \quad (9)$$

where \mathbf{K} and \mathbf{e} denote the eigenpairs, the inverse of \mathbf{G} can be readily computed as

$$\mathbf{G}^{-1} = \mathbf{Ke}^{-1}\mathbf{K}^\top, \quad (10)$$

or in scalar notation,

$$(G^{-1})_{ij} = \sum_n \frac{K_{ni} K_{nj}}{e_n}. \quad (11)$$

If the IC set includes m redundancies, the \mathbf{G} matrix becomes singular. One can then compute the inverse by neglecting the m sum elements in Eq. (11) where the eigenvalue e_n is zero. Finally, by taking the inverse of (8), one can rearrange to obtain the expression to compute $\tilde{\mathbf{B}}^{-1}$,

$$\mathbf{G}^{-1} = (\tilde{\mathbf{B}}^{-1})^\top \mathbf{m} \tilde{\mathbf{B}}^{-1}, \quad (12)$$

$$\tilde{\mathbf{B}}^{-1} = \mathbf{m}^{-1} \tilde{\mathbf{B}}^\top \mathbf{G}^{-1}. \quad (13)$$

4. Normal modes

The mass-weighted Cartesian force constant matrix

$$\mathbf{f}' = (\mathbf{M}^{-1/2})^\top \mathbf{f} \mathbf{M}^{-1/2} \quad (14)$$

can be diagonalized as

$$\Lambda = \mathbf{L}^\top \mathbf{f}' \mathbf{L}, \quad (15)$$

where the Λ -matrix is the diagonal and contains the eigenvalues λ_i and the \mathbf{L} -matrix ($3N \times 3N$) contains the eigenvectors usually denoted as *normal modes*. The diagonal matrix $\mathbf{M}^{-1/2}$ contains the reciprocal square roots of the atomic masses. One can rewrite Eq. (15) in terms of the (not mass-weighted) Cartesian force constant matrix

$$\Lambda = \mathbf{I}^\top \mathbf{f} \mathbf{l}, \quad (16)$$

where we have used the definition

$$\mathbf{l} = \mathbf{M}^{-1/2} \mathbf{L} \quad (17)$$

The \mathbf{l} -matrix contains eigenvectors that can be denoted as mass-weighted normal modes in Cartesian coordinates. Now, we can use the \mathbf{B} -matrix to transform the Cartesian force constant matrix \mathbf{f} to the internal force constant matrix \mathbf{F} and perform an analogous diagonalization,

$$\Lambda = \mathbf{I}^\top \mathbf{f} \mathbf{l} = \mathbf{I}^\top (\mathbf{B}^\top \mathbf{F} \mathbf{B}) \mathbf{l} \quad (18)$$

By multiplying the \mathbf{B} -matrix by the \mathbf{l} -matrix, the mass-weighted normal modes in Cartesian coordinates transform into the \mathbf{D} -matrix ($S \times 3N$),

$$\mathbf{D} = \mathbf{Bl}, \quad (19)$$

which contains eigenvectors that describe the mass-weighted normal modes in internal coordinates. The diagonalization in Eq. (18) using the \mathbf{D} -matrix reads as

$$\Lambda = \mathbf{D}^\top \mathbf{FD}. \quad (20)$$

One can now verify from Eq. (19) that the \mathbf{D} -matrix is identical to the transformation matrix between internal and normal coordinates. Note that \mathbf{D} equals the \mathbf{L} -matrix encountered in studies dealing with PED analysis.^{117,120,121,130,134} The latter should not be confused with \mathbf{L} in Eq. (15). When applying the strategy outlined in Sec. III A 3 to compute \mathbf{G}^{-1} (and, hence, $\tilde{\mathbf{B}}^{-1}$) in the presence of redundancies, the \mathbf{D} -matrix can be computed with the equations above. Alternatively, one can also use the definition provided by Brunvoll *et al.*¹³⁵ to compute a unique representation of \mathbf{D} (denoted as $\hat{\mathbf{D}}$) in the presence of redundant coordinates,

$$\hat{\mathbf{D}} = \tilde{\mathbf{B}} \mathbf{m}^{-1} \mathbf{B}^\top (\mathbf{D}^{-1})^\top. \quad (21)$$

Here, $\tilde{\mathbf{B}}$ corresponds to the \mathbf{B} -matrix when using redundant internal coordinates, while \mathbf{B} and \mathbf{D} correspond to the matrices for non-redundant internal coordinates. The definition (21) is not dependent on the particular choice of the complete set of independent internal coordinates. Using Eq. (21) has the drawback that whenever redundancies are present, one needs to construct additional matrices for a complete set of non-redundant internal coordinates. This can be avoided using the method in Sec. III A 3.

5. Potential, kinetic, and total energy distribution matrices

As shown by several authors,^{2,117,120,123,130,134} the harmonic vibrational potential energy V in terms of internal coordinates is

$$V = \frac{1}{2} \sum_m \sum_n F_{mn} R_m R_n. \quad (22)$$

Analogously, one can also define the vibrational kinetic energy T in internal coordinates,

$$T = \frac{1}{2} \sum_m \sum_n (G^{-1})_{mn} \dot{R}_m \dot{R}_n. \quad (23)$$

As established in Sec. III A 4, the transformation between internal coordinates \vec{R} and normal coordinates (from here on denoted as \vec{Q}) is given by the \mathbf{D} -matrix,

$$\vec{R} = \mathbf{D} \vec{Q}. \quad (24)$$

Inserting Eq. (24) into Eqs. (22) and (23) yields

$$V = \frac{1}{2} \sum_i Q_i^2 \sum_m \sum_n D_{ni} D_{mi} F_{mn}, \quad (25)$$

$$T = \frac{1}{2} \sum_i Q_i^2 \sum_m \sum_n D_{ni} D_{mi} (G^{-1})_{mn}. \quad (26)$$

The sums in Eqs. (25) and (26) illustrate how the contributions to the potential and kinetic energy can be decomposed into individual contributions per mode. Hence, one can formulate potential and kinetic energy for a system vibrating in the i th normal mode, Q_i as

$$V^i = \frac{1}{2} Q_i^2 \sum_m \sum_n D_{ni} D_{mi} F_{mn}, \quad (27)$$

$$T^i = \frac{1}{2} \dot{Q}_i^2 \sum_m \sum_n D_{ni} D_{mi} (G^{-1})_{mn}. \quad (28)$$

The individual sum terms in Eqs. (27) and (28) are the elements of the PED and KED matrix. Since

$$\lambda_i = \sum_m \sum_n D_{mi} D_{ni} F_{mn} \quad i \in [1, 3N], \quad (29)$$

$$1 = \sum_m \sum_n D_{mi} D_{ni} (G^{-1})_{mn} \quad i \in [1, 3N], \quad (30)$$

the contributions (normalized to unity) of each term to the potential or kinetic energy are given by the following expressions:

$$V_{mn}^i = \frac{D_{mi} F_{mn} D_{ni}}{\lambda_i}, \quad (31)$$

$$T_{mn}^i = D_{mi} (G^{-1})_{mn} D_{ni}. \quad (32)$$

Hence, for one mode i , a pair of matrices can be computed. The non-diagonal elements are coupling terms, while the diagonal elements are the contribution of the internal coordinate to the respective vibrational energy. Finally, one can introduce the TED matrix, whose elements are the average of the corresponding PED and KED terms,

$$E_{mn}^i = \frac{1}{2} [V_{mn}^i + T_{mn}^i]. \quad (33)$$

The vibrational density matrix P_{mn}^i used by Boatz and Gordon¹³¹ equals the PED elements, as one can readily see from their definition,

$$P_{mn}^i = \frac{D_{mi} F_{mn} D_{ni}}{\lambda_i}. \quad (34)$$

From now on, we will use P_{mn}^i to denote elements of the PED matrix while maintaining the notations for the KED and TED matrix as previously introduced.

6. Vibrational energy distribution matrix

One can simplify the complete energy distributions via two approaches (cf. Sec. I B) to obtain one matrix for all modes instead of two (or potentially three) matrices for each mode separately. The first approach involves summing up all elements in a row (or column, as the complete PED, KED, and TED are symmetric¹³⁰), resulting in the set of i matrices becoming a single $i \times i$ matrix. The elements of these simplified energy distributions are, hence, given as

$$P_m^i = \sum_n P_{mn}^i = \sum_n \frac{D_{mi} F_{mn} D_{ni}}{\lambda_i}, \quad (35)$$

$$T_m^i = \sum_n T_{mn}^i = \sum_n D_{mi} (G^{-1})_{mn} D_{ni}, \quad (36)$$

$$E_m^i = \sum_n E_{mn}^i. \quad (37)$$

Note that expressions (35)–(37) are normalized. It has been shown^{119,130} that in this simplified form, matrices (35)–(37) are numerically identical and can be computed as

$$P_m^i = T_m^i = E_m^i = D_{ik} (D^{-1})_{ki}. \quad (38)$$

As PED, KED, and TED are identical in this approximation, we simply term Eq. (38) as a *vibrational energy distribution* (VED) matrix, where the columns are the normal modes and the rows are the internal coordinates. This approximation is considered more exact than the strategy shown below, as neither diagonal nor off-diagonal elements are neglected in the summation. However, one cannot identify the specific contributions of diagonal and off-diagonal elements as they are combined in the new matrix elements.

The other approach, which is employed, e.g., by the VEDA program,¹³⁴ neglects the off-diagonal elements. One can again reduce one PED/KED/TED matrix to a single column,

$$[P_m^i]_{\text{approx.}} = \frac{F_{mm} D_{mi}^2 \lambda_i^{-1}}{\sum_n F_{nn} D_{ni}^2 \lambda_i^{-1}} = \frac{F_{mm} D_{mi}^2}{\sum_n F_{nn} D_{ni}^2}, \quad (39)$$

$$[T_m^i]_{\text{approx.}} = \frac{(G^{-1})_{mm} D_{mi}^2}{\sum_n (G^{-1})_{nn} D_{ni}^2}, \quad (40)$$

$$[E_m^i]_{\text{approx.}} = \frac{[F_{mm} \lambda_i^{-1} + (G^{-1})_{mm}] D_{mi}^2}{\sum_n [F_{nn} \lambda_i^{-1} + (G^{-1})_{nn}] D_{ni}^2}. \quad (41)$$

Note that the implicit normalization in (31)–(33) and (35)–(37) is no longer valid here. Hence, we divide through the appropriate sums in (39)–(41). When using internal coordinates, the neglect of the off-diagonal elements is only valid for PED but not for KED and TED.¹³⁰ Consequently, the approximate matrices are not identical, which presents a disadvantage. On the other hand, one gains the advantage of easily formulating a relative contribution, as the diagonal elements are always non-negative.¹²⁰ Since the choice of either of the aforementioned definitions depends on the user's interest, the NOMODECO toolkit can calculate all of them.

7. Intrinsic vibrational frequencies

Finally, we define the intrinsic vibrational frequencies introduced by Boatz and Gordon¹³¹ as

$$\nu_n = \sum_m \sum_i P_{mn}^i \lambda_i. \quad (42)$$

One can readily see in this definition that an intrinsic frequency equals the contribution of all normal modes i to a particular internal coordinate. If we substitute Eq. (34) in Eq. (42), we further obtain

$$\nu_n = \sum_m \sum_i \frac{D_{mi} F_{mn} D_{ni}}{\lambda_i} \lambda_i = \sum_m \sum_i D_{mi} F_{mn} D_{ni}. \quad (43)$$

Intrinsic frequencies represent the characteristic frequency of a mode comprised of the motion of a single IC. In other words, the more a mode can be characterized by a single IC, the closer its harmonic frequency gets to its intrinsic frequency.

B. Selection of internal coordinates

The here used ICs (cf. Sec. III) are known as *primitive internal coordinates*,²⁹ i.e., geometrical parameters derived from interatomic distances and angles between chemical bonds. Figure 2 shows the five most frequently used IC types for molecular vibration: bonds, (linear) angles, out-of-plane angles, and proper dihedrals. These five coordinates have all been defined by Wilson *et al.*² and are, therefore, sometimes called, in the literature, *Wilson-type internal coordinates*. The entries for the transformation matrix, i.e., the B-matrix, have appeared elsewhere already^{2,9,91} and will hence not be shown here. It should be pointed out that, in general, other IC types can be defined,^{142–145} as long as they fulfill the Eckart conditions.¹⁷ Such definitions are often only useful for specific vibrational problems, while Wilson-type ICs can be applied to a general set of problems in molecular vibration.

Determining the number of IC types per set can be aided by investigating the molecule's topology. Once the ICs are defined for a system, two major questions arise: (1) Can one find a set that describes all vibrational degrees of freedom, i.e., a “complete” set? (2) Can this be done without introducing redundancies in a set? Already in 1949, Decius¹³⁶ discussed these problems, providing a scheme that yields the number of a specific IC type for a complete set. It should be mentioned that redundancies can occur for cyclic (sub)systems. However, Decius also suggests how to remove these redundancies if

needed (cf. Sec. IV D). Furthermore, redundancies can also be introduced to exploit symmetry (cf. Sec. II). The prescriptions given by Decius¹³⁶ are summarized in Fig. 3.

C. Metric for determining the quality of an internal coordinate set

We introduce a metric to distinguish between the generated complete IC sets. This metric depends on the application area of the chosen coordinate set. Here, we want to provide for the least coupling between the ICs and maximize the contribution of an IC to a normal mode, providing the clearest possible description of this mode in terms of ICs. The same properties are also essential for VSCF/VCI schemes in curvilinear ICs, where one wants to minimize the coupling between the coordinates for a better separability of the wavefunction ansatz and a more beneficial convergence of the PES. Hence, the metric \mathcal{M} shown below may also be applicable for such applications, although modifications need to be considered (cf. Sec. V).

As the contributions of ICs to the normal modes are given by the PED matrices (cf. Sec. III A), we can use the information of the PED matrices for computing our metric \mathcal{M} . Let the matrix A be either the VED matrix P_m^i or the normalized diagonal elements of the PED matrix $[P_m^i]_{\text{approx}}$. Furthermore, let a_{ij} denote the individual matrix elements. We define the metric \mathcal{M} as

$$\mathcal{M} = \frac{\sum_{i=1}^S \max_{j=1,\dots,3N-6(5)} a_{ij}}{S}. \quad (44)$$

Note that the number of rows i and columns j are the indices over the number of ICs and the number of normal modes, respectively. Depending on A , \mathcal{M} can take values of 0–1 for $A = P_m^i$ or 0–100 for $A = [P_m^i]_{\text{approx}}$. Finally, for specific cases, penalties \mathcal{P}_i can be introduced to modify the metric given above. As they are irrelevant for the following, we will only comment on them in the supplementary material.

IV. RESULTS AND DISCUSSION

For different topological molecules, we demonstrate the performance of the NOMODECO toolkit in selecting the best-suited IC set. We illustrate the increase in IC sets with molecule size and show how symmetry and topology considerations significantly reduce the

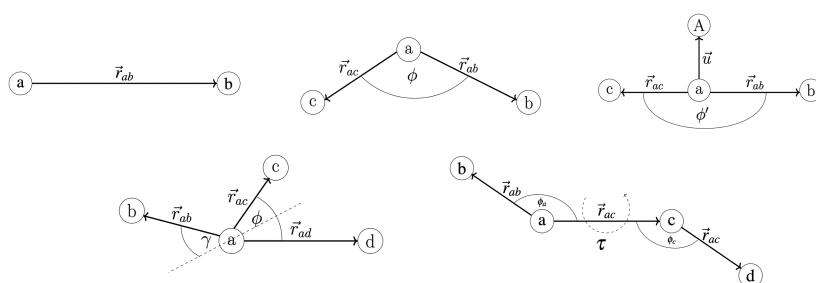


FIG. 2. All coordinate sets generated in the NOMODECO toolkit rely on primitive internal coordinates, namely bonds \vec{r}_{ab} , (in-plane) bond angles ϕ (linear angles are denoted as ϕ'), out-of-plane angles γ , and proper dihedrals τ .

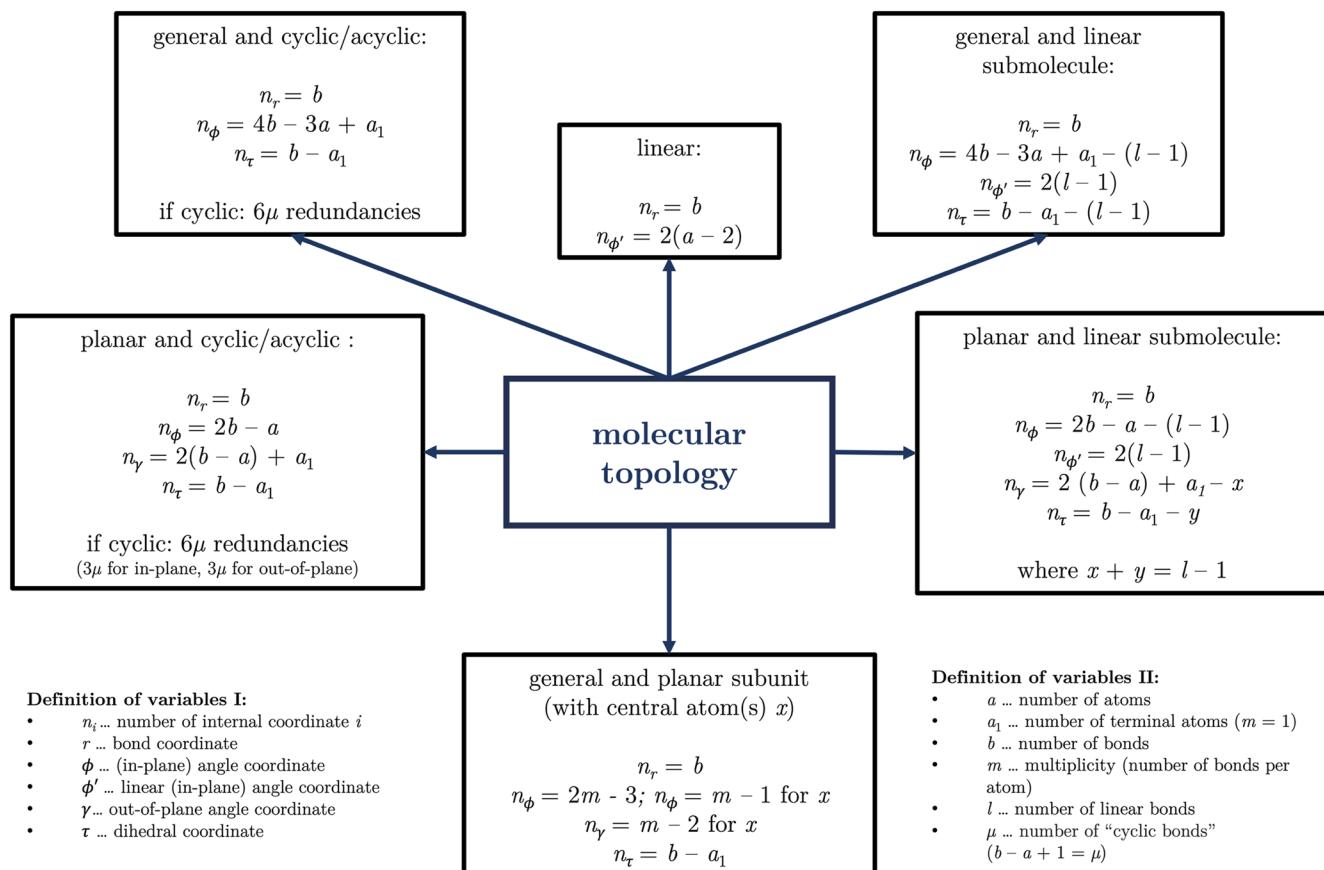


FIG. 3. IC sets can be restricted by molecular topology parameters, which determine the allowed number of ICs of a given coordinate type.

space of definable IC sets (see Table I). We present the best IC set chosen by the NOMODECO toolkit and show how harmonic frequencies and intrinsic frequencies are connected, ultimately evaluating normal modes based on IC contribution heatmaps. The PED analysis provides either the VED matrix or the normalized diagonal elements of the PED matrix. We only show the latter as (1) both matrices propose (nearly) the same contributions and (2) the diagonal elements can be easily converted to percentages, making the contribution heatmaps even more intuitive.

A. Linear molecules

For linear molecules, the only possible IC set comprises all bonds and linear angles.¹³⁶ Therefore, NOMODECO suggests exactly one IC set (cf. Table I). Because of its simplicity, we introduce the NOMODECO analysis with the example of linear CO₂ (cf. Table II). The $\delta(\text{OCO})$ mode (OCO angle deformation) and the $\phi(\text{OCO})$ angle match, as seen by comparing the respective harmonic and intrinsic frequencies. The two degenerate $\phi(\text{OCO})$ angles have the same intrinsic frequencies, similar to the harmonic frequency of the degenerate $\delta(\text{OCO})$ modes. Consequently, one of these ICs contributes 100% to one of the $\delta(\text{OCO})$ modes. In contrast, the intrinsic frequencies of the two r(CO) bonds at 2125 cm⁻¹ are centered

between the harmonic frequencies for the $\nu(\text{CO})$ modes at 1522 and 2592 cm⁻¹. The contribution heatmap (see Fig. 4) shows that both r(CO) bonds contribute equally to the two $\nu(\text{CO})$ modes. Although the decomposition does not show the symmetry of the stretching modes, one could qualitatively determine the direction of IC changes using the D-matrix.

B. Planar, acyclic systems

For planar, acyclic molecules, several possibilities exist to define a non-redundant IC set. Including symmetry and topology criteria (cf. Fig. 3) significantly reduces the number of IC sets. Here, we elaborate on this with the example of formic acid (HCOOH) and formyl cyanide (HCOCN), as shown in Fig. 5. For HCOOH, one can choose from 13 ICs (cf. the supplementary material). The number of ICs in a non-redundant set for HCOOH is $3 \times 5 - 6 = 9$. Generation of all sets comprising 9 out of 13 ICs leads to $\binom{13}{9}$, i.e., 715 definable sets. Even when ensuring a set contains all bonds, the random variation of angles and dihedrals leads to $\binom{9}{5} = 126$ possible sets. These large numbers suggest that a random generation of IC sets is not feasible when considering larger systems.

TABLE I. Number of IC sets for molecules with different topologies and sizes. Including symmetry and topology criteria in generating an IC set significantly reduces the number of possible sets.

Criterion	No. of IC sets					
	CO ₂	HCOOH	HCOCN	CH ₃ OH	HCO ₃ CH ₃	C ₆ H ₁₂
Complete random generation	1	715	2002	455	4 292 145	>10 ³¹
Random generation except bonds	1	126	126	120	203 490	>10 ²³
Symmetry	1	126	126	16	9 312	89
Topology	1	6	2	21	5 775	>10 ¹⁰
Symmetry + topology	1	6	2	3	270	2

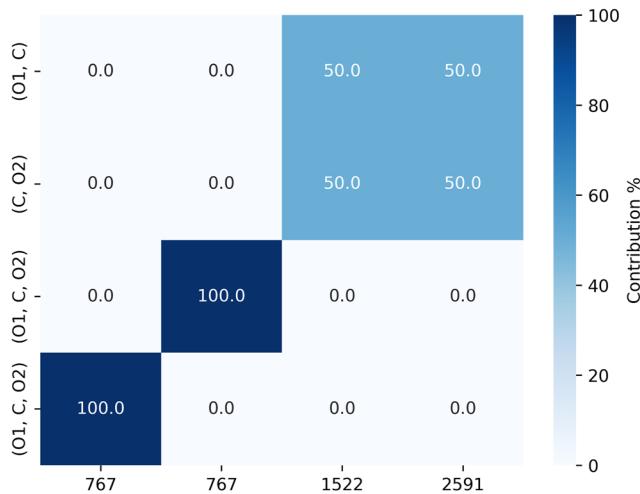
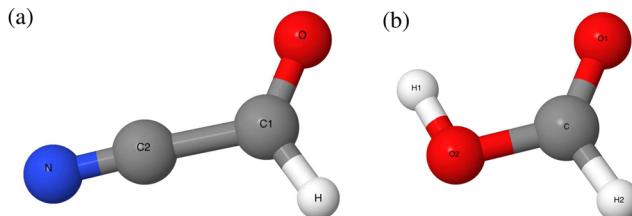
NOMODECO significantly reduces the number of IC sets (see Table I). Symmetry alone has no benefit, as there are no symmetry-equivalent atoms present in HCOOH. From topological parameters (cf. Fig. 3), the number of ICs per type is $n_r = 4$, $n_\phi = 3$, $n_y = 1$, $n_\tau = 1$. We automatically neglect angles that are linearly dependent, which, in the case of *trans*-HCOOH, occurs for one of the three angles that are defined with carbon as the central atom.¹³⁶ Finally, the number

TABLE II. Notation of vibrational frequencies for CO₂.^a

Harmonic frequencies	Assignment ^b	Intrinsic frequencies	Internal coordinates
767 (2)	$\delta(\text{OCO})$	767 (2)	$\phi(\text{OCO})$
1522	$\nu_s(\text{CO})$	2125 (2)	$\text{r}(\text{CO})$
2591	$\nu_{as}(\text{CO})$		

^aAll frequency values are given in cm⁻¹. The frequency degeneracies are indicated in parentheses. The normal modes with the corresponding harmonic frequencies are given in Fig. 1 of the supplementary material.

^bAssignment is formulated in the chemist's notation.

**FIG. 4.** Contribution table for CO₂. Rows represent ICs, and columns represent the harmonic frequencies (in cm⁻¹) for each normal mode.**FIG. 5.** (a) Planar, acyclic molecule with a linear angle: HCOCN in C_s symmetry. (b) Planar, acyclic molecule: *trans*-HCOOH in C_s symmetry.

of possible IC sets to choose from reduces to $\binom{2}{1} \times \binom{3}{1} = 6$. From the six IC sets, NOMODECO picks the optimal set with $\mathcal{M} = 75.2$. For this IC set, the computed harmonic and intrinsic frequencies are shown in Table III and the contribution heatmap in Fig. 6. The main aspect to emphasize here is the increasing delocalized character of the normal modes, especially for the deformational modes $\delta(\text{OCO})$ and $\delta(\text{COH})$.

For planar, acyclic molecules with linear angles, such as formyl cyanide (HCOCN), the topology prescription changes: (1) the number of non-linear angles decreases by the number of linear angles included, and (2) due to the perpendicular linear angle, the number of out-of-plane coordinates (out-of-plane angles and/or dihedral angles) must also decrease. For general non-planar molecules, this is straightforward as there are only dihedral angles. However, for planar molecules, the challenge is to determine the correct decrease in the number of out-of-plane angles and/or dihedral angles. One approach could involve generating different combinations of the topological parameters n_y, n_τ , depending on the values for x and y (cf. Fig. 3). As this leads to an unfeasible number of IC sets, NOMODECO checks if the bonds involved in the introduced linear angles also contribute to out-of-plane and/or terminal dihedral coordinates. If so, the corresponding coordinate is removed, and n_y, n_τ are redefined based on the removed coordinates. To illustrate this, we introduce a linear angle to formic acid by replacing OH with CN [cf. Fig. 5(b)].

From the 14 definable ICs for HCOCN (cf. the supplementary material), $3 \times 5 - 6 = 9$ ICs are needed for a non-redundant IC set. Random generation yields $\binom{13}{9} = 2002$ possible IC sets. Fixing the bonds for each set reduces the number of IC sets to $\binom{9}{5} = 126$. Again, for this molecule, symmetry does not reduce the number of sets. From topological prescriptions, we compute the number

of ICs per type for each set as $n_r = 4$, $n_\phi = 2$, $n_{\phi'} = 2$, $n_y = 1$, $n_\tau = 0$. Note that the number of dihedral angles n_τ is reduced automatically due to the terminal dihedral angles being removed. In addition, one potential out-of-plane coordinate, where atom C2 moves out of the plane, is excluded. Finally, by considering that, analogous to *trans*-HCOOH, one of the three angles centered at C1 needs to be removed, the number of IC sets drastically reduces to $\binom{2}{1} = 2$. The two sets differ in the selected out-of-plane angle, i.e., either $\gamma(\text{C1O})$ or $\gamma(\text{C1H})$. NOMODECO selects the latter since it has a higher value of $\mathcal{M} = 77.3$, compared to $\mathcal{M} = 74.3$. This choice is based on the better description of the out-of-plane mode with the $\gamma(\text{C1H})$ coordinate, mirroring the optimal set for HCOOH.

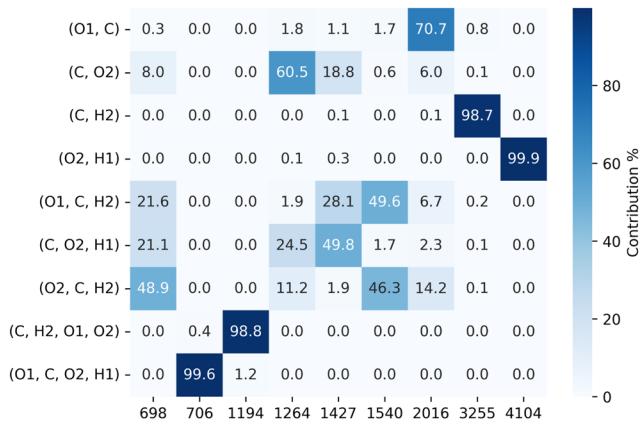
The vibrational notations and the contribution heatmap for the optimal IC set are given in Table IV and Fig. 7. Similar to HCOOH, one can distinguish between localized and delocalized modes by comparing harmonic and intrinsic frequencies, along with the contributions per harmonic frequency. The CCN linear angle has a

TABLE III. Notation of vibrational frequencies for *trans*-HCOOH.^a

Harmonic frequencies	Assignment ^b	Intrinsic frequencies	Internal coordinates
698	$\delta(\text{OCO})$	1242	$\phi(\text{O2CH}_2)$
706	$\tau(\text{COH})$	815	$\tau(\text{O1CO}_2\text{H}_1)$
1194	$\gamma(\text{CH})$	1341	$\gamma(\text{C1H})$
1264	$\nu(\text{C-O})$	1286	$r(\text{CO}_2)$
1427	$\delta(\text{COH})$	1370	$\phi(\text{CO}_2\text{H}_1)$
1540	$\delta(\text{CH})$	1291	$\phi(\text{O1CH}_2)$
2016	$\nu(\text{C=O})$	1961	$r(\text{CO}_1)$
3255	$\nu(\text{CH})$	3248	$r(\text{CH}_2)$
4104	$\nu(\text{OH})$	4105	$r(\text{O2H}_1)$

^aAll frequency values are given in cm^{-1} . The frequency degeneracies are indicated in parentheses. The normal modes with the corresponding harmonic frequencies are given in Fig. 2 of the supplementary material.

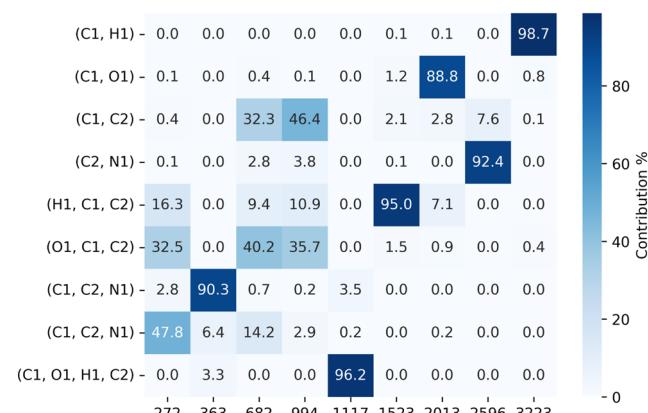
^bAssignment is formulated in the so-called chemist's notation and based on earlier studies.^{146–148}

FIG. 6. Contribution table for *trans*-HCOOH. Rows represent ICs, and columns represent the harmonic frequencies (in cm^{-1}) for each normal mode.TABLE IV. Notation of vibrational frequencies for HCOCN.^a

Harmonic frequencies	Assignment ^b	Intrinsic frequencies	Internal coordinates
272	$\delta(\text{CCN})$	486	$\phi'(\text{C1C2N})$
363	$\delta(\text{CCN})$	449	$\phi'(\text{C1C2N})$
682	$\delta(\text{OCC})$	795	$\phi(\text{OC1C2})$
994	$\nu(\text{CC})$	1114	$r(\text{C1C2})$
1117	$\gamma(\text{CH})$	927	$\gamma(\text{C1H})$
1523	$\delta(\text{CH})$	1507	$\phi(\text{C2C1H})$
2013	$\nu(\text{C=O})$	1999	$r(\text{C1O})$
2596	$\nu(\text{CN})$	2508	$r(\text{C1N})$
3223	$\nu(\text{CH})$	3218	$r(\text{C1H})$

^aAll frequency values are given in cm^{-1} . The frequency degeneracies are indicated in parentheses. The normal modes with the corresponding harmonic frequencies are given in Fig. 3 of the supplementary material.

^bAssignment is formulated in the so-called chemist's notation and derived from the HCOOH assignment.

FIG. 7. Contribution table for HCOCN. Rows represent ICs, and columns represent the harmonic frequencies (in cm^{-1}) for each normal mode.

slightly higher intrinsic frequency than the CCN coordinate perpendicular to it, as it also contributes to the mode at 682 cm^{-1} . Furthermore, comparing the contribution heatmaps of HCOOH (Fig. 6) and HCOCN (Fig. 7) reveals similar vibrating molecular fragments, such as C=O and C–H stretchings.

C. General, acyclic systems

For general (non-planar) molecules, out-of-plane angles are not needed, e.g., methanol [CH_3OH , cf. Fig. 8(a)]. If planar subunits occur, the inclusion of specific out-of-plane angles is needed, e.g., carbonic acid methyl ester [HCO_3CH_3 , cf. Fig. 8(b)].

For methanol, a non-redundant set requires $3 \times 6 - 6 = 12$ ICs out of 15 definable (cf. the supplementary material). There could be $\binom{15}{12} = 455$ IC sets with random generation. Considering fixed bonds and varying angles and dihedrals, there are $\binom{10}{7} = 120$ definable sets. Symmetry considerations can further reduce this number. CH_3OH has a mirror plane, making certain ICs

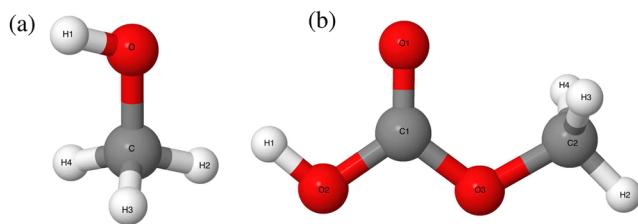


FIG. 8. (a) General, acyclic molecule: CH₃OH in C_s symmetry. (b) General, acyclic molecule with a planar subunit: HCO₃CH₃ in C_s symmetry.

TABLE V. Notation of vibrational frequencies for CH₃OH.^a

Harmonic frequencies	Assignment ^b	Intrinsic frequencies	Internal coordinates
313	$\delta_{\text{oop}}(\text{OH})$	267	$\tau(\text{H}_2\text{OCH}_2)$
1143	$\nu(\text{CO})$	1166	$\text{r}(\text{CO})$
1186	$\omega(\text{CH}_3)$	1323 (2)	$\phi(\text{OCH}_3)/\phi(\text{OCH}_4)$
1286	$\tau(\text{CH}_3)$	1374	$\phi(\text{OCH}_2)$
1487	$\delta_{\text{ip}}(\text{OH})$	1557	$\phi(\text{H}_3\text{CH}_4)$
1617	$\delta_s(\text{CH}_3)$	1627 (2)	$\phi(\text{H}_2\text{CH}_3)/\phi(\text{H}_2\text{CH}_4)$
1624	$\delta_{\text{as}}(\text{CH}_3)$	3187 (2)	$\text{r}(\text{CH}_3)/\text{r}(\text{CH}_4)$
1624	$\delta_{\text{as}}(\text{CH}_3)$	3212	$\text{r}(\text{CH}_2)$
3152	$\nu_s(\text{CH}_3)$	4191	$\text{r}(\text{OH}_1)$
3205	$\nu_{\text{as}}(\text{CH}_2)$	-	-
3226	$\nu_{\text{as}}(\text{CH}_3)$	-	-
4191	$\nu(\text{OH})$	-	-

^a All frequency values are given in cm⁻¹. The frequency degeneracies are indicated in parentheses. The normal modes with the corresponding harmonic frequencies are given in Fig. 4 of the supplementary material.

^b Assignment is formulated in the so-called chemist's notation and based on earlier studies.^{149–151}

symmetry-equivalent: the H₂CH₃ and H₂CH₄ angles, the H₃CO and H₄CO angles, and two dihedral angles, i.e., H₃COH₁ and H₄COH₁. To preserve symmetry in intrinsic frequencies, only IC sets, including all symmetrical counterparts, are allowed. If not possible, NOMODECO tries redundancies to fulfill symmetry criteria; otherwise, it breaks symmetry. Here, the symmetry criterion reduces sets from 120 to 16. Regarding topology (cf. Fig. 3), we calculate the number of ICs per type for CH₃OH as $n_r = 5$, $n_\phi = 6$, $n_\tau = 1$. Without symmetry, the topology criterion yields $\binom{7}{6} \times \binom{3}{1} = 21$ IC sets. Combining symmetry and topology, NOMODECO reduces it to three IC sets. These sets only differ in the chosen angle coordinates. One can either exclude the COH₁ angle or one of the non-symmetry equivalent HCO (H₂CO) or HCH (H₃CH₄) angles. In the first case, the IC set would be incomplete, leaving only two options to be evaluated using the metric \mathcal{M} . The IC set, including all HCH angles and excluding the H₂CO angle, has a $\mathcal{M} = 66.2$. The IC set, including all HCO angles and excluding the H₃CH₄ angle, has a $\mathcal{M} = 61.1$. Thus, the former IC set is preferable. For this IC set, the harmonic and intrinsic frequencies are summarized in Table V, and

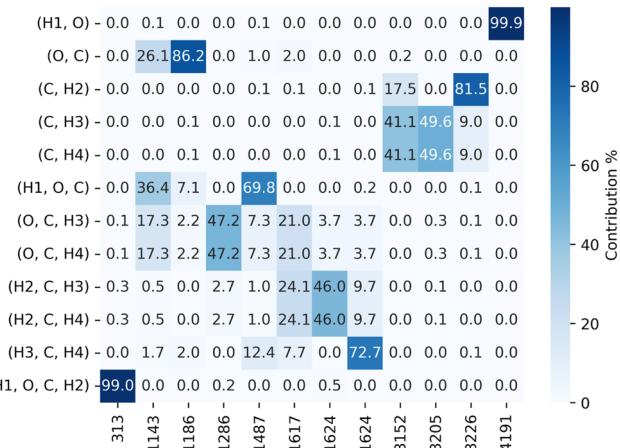


FIG. 9. Contribution table for CH₃OH. Rows represent ICs, and columns represent the harmonic frequencies (in cm⁻¹) for each normal mode.

the contribution heatmap is shown in Fig. 9. While for HCOOH, the contribution of a single IC to specific normal modes is quite clear, this is not the case for CH₃OH. Thus, Table V lists intrinsic frequencies in ascending order instead of listing the ICs most contributing to every normal mode (cf. Table III for HCOOH).

The notation for the CH₃OH normal modes is more involved than in previous examples. As shown in the contribution heatmap (cf. Fig. 9), this is most pronounced for the modes q_2 , q_3 , and q_6 . Contrary to the literature assignments we used in Table V, the contribution heatmap suggests that the CO stretching notation better suits q_3 than q_2 . On the other hand, the intrinsic frequency of the CO bond is centered nearly perfectly between the harmonic frequencies of modes q_2 and q_3 . This illustrates the delocalized character of the NCs as well as the ambiguity arising in vibrational notations. We will discuss this circumstance in an upcoming paper.¹⁵²

If a general (non-planar) molecule has atom(s) X with multiplicities greater than two and all incident bonds are co-planar, these atoms form *planar subunits*. From topology (cf. Fig. 3), three considerations are taken in the NOMODECO toolkit when a planar subunit is detected: (1) Out-of-plane angles are defined for the atom(s) X, (2) the number of angles is recomputed (cf. Fig. 3), and (3) linearly dependent angles with the atom(s) X as the central atom are excluded (see Sec. IV B). As an example, we consider the carbonic acid methyl ester HCO₃CH₃ [cf. Fig. 8(b)].

From the 29 possible ICs (see Table 4 in the supplementary material), $3 \times 9 - 6 = 21$ need to be selected for a non-redundant IC set. A random generation would lead to $\binom{29}{21} = 4\,292\,145$ definable sets. Fixing the bonds for all sets yields $\binom{21}{13} = 203\,490$ IC sets. Selecting one set out of these numbers is computationally not feasible. However, the combined use of topology and symmetry in the NOMODECO toolkit solves this issue. Applying symmetry, i.e., the mirror plane of the C_s point group, one can group two of the HCH (H₂C₂H₃, H₂C₂H₄) and HCO (H₃C₂O₃, H₄C₂O₃) angles, as well as two HCOC dihedral angles (H₃C₂O₃C₁, H₄C₂O₃C₁). By exploiting the symmetry in each IC set analogous to the cases

TABLE VI. Notation of vibrational frequencies for HCO_3CH_3 .^a

Harmonic frequencies	Assignment ^b	Intrinsic frequencies	Internal coordinates
144	CH_3 rotation	271	$\tau(\text{C}2\text{O}3\text{C}1\text{O}2)$
190	$\tau(\text{CO}-\text{CH}_3)$	352	$\tau(\text{H}2\text{C}2\text{O}3\text{C}1)$
320	$\delta(\text{COC})$	550	$\tau(\text{O}3\text{C}1\text{O}2\text{H}1)$
578	$\delta(\text{CO}_3)$	601	$\phi(\text{C}2\text{O}3\text{C}1)$
594	$\delta_{\text{oop}}(\text{OH})$	775	$\phi(\text{O}3\text{C}1\text{O}2)$
738	$\rho(\text{CO}_3)$	854	$\gamma(\text{C}1\text{O}1)$
911	$\delta_{\text{oop}}(\text{CO}_3)$	903	$\phi(\text{O}3\text{C}1\text{O}1)$
1012	$\nu(\text{O}-\text{CH}_3)$	1101	$r(\text{C}2\text{O}3)$
1223	$\nu(\text{CO})$	1302	$r(\text{C}1\text{O}2)$
1291	$\tau(\text{CH}_3)$	1319	$r(\text{C}1\text{O}3)$
1339	$\omega(\text{CH}_3)$	1330 (2)	$\phi(\text{H}3\text{C}2\text{O}3)/\phi(\text{H}4\text{C}2\text{O}3)$
1362	$\delta_{\text{ip}}(\text{OH})$	1404	$\phi(\text{C}1\text{O}2\text{H}1)$
1547	$\delta_{\text{ip}}(\text{OH}), \delta_{\text{ip}}(\text{CO}_3)$	1532	$\phi(\text{H}3\text{C}2\text{H}4)$
1613	$\delta_{\text{as}}(\text{CH}_3)$	1620 (2)	$\phi(\text{H}2\text{C}2\text{H}3)/\phi(\text{H}2\text{C}2\text{H}4)$
1619	$\delta_{\text{as}}(\text{CH}_3)$	1873	$r(\text{C}1\text{O}1)$
1630	$\delta_s(\text{CH}_3)$	3261 (2)	$(\text{C}2\text{H}3)/r(\text{C}2\text{H}4)$
1999	$\nu(\text{C}=\text{O})$	3294	$r(\text{C}2\text{H}2)$
3211	$\nu_s(\text{CH}_3)$	4153	$r(\text{O}2\text{H}1)$
3289	$\nu_{\text{as}}(\text{CH}_2)$		
3315	$\nu_{\text{as}}(\text{CH}_3)$		
4152	$\nu(\text{OH})$		

^a All frequency values are given in cm^{-1} . The frequency degeneracies are indicated in parentheses. The normal modes with the corresponding harmonic frequencies are given in Fig. 5 of the supplementary material.

^b Assignment is formulated in the so-called chemist's notation and based on earlier studies.^{153,154}

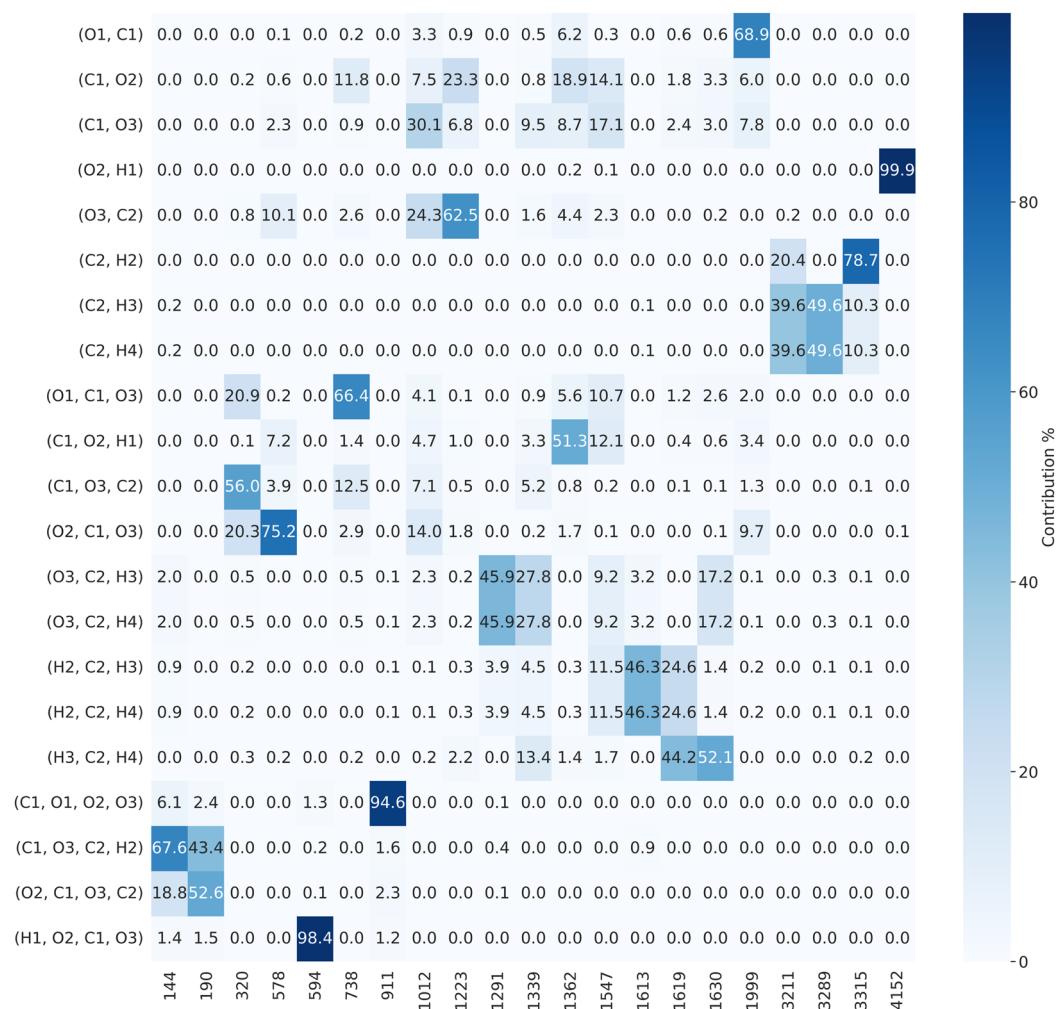
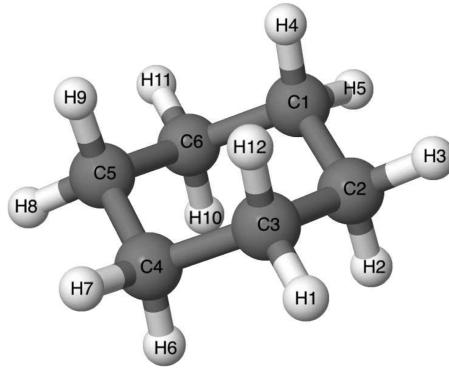
discussed above, the number of IC sets can be reduced to 9312. From topology (cf. Fig. 3), the number of ICs for each type is given as $n_r = 8, n_\phi = 9, n_\gamma = 1, n_\tau = 3$. By imposing topological constraints, only $\binom{11}{9} \times \binom{3}{1} \times \binom{7}{3} = 5575$ IC sets are allowed. When combining symmetry and topology, the number of definable IC sets is reduced to 270, significantly enhancing computational feasibility. Still, 270 normal mode decompositions must be performed to find the IC set with the optimal metric. The computed harmonic and intrinsic frequencies, as well as the contribution heatmap for the optimal IC set, are given in Table VI and Fig. 10, respectively. While most normal modes exhibit a delocalized character, the contribution heatmap offers a clear and intuitive view of individual IC contributions.

D. Cyclic systems

Generating IC sets for cyclic systems is a significant challenge, as adhering to the topological prescriptions (see Fig. 3), these systems inherently introduce six μ redundant ICs. Here, μ represents the number of bonds that must be removed to render the molecule acyclic. These cyclic redundancies can be distinguished from local ones¹³⁶ included to preserve symmetry within an IC set. NOMODECO automatically eliminates such redundancies by removing μ bond(s) along with their associated ICs and then applying the original acyclic system scheme shown in Fig. 3. The only requirement is that after bond exclusion, the resulting coordinate

set still describes a molecule where all atoms are connected either directly or indirectly. While this method is universally applicable, it has the crucial drawback that excluding bonds and their associated ICs leads to broken symmetry in the intrinsic frequencies for some cyclic systems. Nevertheless, these wavenumber shifts are typically small and should not significantly affect the intrinsic frequency analysis. One strategy to address this issue for highly symmetrical systems involves applying a correction scheme to ensure that intrinsic frequency values, which should be symmetrically equivalent, match.¹³¹ NOMODECO aims to remove symmetry-equivalent bonds to temper the impact of symmetry breaking.

As an example, we demonstrate the generation of IC sets for cyclohexane (C_6H_{12}) in chair conformation (see Fig. 11). Due to size, we show neither the vibrational modes nor a full assignment and discuss merely the key results from the combinatorial analysis and the normal mode decomposition. From the 108 possible ICs (cf. the supplementary material), $3 \times 18 - 6 = 48$ ICs must be selected for a non-redundant set. The number of possibilities, with and without fixed bonds, during random generation is so high that we only approximate them using decimal powers in Table I. Regarding topology, the molecule is rendered acyclic, i.e., in the case of C_6H_{12} , a bond with the associated ICs is excluded (cf. the supplementary material). The numbers for each IC type per set, therefore, become $n_r = 17, n_\phi = 26, n_\tau = 5$. Imposing topological constraints still results in $\binom{30}{26} \times \binom{39}{5}$ IC sets, which remains

**FIG. 10.** Contribution table for HCO_3CH_3 . Rows represent ICs, and columns represent the harmonic frequencies (in cm^{-1}) for each normal mode.**FIG. 11.** General, cyclic molecule: C_6H_{12} in D_{3d} symmetry.

computationally unfeasible. However, considering only symmetry leads to a dramatic decrease to 89 definable sets. Ultimately, combining both topology and symmetry yields two IC sets being generated in NOMODECO. Once again, this emphasizes the substantial impact of combining topology and symmetry considerations in generating IC sets. Table VII presents the computed harmonic and intrinsic frequencies for the optimal IC set. We show the contribution heatmap in the supplementary material due to its size. It visually groups normal modes into stretching, bending, or torsion motions.

Finally, we want to point out again that, as per symmetry expectations, one would anticipate uniform intrinsic frequency values for all CC stretches and two sets of CH stretches and HCC bends in Table VII.¹³¹ However, this uniformity is not achieved due to the aforementioned algorithmic treatment of cyclic systems.

TABLE VII. Computed harmonic and intrinsic frequencies for C₆H₁₂.^a

Harmonic frequencies	Intrinsic frequencies	Internal coordinates
247 (2)	328 (2)	$\tau(\text{H1C3C4H7})/\tau(\text{H8C5C6H11})$
402	636	$\tau(\text{H7C4C5H8})$
457 (2)	989 (2)	$\tau(\text{H1C3C2H3})/\tau(\text{H5C1C6H11})$
559	1050 (2)	$r(\text{C3C4})/r(\text{C5C6})$
847	1058	$r(\text{C4C5})$
855 (2)	1062 (2)	$\phi(\text{C4C5H9})/\phi(\text{C5C4H6})$
917 (2)	1107 (2)	$\phi(\text{C5C6H10})/\phi(\text{C4C3H12})$
987 (2)	1121 (2)	$\phi(\text{C4C3H1})/\phi(\text{C5C6H11})$
1100 (2)	1152 (2)	$r(\text{C1C6})/r(\text{C2C3})$
1126	1204 (2)	$\phi(\text{C4C5H8})/\phi(\text{C5C4H7})$
1149 (2)	1222 (2)	$\phi(\text{C6C5H8})/\phi(\text{C3C4H7})$
1170	1243 (2)	$\phi(\text{C3C4H6})/\phi(\text{C6C5H9})$
1228	1252 (2)	$\phi(\text{C2C3H12})/\phi(\text{C1C6H10})$
1283	1297 (2)	$\phi(\text{C3C2H3})/\phi(\text{C6C1H5})$
1385 (2)	1385 (2)	$\phi(\text{C6C1H4})/\phi(\text{C3C2H2})$
1396 (2)	1457 (2)	$\phi(\text{H6C4H7})/\phi(\text{H8C5H9})$
1460	1486 (2)	$\phi(\text{C1C6H11})/\phi(\text{C2C3H1})$
1503 (3)	1495 (2)	$\phi(\text{H1C3H12})/\phi(\text{H10C6H11})$
1510 (2)	1609 (2)	$\phi(\text{H2C2H3})/\phi(\text{H4C1H5})$
1603 (2)	3150 (2)	$r(\text{C1H4})/r(\text{C2H2})$
1613 (3)	3151 (2)	$r(\text{C4H6})/r(\text{C5H9})$
1634	3153 (2)	$r(\text{C3H12})/r(\text{C6H10})$
3137 (2)	3178 (2)	$r(\text{C1H5})/r(\text{C2H3})$
3141 (3)	3178 (2)	$r(\text{C4H7})/r(\text{C5H8})$
3143	3179 (2)	$r(\text{C3H1})/r(\text{C6H11})$
3184 (2)		
3186 (2)		
3191		
3198		

^a All frequency values are given in cm⁻¹. The frequency degeneracies are indicated in parentheses.

V. CONCLUSION

Using internal coordinates (ICs) in quantum chemistry problems, e.g., molecular spectroscopy, provides a means to overcome limitations posed by rectilinear normal coordinates (NCs). Although very promising applications of ICs in terms of molecular spectroscopy already exist,^{87–90,111} these applications are usually demonstrated for systems where the choice of ICs is trivial and based on what one may call “chemical intuition.” However, as such intuitive choices can be elusive, especially for larger molecules, a systematic choice of ICs is preferable. In the present work, we implemented an approach based on the harmonic approximation and a normal mode decomposition scheme.

The number of IC sets increases drastically with molecule size. Thus, for mid-sized molecules, it is computationally not feasible to determine an optimal set by considering the whole space of definable sets. We here show that straightforward symmetry and topology rules can be automatically considered for screening reasonable sets. This greatly reduces the space of IC sets and allows the use of normal mode decomposition schemes for selecting an optimal one.

The NOMODECO toolkit applies to mid-sized rigid, semi-rigid, and floppy systems, and so far, we have tested molecules with up to 21 atoms. We chose a simple metric to select an optimal IC set based on the application in molecular spectroscopy. Therefore, we defined the optimal set to be the one with the highest separability in terms of the harmonic potential energy, i.e., it leads to a unique description of normal modes in terms of the ICs. This core idea can be transferred to other applications, e.g., IC-based VSCF/VCI, although the metric has to be redefined to include additional potential information.

The normal mode decomposition is derived from the established potential energy distribution (PED) analysis.^{11,117,119–121} While more sophisticated theoretical schemes, such as the local mode theory,¹⁵⁵ provide even more information for delocalized vibrations, the simplicity of the PED scheme makes it suitable for probing multiple IC sets for a given molecule. The clarity of PED results is also attributed to the employed ICs, which equal simple structural parameters of a molecule (bonds, angles, and dihedrals). The present work provides such information almost fully automated, as it goes hand in hand with selecting an optimal IC set.

We rely on primitive (Wilson-type) ICs. Employing mixed coordinates based on linear combinations of these¹³⁴ can address redundancy issues and enhance nominal separability. However, using such coordinates may complicate the interpretability of a normal mode decomposition. Still, adding new IC types to the NOMODECO toolkit, such as combined torsion coordinates for methyl groups,^{144,145} and adapting the guidelines from Fig. 3 for these coordinates, represents a next, natural step. Furthermore, the recent effort to compute anharmonic vibrational spectra for dimers¹⁵⁶ opens the question of how to expand the NOMODECO scheme to such systems. In conclusion, for low-symmetry and/or large molecules, additional strategies beyond symmetry and topology constraints are needed to limit the space of definable IC sets.

SUPPLEMENTARY MATERIAL

In the supplementary material, we provide additional information on introducing penalties to modify the metric for specific cases. Furthermore, the supplementary material contains pictures of the vibrational modes (except C₆H₁₂), all definable internal coordinates for each of the shown systems, and the contribution heatmap for C₆H₁₂.

ACKNOWLEDGMENTS

We are grateful to Guntram Rauhut for making several helpful suggestions regarding the theoretical concepts and key messages of this work. This work was supported by the Austrian Science Fund via Grant No. P34518.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Kemal Oenen: Conceptualization (lead); Methodology (lead); Software (lead); Writing – original draft (lead); Writing – review & editing (equal). **Dennis F. Dinu:** Conceptualization (supporting); Methodology (supporting); Writing – review & editing (equal). **Klaus R. Liedl:** Conceptualization (supporting); Resources (lead); Supervision (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available within the article and its supplementary material. The used Hessians can be found in the NOMODECO GitHub repository.¹³³

REFERENCES

- ¹E. B. Wilson and J. B. Howard, *J. Chem. Phys.* **4**, 260 (1936).
²E. B. Wilson, J. C. Decius, P. C. Cross, and B. R. Sundheim, *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra* (Dover Publications, 1955).
³J. K. G Watson, *Mol. Phys.* **15**, 479 (1968).
⁴J. K. G Watson, *Mol. Phys.* **79**, 943 (1993).
⁵R. Meyer and H. H. Günthard, *J. Chem. Phys.* **49**, 1510 (1968).
⁶R. Meyer and H. H. Günthard, *J. Chem. Phys.* **50**, 353 (1969).
⁷J. K. G Watson, *J. Mol. Spectrosc.* **228**, 645 (2004).
⁸J. Zúñiga, A. Bastida, and A. Requena, *J. Chem. Soc., Faraday Trans.* **93**, 1681 (1997).
⁹S. Califano, *Vibrational States* (Wiley, London, 1976).
¹⁰Z. Baćić and J. C. Light, *Annu. Rev. Phys. Chem.* **40**, 469 (1989).
¹¹E. Whittaker, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies, with an Introduction to the Problem of Three Bodies*, 2nd ed. (Cambridge University Press, 1917).
¹²D. M. Dennison, *Rev. Mod. Phys.* **3**, 280 (1931).
¹³E. B. Wilson, *J. Chem. Phys.* **7**, 1047 (1939).
¹⁴E. B. Wilson, *J. Chem. Phys.* **9**, 76 (1941).
¹⁵D. M. Bishop and D. J. Klein, *Group Theory and Chemistry* (Dover Publications, 1975).
¹⁶E. L. Klinting, D. Lauvergnat, and O. Christiansen, *J. Chem. Theory Comput.* **16**, 4505 (2020).
¹⁷C. Eckart, *Phys. Rev.* **47**, 552 (1935).
¹⁸E. Mátyus, G. Czakó, and A. G. Császár, *J. Chem. Phys.* **130**, 134112 (2009).
¹⁹O. Christiansen, *Phys. Chem. Chem. Phys.* **9**, 2942 (2007).
²⁰R. J. Whitehead and N. C. Handy, *J. Mol. Spectrosc.* **55**, 356 (1975).
²¹J. M. Bowman, K. Christoffel, and F. Tobin, *J. Phys. Chem.* **83**, 905 (1979).
²²K. M. Dunn, J. E. Boggs, and P. Pulay, *J. Chem. Phys.* **85**, 5838 (1986).
²³K. M. Dunn, J. E. Boggs, and P. Pulay, *J. Chem. Phys.* **86**, 5088 (1987).
²⁴D. J. Searles and E. I. von Nagy-Felsobuki, *J. Chem. Phys.* **95**, 1107 (1991).
²⁵J. O. Jung and R. B. Gerber, *J. Chem. Phys.* **105**, 10332 (1996).
²⁶J. O. Jung and R. B. Gerber, *J. Chem. Phys.* **105**, 10682 (1996).
²⁷K. Yagi, T. Taketsugu, K. Hirao, and M. S. Gordon, *J. Chem. Phys.* **113**, 1005 (2000).
²⁸J. M. Bowman, S. Carter, and X. Huang, *Int. Rev. Phys. Chem.* **22**, 533 (2003).
²⁹C. Léonard, N. C. Handy, S. Carter, and J. M. Bowman, *Spectrochim. Acta, Part A* **58**, 825 (2002).
³⁰T. Yonehara, T. Yamamoto, and S. Kato, *Chem. Phys. Lett.* **393**, 98 (2004).
³¹G. Rauhut, *J. Chem. Phys.* **121**(19), 9313 (2004).
³²E. Mátyus, G. Czakó, B. T. Sutcliffe, and A. G. Császár, *J. Chem. Phys.* **127**, 084102 (2007).
³³E. Kauppi, *J. Chem. Phys.* **105**, 7986 (1996).
³⁴A. E. Roitberg, R. B. Gerber, and M. A. Ratner, *J. Phys. Chem. B* **101**, 1700 (1997).
³⁵S. Carter, S. J. Culik, and J. M. Bowman, *J. Chem. Phys.* **107**, 10458 (1997).
³⁶S. Carter, J. M. Bowman, and L. B. Harding, *Spectrochim. Acta, Part A* **53**, 1179 (1997).
³⁷S. Carter, J. M. Bowman, and N. C. Handy, *Theor. Chem. Acc.* **100**, 191 (1998).
³⁸S. Carter and J. M. Bowman, *J. Chem. Phys.* **108**, 4397 (1998).
³⁹G. M. Chaban, J. O. Jung, and R. B. Gerber, *J. Chem. Phys.* **111**, 1823 (1999).
⁴⁰T. C. Thompson and D. G. Truhlar, *J. Chem. Phys.* **77**, 3031 (1982).
⁴¹N. Moiseyev, *Chem. Phys. Lett.* **98**, 233 (1983).
⁴²K. Yagi, M. Keçeli, and S. Hirata, *J. Chem. Phys.* **137**, 204118 (2012).
⁴³G. Rauhut, *J. Chem. Phys.* **127**, 184109 (2007).
⁴⁴X. Cheng and R. P. Steele, *J. Chem. Phys.* **141**, 104105 (2014).
⁴⁵X. Cheng, J. J. Talbot, and R. P. Steele, *J. Chem. Phys.* **145**, 124112 (2016).
⁴⁶A. Molina, P. Smereka, and P. M. Zimmerman, *J. Chem. Phys.* **144**, 124111 (2016).
⁴⁷B. Ziegler and G. Rauhut, *J. Chem. Phys.* **149**, 164110 (2018).
⁴⁸B. Ziegler and G. Rauhut, *J. Chem. Theory Comput.* **15**, 4187 (2019).
⁴⁹P. T. Panek, A. A. Hoeske, and C. R. Jacob, *J. Chem. Phys.* **150**, 054107 (2019).
⁵⁰C. R. Jacob and M. Reiher, *J. Chem. Phys.* **130**, 084106 (2009).
⁵¹C. R. Jacob, S. Luber, and M. Reiher, *J. Phys. Chem. B* **113**, 6558 (2009).
⁵²C. R. Jacob, *ChemPhysChem* **12**, 3291 (2011).
⁵³P. T. Panek and C. R. Jacob, *ChemPhysChem* **15**, 3365 (2014).
⁵⁴P. T. Panek and C. R. Jacob, *J. Phys. Chem. Lett.* **7**, 3084 (2016).
⁵⁵E. L. Klinting, C. König, and O. Christiansen, *J. Phys. Chem. A* **119**, 11007 (2015).
⁵⁶E. Klinting, B. Thomsen, I. Godtliebsen, and O. Christiansen, *J. Chem. Phys.* **148**, 064113 (2018).
⁵⁷F. Gatti, C. Iung, M. Menou, Y. Justum, A. Nauts, and X. Chapuisat, *J. Chem. Phys.* **108**, 8804 (1998).
⁵⁸F. Gatti, C. Iung, M. Menou, and X. Chapuisat, *J. Chem. Phys.* **108**, 8821 (1998).
⁵⁹F. Gatti, *J. Chem. Phys.* **111**, 7225–7235 (1999).
⁶⁰C. Leforestier, A. Viel, F. Gatti, C. Muñoz, and C. Iung, *J. Chem. Phys.* **114**, 2099 (2001).
⁶¹R. Radau, *Ann. Sci. Éc. Norm. Supér.* **5**, 311 (1868).
⁶²F. T. Smith, *Phys. Rev. Lett.* **45**, 1157 (1980).
⁶³X.-G. Wang and T. Carrington, *J. Chem. Phys.* **129**, 234102 (2008).
⁶⁴M. J. Bramley and N. C. Handy, *J. Chem. Phys.* **98**, 1378 (1993).
⁶⁵L. Hedberg and I. M. Mills, *J. Mol. Spectrosc.* **160**, 117 (1993).
⁶⁶B. R. Johnson, *J. Chem. Phys.* **73**, 5051 (1980).
⁶⁷X. Chapuisat, J. P. Brunet, and A. Nauts, *Chem. Phys. Lett.* **136**, 153 (1987).
⁶⁸T. R. Horn, R. B. Gerber, and M. A. Ratner, *J. Chem. Phys.* **91**, 1813 (1989).
⁶⁹X. Chapuisat, A. Nauts, and J. P. Brunet, *Mol. Phys.* **72**, 1 (1991).
⁷⁰X. Chapuisat, *Mol. Phys.* **72**, 1233 (1991).
⁷¹T. R. Horn, R. B. Gerber, J. J. Valentini, and M. A. Ratner, *J. Chem. Phys.* **94**, 6728 (1991).
⁷²L. L. Gibson, R. M. Roth, M. A. Ratner, and R. B. Gerber, *J. Chem. Phys.* **85**, 3425 (1986).
⁷³Z. Baćić, R. B. Gerber, and M. A. Ratner, *J. Phys. Chem.* **90**, 3606 (1986).
⁷⁴F. Gatti and C. Iung, *Phys. Rep.* **484**, 1 (2009).
⁷⁵C. Iung, F. Gatti, A. Viel, and X. Chapuisat, *Phys. Chem. Chem. Phys.* **1**, 3377 (1999).
⁷⁶X. Chapuisat and C. Iung, *Phys. Rev. A* **45**, 6217 (1992).
⁷⁷M. Mladenović, *J. Chem. Phys.* **112**, 1070 (2000).
⁷⁸M. Mladenović, *J. Chem. Phys.* **112**, 1082 (2000).
⁷⁹F. Gatti, C. Muñoz, and C. Iung, *J. Chem. Phys.* **114**, 8275 (2001).
⁸⁰F. Gatti and A. Nauts, *Chem. Phys.* **295**, 167 (2003).
⁸¹M. Ndong, L. Joubert-Doriol, H. D. Meyer, A. Nauts, F. Gatti, and D. Lauvergnat, *J. Chem. Phys.* **136**, 034107 (2012).
⁸²M. Ndong, A. Nauts, L. Joubert-Doriol, H. D. Meyer, F. Gatti, and D. Lauvergnat, *J. Chem. Phys.* **139**, 204107 (2013).
⁸³K. Sadri, D. Lauvergnat, F. Gatti, and H. D. Meyer, *J. Chem. Phys.* **136**, 234112 (2012).

- ⁸⁴K. Sadri, D. Lauvergnat, F. Gatti, and H.-D. Meyer, *J. Chem. Phys.* **141**, 114101 (2014).
- ⁸⁵C. Iung and F. Gatti, *Int. J. Quantum Chem.* **106**, 130 (2006).
- ⁸⁶D. Mendive-Tapia, H. D. Meyer, and O. Vendrell, *J. Chem. Theory Comput.* **19**, 1144 (2023).
- ⁸⁷D. Lauvergnat and A. Nauts, *J. Chem. Phys.* **116**, 8560 (2002).
- ⁸⁸I. Suwan and R. B. Gerber, *Chem. Phys.* **373**, 267 (2010).
- ⁸⁹D. Strobusch and C. Scheurer, *J. Chem. Phys.* **135**, 124102 (2011).
- ⁹⁰D. Strobusch and C. Scheurer, *J. Chem. Phys.* **135**, 144101 (2011).
- ⁹¹D. F. McIntosh, K. H. Michaelian, and M. R. Peterson, *Can. J. Chem.* **56**, 1289 (1978).
- ⁹²A. Banerjee, N. Adams, J. Simons, and R. Shepard, *J. Phys. Chem.* **89**, 52–57 (1985).
- ⁹³J. Baker, *J. Comput. Chem.* **7**, 385 (1986).
- ⁹⁴P. Császár and P. Pulay, *J. Mol. Struct.* **114**, 31 (1984).
- ⁹⁵H. B. Schlegel, *J. Comput. Chem.* **3**, 214 (1982).
- ⁹⁶P. Pulay and G. Fogarasi, *J. Chem. Phys.* **96**, 2856 (1992).
- ⁹⁷F. Eckert, P. Pulay, and H.-J. Werner, *J. Comput. Chem.* **18**, 1473 (1997).
- ⁹⁸J. Baker and W. J. Hehre, *J. Comput. Chem.* **12**, 606 (1991).
- ⁹⁹J. Baker, *J. Comput. Chem.* **14**, 1085 (1993).
- ¹⁰⁰J. A. Pople, P. V. Schleyer, W. J. Hehre, and L. Radom, *AB INITIO Molecular Orbital Theory* (Wiley, 1986).
- ¹⁰¹J. Baker, A. Kessi, and B. Delley, *J. Chem. Phys.* **105**, 192 (1996).
- ¹⁰²G. Fogarasi, X. Zhou, P. W. Taylor, and P. Pulay, *J. Am. Chem. Soc.* **114**, 8191–8201 (1992).
- ¹⁰³V. Bakken and T. Helgaker, *J. Chem. Phys.* **117**, 9160 (2002).
- ¹⁰⁴J. M. Bowman, *J. Chem. Phys.* **68**, 608 (1978).
- ¹⁰⁵M. Neff and G. Rauhut, *J. Chem. Phys.* **131**, 124129 (2009).
- ¹⁰⁶G. D. Carney, L. L. Sprandel, and C. W. Kern, “Variational approaches to vibration-rotation spectroscopy for polyatomic molecules,” in *Advances in Chemical Physics* edited by I. Prigogine and S. A. Rice (John Wiley & Sons, Ltd., 1978), Vol. 37, pp. 305–379.
- ¹⁰⁷M. Bounouar and C. Scheurer, *Chem. Phys.* **347**, 194 (2008).
- ¹⁰⁸B. Njegic and M. S. Gordon, *J. Chem. Phys.* **125**, 224102 (2006).
- ¹⁰⁹B. Njegic and M. S. Gordon, *J. Chem. Phys.* **129**, 164107 (2008).
- ¹¹⁰Y. Scribano, D. M. Lauvergnat, and D. M. Benoit, *J. Chem. Phys.* **133**, 094103 (2010).
- ¹¹¹I. Bulik, M. Frisch, and P. Vaccaro, *J. Chem. Phys.* **147**, 044110 (2017).
- ¹¹²R. Mecke, *Z. Phys. Chem.* **16B**, 409 (1932).
- ¹¹³R. Mecke, *Z. Phys. Chem.* **16B**, 421 (1932).
- ¹¹⁴B. Schrader, *Infrared and Raman Spectroscopy: Methods and Applications* (VCH, 1996).
- ¹¹⁵P. Torkington, *J. Chem. Phys.* **17**, 357 (1949).
- ¹¹⁶W. J. O. Thomas, *J. Chem. Phys.* **19**, 1162 (1951).
- ¹¹⁷Y. Morino and K. Kuchitsu, *J. Chem. Phys.* **20**, 1809 (1952).
- ¹¹⁸R. A. Munos, Y. N. Panchenko, G. S. Koptev, and N. F. Stepanov, *J. Appl. Spectrosc.* **12**, 428 (1970).
- ¹¹⁹G. Kereszty and G. Jalszovszky, *J. Mol. Struct.* **10**, 304 (1971).
- ¹²⁰W. J. Taylor, *J. Chem. Phys.* **22**, 1780 (1954).
- ¹²¹E. Rytter, *J. Chem. Phys.* **60**, 3882 (1974).
- ¹²²P. Adámek, *J. Mol. Spectrosc.* **57**, 164 (1975).
- ¹²³A. J. P. Alix and A. Müller, *J. Mol. Struct.* **24**, 229 (1975).
- ¹²⁴A. J. P. Alix *et al.*, *Z. Naturforsch. A* **29**, 1454 (1974).
- ¹²⁵A. J. P. Alix and E. Rytter, *Z. Naturforsch. A* **35**, 1142 (1980).
- ¹²⁶A. J. P. Alix, *Spectrosc. Lett.* **14**, 441 (1981).
- ¹²⁷G. Bánhegyi, G. Fogarasi, and P. Pulay, *J. Mol. Struct.: THEOCHEM* **89**, 1 (1982).
- ¹²⁸A. Natarajan and S. Somasundaram, *Acta Phys. Acad. Sci. Hung.* **52**, 237 (1982).
- ¹²⁹N. C. Craig and S. V. Krasnoshchekov, *Mol. Phys.* **117**, 1059 (2019).
- ¹³⁰J. C. Whitmer, *J. Mol. Spectrosc.* **68**, 326 (1977).
- ¹³¹J. A. Boatz and M. S. Gordon, *J. Phys. Chem.* **93**, 1819 (1989).
- ¹³²J. A. Boatz and M. S. Gordon, *J. Phys. Chem.* **93**, 5774 (1989).
- ¹³³K. Oenen (2023). “GitHub repository: Nomodeco toolkit,” Figshare. <https://doi.org/10.6084/m9.figshare.24716127.v1>
- ¹³⁴M. H. Jamróz, *Spectrochim. Acta, Part A* **114**, 220 (2013).
- ¹³⁵J. Brunvoll, B. N. Cyvin, and S. J. Cyvin, *Z. Naturforsch. A* **37**, 342–345 (1982).
- ¹³⁶J. C. Decius, *J. Chem. Phys.* **17**, 1315–1318 (1949).
- ¹³⁷E. Mátyus, C. Fábri, T. Szidárovszky, G. Czakó, W. D. Allen, and A. G. Császár, *J. Chem. Phys.* **133**, 034113 (2010).
- ¹³⁸H. J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2**, 242 (2012).
- ¹³⁹M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Yule, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, *Gaussian 16, Revision C.01*, Gaussian, Inc., Wallingford, CT, 2016.
- ¹⁴⁰F. Neese, F. Wennmohs, U. Becker, and C. Ripplinger, *J. Chem. Phys.* **152**, 224108 (2020).
- ¹⁴¹B. P. Winnewisser and J. K. G. Watson, *J. Mol. Spectrosc.* **205**, 227 (2001).
- ¹⁴²S. J. Cyvin and J. Brunvoll, *J. Mol. Spectrosc.* **40**, 431 (1971).
- ¹⁴³L. A. Curtiss, *J. Mol. Spectrosc.* **44**, 605 (1972).
- ¹⁴⁴R. L. Hilderbrandt, *J. Mol. Spectrosc.* **44**, 599 (1972).
- ¹⁴⁵I. H. Williams, *J. Mol. Spectrosc.* **66**, 288 (1977).
- ¹⁴⁶I. D. Reva, A. M. Plokhotnichenko, E. D. Radchenko, G. G. Sheina, and Y. P. Blagoi, *Spectrochim. Acta, Part A* **50**, 1107 (1994).
- ¹⁴⁷K. Marushkevich, L. Khriachtchev, J. Lundell, A. V. Domanskaya, and M. Räsänen, *J. Mol. Spectrosc.* **259**, 105 (2010).
- ¹⁴⁸E. M. Maçôas, J. Lundell, M. Pettersson, L. Khriachtchev, R. Fausto, and M. Räsänen, *J. Mol. Spectrosc.* **219**, 70 (2003).
- ¹⁴⁹D. F. Dinu, M. Podewitz, H. Grothe, T. Loerting, and K. R. Liedl, *Theor. Chem. Acc.* **139**, 174 (2020).
- ¹⁵⁰A. Serrallach, R. Meyer, and H. H. Günthard, *J. Mol. Spectrosc.* **52**, 94 (1974).
- ¹⁵¹M. Falk and E. Whalley, *J. Chem. Phys.* **34**, 1554–1568 (1961).
- ¹⁵²D. F. Dinu, K. Oenen, J. Schlagin, M. Podewitz, H. Grothe, T. Loerting, and K. R. Liedl, “Explaining resonances in the IR spectrum of CO₂ and CH₃OH: Limits of conventional vibrational notations” (unpublished) (2023).
- ¹⁵³J. Bernard, E. M. Köck, R. G. Huber, K. R. Liedl, L. Call, R. Schlögl, H. Grothe, and T. Loerting, *RSC Adv.* **7**, 22222–22233 (2017).
- ¹⁵⁴E. M. Köck, J. Bernard, M. Podewitz, D. F. Dinu, R. G. Huber, K. R. Liedl, H. Grothe, E. Bertel, R. Schlögl, and T. Loerting, *Chem. - Eur. J.* **26**, 285 (2020).
- ¹⁵⁵E. Kraka, M. Quintana, H. W. La Force, J. J. Antonio, and M. Freindorf, *J. Phys. Chem. A* **126**, 8781–8798 (2022).
- ¹⁵⁶D. F. Dinu, P. Bartl, P. K. Quoika, M. Podewitz, K. R. Liedl, H. Grothe, and T. Loerting, *J. Phys. Chem. A* **126**, 2966–2975 (2022).