JCTC Journal of Chemical Theory and Computation

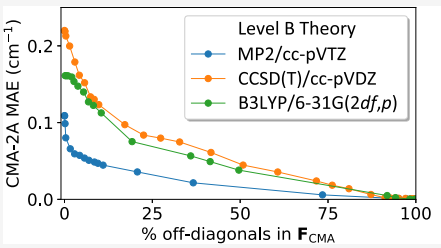
pubs.acs.org/JCTC Article

Convergent Concordant Mode Approach for Molecular Vibrations: CMA-2

Nathaniel L. Kitzmiller, Mitchell E. Lahm, Laura N. Olive Dornshuld, Jincan Jin, Wesley D. Allen, and Henry F. Schaefer III*

Cite This: J. Chem. Theory Comput. 2024, 20, 10886−10898 Read Online

ABSTRACT: The concordant mode approach (CMA) is a promising new scheme for dramatically increasing the system size and level of theory achievable in quantum chemical computations of molecular vibrational frequencies. Here, we achieve advances in the CMA hierarchy by computations targeting CCSD(T)/cc-pVTZ (coupled cluster singles and doubles with perturbative triples using a correlation-consistent polarized-valence triple-ζ basis set) benchmarks within the G2 molecular test set, executing a statistical analysis for 1501 frequencies from 111 compounds and then separately solving the refractory case of pyridine. First, MP2/cc-pVTZ (second-order Møller−Plesset perturbation theory with the same basis set) proves to be an excellent and preferred choice for generating the underlying (Level B) normal modes of the CMA scheme. Utilizing this Level B within the CMA-0A method reproduces the 1501 benchmark frequencies with a mean absolute error (MAE) of only 0.11 cm^{-1} and an attendant standard deviation of 0.49 cm^{-1}. Second, a convergent CMA-2 method is constituted that allows efficient computation of higher level (Level A) frequencies to any reasonable accuracy threshold by using only Hartree−Fock (HF) and MP2 or density functional theory (DFT) data to generate ξ parameters, which select the sparse off-diagonal force field elements for explicit evaluation at Level A. When Level B = MP2/cc-pVTZ, a cutoff of ξ = 0.02 provides an average maximum absolute error per molecule of only 0.17 cm^{-1} by incurring merely a 33% increase in average cost over CMA-0A. This CMA-2 method also eradicates the 4 problematic CMA-0A outliers of pyridine with even less effort (ξ = 0.04, 22% increase). Finally, the newly developed CMA procedures are shown to be highly successful when applied to 1-(1*H*-pyrrol-3-yl)ethanol, a new test molecule with diverse types of vibration.



G2 test set of 1501 benchmark vibrational frequencies at CCSD(T)/cc-pVTZ

1. INTRODUCTION

The force field for molecular vibrations is an indispensable tool for contemporary quantum chemistry, facilitating investigations on molecular structure, spectroscopy, and thermochemistry. Even after decades of research, transformative advances in the computation of force fields are still ongoing.[1] For highly accurate *ab initio* quantum chemical methods such as coupled cluster singles and doubles with perturbative triples [CCSD(T)], computing the vibrational Hessian fully analytically[2−4] is possible for small systems but rapidly becomes intractable for larger molecules. Alternatively, one can numerically compute the Hessian using finite differences of analytic gradients or single-point energies, coarsely parallelizing the task. In such schemes, numerical differentiation is less resource-demanding and a necessity for theoretical methods for which analytical gradients or Hessians have not been implemented, let alone derived.

The number of single-point energies required to compute the force constants in a Hessian for a polyatomic molecule scales quadratically with the number of atoms in the molecule. Computations of CCSD(T)/cc-pVTZ quadratic force constants for systems with up to 500 basis functions and 75 vibrational degrees of freedom are still currently feasible in our laboratory, but the inherent scaling causes the full process to rapidly become intractable as the size of the system increases further. Rectilinear Cartesian displacement coordinates are often used as a basis for computing the force constants numerically; however, this set of coordinates is not optimal for describing molecular motion on inherently curvilinear potential energy surfaces. Employing intuitive curvilinear internal coordinates as a basis reduces the coupling encountered with rectilinear coordinates, and a litany of research exists on this topic.[5−21] The Natural Internal Coordinates (NICs) of Pulay and co-workers are a trenchant choice for describing molecular vibrations.[22,23]

NICs are linear combinations of chemically motivated internal coordinates, and algorithms exist for their automatic

Received: September 20, 2024
Revised: November 4, 2024
Accepted: November 8, 2024
Published: December 13, 2024

https://doi.org/10.1021/acs.jctc.4c01240
J. Chem. Theory Comput. 2024, 20, 10886−10898

generation.[23] While NICs are a felicitous choice to represent the force field, it is difficult to know *a priori* which force constants are negligible. The subject of vibrational coordinate optimization for the potential energy surface beyond the harmonic approximation is well-researched.[24−26] The optimal coordinates for the harmonic oscillator are the normal modes of vibration, which simultaneously diagonalize both the quadratic kinetic and potential energy matrices. Accordingly, to maximize the sparsity of the force constant matrix, the coordinates selected should target the normal mode basis.

A breakthrough solution for selecting vibrational coordinates is the recently developed Concordant Mode Approach (CMA),[1] whose protocol and notation scheme are fully specified in Section 3 below. The CMA-0A method is a highly accurate starting point centered on the computation of only diagonal force constants at the higher level theory A in a normal mode basis generated by a lower-level theory B. The remarkable accuracy of CMA-0A is demonstrated in the 1580 targeted CCSD(T)/cc-pVTZ benchmark vibrational frequencies of the G2 test set.[27] Utilizing NICs to represent the normal modes, out of the 1580 cases CMA-0A is successful in reducing all but three and seven frequency residuals to less than 2.5 cm^{-1} for the Level B choices B3LYP/6-31G(2*df*,*p*) and CCSD(T)/cc-pVDZ, respectively. Simultaneously, CMA-0A reduces computational times by a factor of 7−10. A question not addressed in the initial study[1] is the choice of preferred Level B theory in the CMA protocol. Therefore, the first section of the present research investigates the overall performance of various Level B methods and makes some explicit recommendations.

The second section of this report focuses on the systematic convergence of the CMA methodology by an astute selection of limited off-diagonal force constants for explicit evaluation at Level A. Our introductory paper[1] formulated a CMA-1 method, where one hand-selects off-diagonal force constants, neglected in CMA-0, to eliminate frequency residuals. When a single off-diagonal force constant per molecule is introduced within CMA-1A(1)[A = CCSD(T)/cc-pVTZ, B = CCSD(T)/cc-pVDZ], all original CMA-0A outlier residuals of the G2 test set are reduced to below 1.2 cm^{-1}. While the aforementioned CMA-1A scheme is impressive, it is not intended to be a generally practical method, because the off-diagonal couplings were judiciously selected with full knowledge of the target Level A force constant matrix. Here we propose CMA-2, an *a priori* method of selecting off-diagonal force constants to manage cost and accuracy. An auxiliary force field is obtained with a method henceforth denoted as Level C at little or no additional cost in the Level B computation. In this paper, we investigate Hartree−Fock (HF) theory as Level C. The Level C force field is transformed into the Level B normal mode basis and then cast into a dimensionless matrix $\boldsymbol{\xi}$. The matrix elements ξ_{ij} above a given threshold correspond to the Level A off-diagonal force constants to be explicitly computed in the CMA basis.

2. COMPUTATIONAL METHODS

The reference geometries optimized at the CCSD(T)/cc-pVTZ[28−30] level of theory, their corresponding quadratic force constants, and the NICs chosen in our first CMA study[1] were also used in the present benchmark research on the G2 test set of molecules. In addition, the CCSD(T)/cc-pVDZ[28−30] and B3LYP/6-31G(2*df*,*p*)[31−36] force constants which were previously computed[1] on top of the CCSD(T)/cc-pVTZ optimized geometries were once again employed with the caveat that symmetry was strictly enforced for all molecules, contrary to the first study where symmetry was only rigorously enforced for molecules belonging to non-Abelian symmetries. Normal modes of vibration that are the sole inhabitants of an irreducible representation will not mix with any other mode and are obtained exactly by all CMA methods. Such modes were retained in the statistical treatment of reference 1 but were omitted here in order improve measures of the efficacy of our CMA methods. For reasons discussed later in this work, pyridine is also omitted from our statistics and analyzed in greater detail in-text, reducing the number of frequencies considered in the G2 test set statistics to 1501. New force fields at the reference geometries were computed with CCSD, MP2,[37] density-fitted MP2 (df-MP2),[38,39] HF, and density-fitted HF (df-HF) electronic wave function methods, also complying with the full symmetry for all molecules.

As before,[1] the cc-pV*X*Z (*X* = D, T) basis sets[40,41] were employed for H and Li-F, while the corresponding cc-pV(*X* +*d*)Z basis sets[42] were chosen for Na-Cl in order to incorporate the flexibility of tight *d*-functions. For brevity, the +*d* designations are assumed but not listed in the basis set notations used here. The Level B quadratic force constants were computed in the NIC basis via finite differences with fourth-order accuracy, and the corresponding vibrational normal modes were computed with the GF-matrix method.[43] A displacement size of 0.01 or 0.005 in the units of the NICs was employed throughout.[43] The Level A reference force constants were transformed from the Cartesian basis to the normal mode basis of Level B to obtain the matrix $\mathbf{F}_{CMA}$. In practical applications, the desired elements of $\mathbf{F}_{CMA}$ would be computed by finite differences of Level A energies. However, to completely eliminate numerical errors in the current benchmark research, we directly obtained CMA frequencies by zeroing out off-diagonal elements of $\mathbf{F}_{CMA}$ according to the CMA protocol being employed. The Molpro[44] program was used to compute any conventional CC, MP2, or HF energies, and the SCF and CC energy residuals were reduced to 10^{-12} Hartrees upon convergence. The df-MP2 and df-HF results were computed with Psi4,[45] and the SCF energy residuals were converged to 10^{-10} Hartrees in these cases. Core orbitals were frozen in all coupled cluster and MP2 computations. The optimized geometries and vibrational frequencies of the CCSD(T)/cc-pVTZ target theory, as well as the NICs employed, are compiled in the Supporting Information of this work.

3. CMA PROTOCOL

The most general CMA protocol[1] can be summarized as follows:

1. Choose a complete (preferably chemically intuitive) set of nonredundant internal coordinates **S**.
2. Optimize the molecular geometry for computing the higher level theory A vibrational frequencies and for constructing the concordant modes at a lower-level of theory B, as denoted by CMA[Level A, Level B].
3. Solve the $\mathbf{GF}_B$ eigenproblem[43] after obtaining the lower-level force constants $\mathbf{F}_B$. The nonorthogonal $\mathbf{L}_B$ eigenvector tensor yields the concordant normal modes ($\mathbf{Q}_B$) according to $\mathbf{S} = \mathbf{L}_B\mathbf{Q}_B$.
4. Adopt the ansatz $\mathbf{F}_A = (\mathbf{L}_B^{-1})^T\mathbf{F}_{CMA}(A)\mathbf{L}_B^{-1}$, where $\mathbf{F}_A$ and $\mathbf{F}_{CMA}(A)$ are the force constants at higher level theory A

https://doi.org/10.1021/acs.jctc.4c01240
J. Chem. Theory Comput. 2024, 20, 10886−10898

Table 1. Summary Statistics within the G2 Test Set[a] of CMA-0A Residuals for CCSD(T)/cc-pVTZ Harmonic Vibrational Frequencies (ϵ, cm^{-1}) and ZPVEs (Δ, cm^{-1}) as Compared to the Corresponding Pure Level B Errors

Level B		MAE ϵ	mean ϵ	ϵ_{max}[b]	σ_ϵ	ϵ_{MAX}[c]	MAE Δ	mean Δ	σ_Δ
MP2/cc-pVTZ[d]	pure	9.44	5.62	28.2	13.68	145	40.6	38.3	39.1
	CMA	0.11	0.02	0.54	0.49	9.3	0.12	0.12	0.24
df-MP2/cc-pVTZ[d,e]	pure	9.00	5.19	27.2	12.55	146	37.9	35.6	32.0
	CMA	0.11	0.02	0.53	0.49	9.3	0.11	0.11	0.24
HF/cc-pVTZ	pure	56.92	54.95	94.7	34.41	251	374.4	371.5	294.7
	CMA	1.09	0.14	5.18	4.93	86.3	0.95	0.95	1.56
B3LYP/6-31G(2*df,p*)	pure	8.82	4.66	23.1	10.86	69	36.0	31.5	33.3
	CMA	0.16	0.04	0.59	0.36	3.4	0.27	0.27	0.32
CCSD(T)/cc-pVDZ	pure	47.54	47.12	120.0	54.46	173	319.0	318.6	244.6
	CMA	0.22	0.04	0.74	0.45	4.0	0.27	0.27	0.29
CCSD/cc-pVDZ	pure	54.34	54.05	128.2	53.55	165	365.5	365.5	267.4
	CMA	0.23	0.04	0.77	0.46	3.6	0.29	0.29	0.29
MP2/cc-pVDZ[d]	pure	52.52	52.15	139.4	60.38	213	355.3	355.3	253.4
	CMA	0.32	0.06	1.18	0.83	11.6	0.38	0.38	0.41
HF/cc-pVDZ	pure	85.11	83.88	132.8	47.55	247	567.8	567.1	440.4
	CMA	0.72	0.12	3.45	3.02	58.4	0.81	0.81	1.17

[a]The instructive case of pyridine is omitted from these statistics and dissected in a later section, leaving 1501 frequencies in the data set. [b]Average maximum absolute ϵ per molecule. [c]Maximum absolute ϵ over the entire data set. [d]Excluding $\omega_3(O_3)$, for which MP2 is a catastrophic failure.[48] [e]Excluding $\omega_3(NO_2)$, for which df-MP2 in Psi4 does not run successfully.

in the **S** and $\mathbf{Q}_B$ basis sets, respectively. Minimally, the elements to be included in $\mathbf{F}_{CMA}(A)$ are the diagonal elements, whereas the transformation becomes exact if all elements are included.

5. Adopt a protocol (CMA-*N*) that specifies which elements of $\mathbf{F}_{CMA}(A)$ to explicitly compute via finite differences at the higher level of theory; the CMA-0, CMA-1, and CMA-2 variants discussed below are those currently available. A chosen displacement $\delta\mathbf{Q}_B$ can be mapped into a final set of Cartesian coordinates by iterative application of the linear relationship $\delta\mathbf{x} = \mathbf{u}\mathbf{B}_A^T\mathbf{G}_A^{-1}\mathbf{L}_B(\delta\mathbf{Q}_B)$, where $\mathbf{B}_A$ is the customary B matrix for the **S** coordinates at the A reference geometry and **u** is a diagonal matrix of reciprocal atomic masses.[46,47]
6. Transform the resulting $\mathbf{F}_{CMA}(A)$ to $\mathbf{F}_A$, and then solve the $\mathbf{GF}_A$ eigenproblem to find the higher level A harmonic vibrational frequencies and corresponding normal modes.

In the CMA protocols labeled as (CMA-*N*A, CMA-*N*B), the Level B force constants are computed on top of the (Level A, Level B) optimum geometries. When necessary for clarity, the various levels of theory are appended in brackets to the end of the overarching protocol, such as CMA-0A[Level A, Level B] or CMA-2A[Level A, Level B, Level C]. All CMA variants require the diagonal elements of $\mathbf{F}_{CMA}(A)$ to be included, and the CMA-0A and CMA-0B variants include only these. In our first study,[1] CMA-0A was found to be more accurate than CMA-0B, since the underlying reference geometry is essential to the accurate prediction of second- and higher-order force constants.[46] The CMA-1(*n*) variant only differs from CMA-0 in that *n* hand-selected off-diagonal force constants are also chosen to be computed along with the diagonals. In the convergent CMA-2 method, the dimensionless matrix $\boldsymbol{\xi}$ is constructed via eq 1 from the Level C force field transformed to the $\mathbf{Q}_B$ normal mode basis $[\mathbf{F}_{CMA}(C)]$:

$$\xi_{ij} = \frac{\left|F_{ij,CMA}(C)\right|}{\sqrt{\left|F_{ii,CMA}(C)F_{jj,CMA}(C)\right|}} \quad (1)$$

In carrying out step 5 of the general protocol for the CMA-2 case, all matrix elements ξ_{ij} greater than a user-given threshold correspond to matrix elements of $\mathbf{F}_{CMA}(A)$ that will be explicitly computed.

4. BENCHMARKING LEVEL B FOR CMA-0A

The CMA-0A procedure was executed for eight Level B theories, as summarized in Table 1. The deviation of the pure Level B results from the CCSD(T)/cc-pVTZ target is generally reduced by at least an order of magnitude for every summary statistic when CMA-0A is applied. The mean absolute error (MAE) is the same statistic designated as the mean absolute deviation (MAD) in our prior work;[1] MAE is used here to avoid confusion with common statistical definitions of MAD. The MAE of the CMA-0A residuals ranges from 0.11 to 1.09 cm^{-1} for all Level B choices tested. The mean CMA residuals fall within a remarkably narrow range of 0.02 to 0.14 cm^{-1}, and the standard deviation (σ_ϵ) ranges from 0.36 to 4.93 cm^{-1}. The maximum absolute error (ϵ_{MAX}) over the entire G2 test set is between 3.4 and 11.6 cm^{-1} for all CMA computations with correlated Level B theories, whereas Hartree−Fock proves to perform much more poorly in this regard. The MAEs for the zero-point vibrational energy (ZPVE) residuals (Δ) are a remarkably tight group of minuscule values across the Level B spectrum, ranging from 0.11 to 0.95 cm^{-1}; similarly, the standard deviation (σ_Δ) of the ZPVE Δ lies between 0.24 and 1.56 cm^{-1}. As observed earlier[1] a beneficial cancellation of individual frequency errors provides exceptional accuracy for the CMA-0A ZPVE values.

For a deeper analysis, we first focus on the CMA-0A results in which a cc-pVDZ basis set was used for Level B. Both CMA-0A with B = CCSD/cc-pVDZ and B = CCSD(T)/cc-pVDZ perform very well, and there is no significant statistical difference differentiating the two. The CMA-0A MP2/cc-pVDZ statistics are almost as good as the coupled cluster results, making MP2 a promising candidate as a Level B theory which scales as N^5 with basis set size, compared to the N^6 and N^7 scaling of conventional CCSD and CCSD(T), respectively. The CMA HF/cc-pVDZ results do not compare as well,

https://doi.org/10.1021/acs.jctc.4c01240
J. Chem. Theory Comput. 2024, 20, 10886−10898

Table 2. Only Cases within the Set of 1501 Benchmark Frequencies with CMA-0A[CCSD(T)/cc-pVTZ, MP2/cc-pVTZ] Residuals (ϵ, cm^{-1}) Greater than 1.5 cm^{-1} in Magnitude, Together with the Corresponding CMA-1A and CMA-2A Results That Target These Outliers

Molecule and mode	Description	Benchmark	ϵ[CMA-0A]	ϵ[CMA-1A(1)]	ϵ[CMA-2A](n)[a]	η[b] (%)
nitrous oxide, $\omega_3(\sigma)$	sym. stretch	1297.09	9.34	0.00	0.00 (1)	25
ketene, $\omega_9(b_2)$	C═C═O bend	514.86	7.12	0.00	0.00 (2)	22
ketene, $\omega_8(b_2)$	CH_2 wag	584.48	−6.28	0.00	0.00 (2)	22
nitrous oxide, $\omega_1(\sigma)$	antisym. stretch	2282.57	−5.33	0.00	0.00 (1)	25
benzene, $\omega_{18}(b_{1u})$	ring def.	1158.88	4.90	0.00	0.00 (10)	36
benzene, $\omega_{17}(b_{1u})$	ring def.	1328.17	−4.29	0.00	0.00 (10)	36
spiropentane, $\omega_4(a_1)$	ring breathing	1054.41	3.67	0.00	0.03 (7)	21
spiropentane, $\omega_3(a_1)$	sym. CH_2 wag	1076.02	−3.60	0.01	0.01 (7)	21
methyl nitrite, $\omega_{12}(a)$	N−O stretch	595.01	3.55	0.68	−0.12 (9)	60
nitromethane, $\omega_4(a)$	NO_2 rock	1642.57	−2.87	−1.11	0.04 (13)	87
nitromethane, $\omega_{10}(a)$	CH_3 rock	1110.76	2.71	0.12	0.12 (13)	87
n−butane, $\omega_{28}(b_u)$	CH_3 stretch	3028.05	1.81	0.00	1.81 (10)	28
n−butane, $\omega_{29}(b_u)$	CH_2 stretch	3023.91	−1.79	−0.03	−1.79 (10)	28
aziridine, $\omega_1(a')$	CH_2 rock	783.89	1.76	1.19	1.11 (12)	67
isobutane, $\omega_3(a_1)$	CH stretch	3019.64	1.62	0.02	1.62 (5)	14
isobutane, $\omega_2(a_1)$	CH_3 sym. stretch	3024.21	−1.56	0.04	−1.56 (5)	14

[a] n = the number of off-diagonal elements included for the ξ = 0.020 cutoff. [b] η = number of $\mathbf{F}_{CMA}(A)$ off-diagonal elements included as a percentage of the vibrational degrees of freedom for the given molecule.

displaying a standard deviation (3.02 cm^{-1}) and maximum absolute residual (58.4 cm^{-1}) that are considerably larger than the corresponding coupled cluster results. Apparently, even a minimal treatment of electron correlation in Level B is important in CMA computations targeting CCSD(T)/cc-pVTZ frequencies. Nonetheless, HF may still be an apt Level B for obtaining ZPVEs, as its MAE is only 0.81 cm^{-1}.

We now turn our attention to the CMA-0A results for which the larger cc-pVTZ and 6-31G(2*df*,*p*) basis sets are used. Selecting Level B = MP2/cc-pVTZ gives the best MAE (0.11 cm^{-1}) and mean error (+0.02 cm^{-1}). The conventional and density-fitted MP2 results are nearly indistinguishable, a quite promising finding because df-MP2 scales as N^4 with basis set size, compared to the N^5 scaling of conventional MP2. The summary statistics for the HF residuals once again show that electron correlation should be included at Level B, but not necessarily in ZPVE computations. While B3LYP/6-31G-(2*df*,*p*) yields excellent CMA-0A results, this DFT method is outperformed by both MP2/cc-pVTZ and df-MP2/cc-pVTZ for the MAE ϵ, mean ϵ, ϵ_{max}, and ZPVE statistics. The only caveat is that MP2 and df-MP2 tend to exhibit a few larger maximum absolute residuals; for example, in six out of 1501 cases this quantity is greater than 4 cm^{-1} for MP2/cc-pVTZ. Overall, the summary statistics indicate that basis set quality is more important than higher-order treatment of electron correlation in the selection of Level B. Pending further research, the key conclusion is that if one is targeting CCSD(T)/cc-pV*X*Z frequencies with the CMA-0A approach, then MP2/cc-pV*X*Z or its df-variant is the preferred Level B method.

5. DEVELOPMENT OF CMA-2A THEORY

For the purpose of developing a convergent CMA-2 theory, in this section we consider a CMA outlier as any frequency residual greater than 1.5 cm^{-1} in magnitude. Such outliers for CMA-0A[CCSD(T)/cc-pVTZ, MP2/cc-pVTZ] are presented in Table 2, where 16 cases are found from 9 distinct molecules out of the test set of 1501 frequencies. The size of these residuals ranges from 1.56 to 9.34 cm^{-1}. Pyridine presented a unique challenge to CMA-0A with four outliers at −22.61, −6.46, 4.87, and 27.60 cm^{-1}. Because this molecule proved to be so abnormal, it is excluded from the statistics in Table 2 and treated in greater detail below. When the vibrational modes highlighted in Table 2 are treated by CMA-1A(1) (fifth column), all absolute residuals are less than 1.2 cm^{-1}. In other words, the inclusion of a single off-diagonal force constant per molecule is sufficient to make all outliers disappear! This mathematical fact shows that CMA-0A is tantalizingly close to being a flawless method for this limited test set. However, finding the optimal mathematical solution without prior knowledge of the full force constant matrix remains a considerable challenge. There may be circumstances in which chemical principles can effectively guide one to choose the correct mode couplings, but the development of a generally applicable, automated approach is clearly warranted.

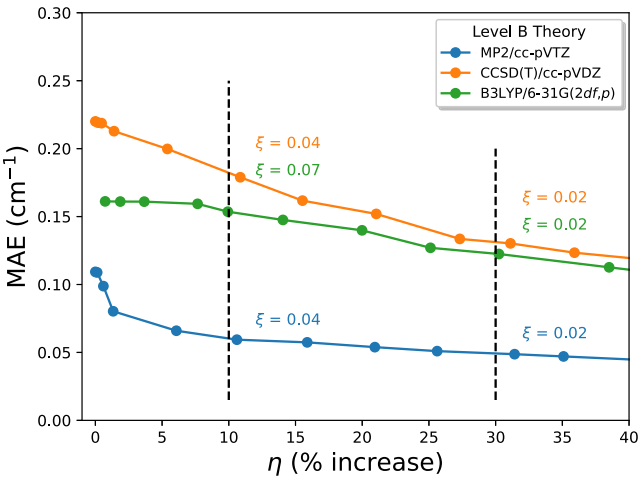


Figure 1. CMA-2A MAE of the 1501 CCSD(T)/cc-pVTZ benchmark frequencies plotted as a function of η, where Level C = HF and the basis set is the same as employed in Level B.

An algorithm for automatically selecting off-diagonal force constants should have a variable tolerance so that the user has

https://doi.org/10.1021/acs.jctc.4c01240
J. Chem. Theory Comput. 2024, 20, 10886−10898

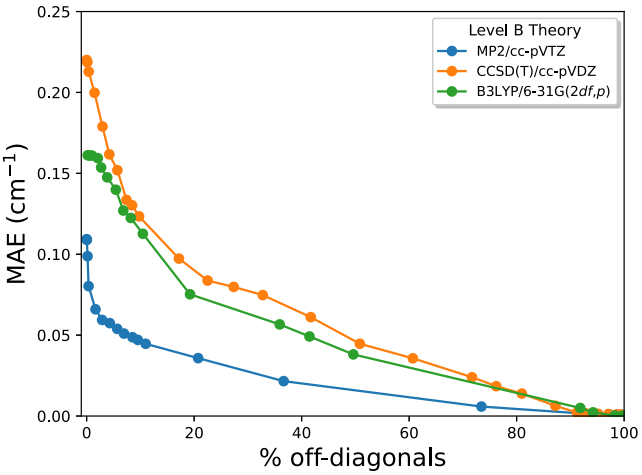


Figure 2. CMA-2A MAE of the 1501 CCSD(T)/cc-pVTZ benchmark frequencies plotted as a function of % nonzero off-diagonal matrix elements included in $\mathbf{F}_{CMA}(A)$.

ultimate control over the cost and accuracy of a convergent CMA protocol. A diagnostic can be imagined with matrix elements that correctly predict the importance of off-diagonal force constants of $\mathbf{F}_{CMA}(A)$. A simple approach would be to compute another set of force constants at a level of theory C in the $\mathbf{Q}_B$ normal mode basis, such that the magnitude of $\mathbf{F}_{CMA}(C)$ off-diagonal force constants mimics the ordering of the $\mathbf{F}_{CMA}(A)$ elements. Preferably, this third Level C should scale better than theory B, and the ideal circumstance would not incur any additional computational cost. If B = MP2/cc-pV*X*Z single-point energies are computed to target A = CCSD(T)/cc-pV*X*Z frequencies using CMA, then the C = HF/cc-pV*X*Z quadratic force field can indeed be obtained without extra cost.

With this choice, the level of theory manifestly increases in the sequence C → B → A. Because $\mathbf{F}_{CMA}(B)$ is a strictly diagonal matrix by construction, then the $\mathbf{F}_{CMA}(C)$ off-diagonal elements of greatest size directly indicate which couplings in the molecule are most sensitive to the C → B increase in level of theory. These are precisely the couplings that are expected to yield significant nonzero values for the off-diagonal elements of $\mathbf{F}_{CMA}(A)$ as a consequence of the B → A change in level of theory. In this sense, $\mathbf{F}_{CMA}(C)$ can be considered a photographic negative of $\mathbf{F}_{CMA}(A)$, although this characterization is not meant to imply that the coupling elements in these matrices are always opposite in sign. In order to translate this correspondence into a viable, universal diagnostic for identifying which parts of $\mathbf{F}_{CMA}(A)$ should be explicitly computed, the $\mathbf{F}_{CMA}(C)$ off-diagonal elements should be cast into a dimensionless form that also takes into account their size relative to the associated diagonal force constants. For this purpose we have adopted eq 1 to define the ξ_{ij} diagnostics for our CMA-2 theory.

For the remainder of the study, when Level B is a correlated wave function method such as CCSD or MP2 with a cc-pV*X*Z basis set, then Level C is chosen as HF/cc-pV*X*Z so that C → B → A forms a series of increasing electron correlation with a fixed basis set. When B = B3LYP/6-31G(2*df*,*p*), we have employed C = HF/6-31G(2*df*,*p*) for lack of a more transparent choice, even though extra cost is incurred because a second set of SCF orbitals must be optimized. The CMA-2A protocol that incorporates all $\mathbf{F}_{CMA}$ force constants corresponding to $\xi > 0.020$ is assessed in Table 2. Overall, 12 of the 16 outliers are successfully eliminated by CMA-2A ($|\epsilon| < 1.11$ cm^{-1}). The remaining four residuals are hardly significant, corresponding to the (1.81, −1.79) and (1.62, −1.56) cm^{-1} pairs for (ω_{28}, ω_{29}) and (ω_3, ω_2) of *n*-butane and isobutane, respectively. For the (*n*-butane, isobutane) C−H stretching modes, the ξ cutoff identified (10, 5) off-diagonal elements for explicit evaluation at Level A, but the key coupling discovered in CMA-1A(1) that eliminates these outliers escaped inclusion. The η parameter appearing in Table 2 is equal to the number of included off-diagonal $\mathbf{F}_{CMA}$ elements as a percentage of the molecular vibrational degrees of freedom that are present. This metric turns out to also be very nearly the percentage increase in computational cost in going beyond CMA-0A to CMA-2A. Demonstrating the success of CMA-2A with a $\xi = 0.020$ cutoff, 8 of the outliers are eliminated with $\eta < 37\%$, and in the remaining instances $\eta = 60-87\%$. Given that CMA-0A already reduces the cost of Level A frequency computations by 700−1000%,[1] any η value less than 100% still constitutes a rather marginal price to pay for the added certainty afforded by CMA-2A.

Table 3. Metrics for CMA-2A Performance of the 1501 CCSD(T)/cc-pVTZ Benchmark Frequencies as a Function of the ξ Cutoff and the Choice of Level B[a]

	B = MP2/cc-pVTZ					B = CCSD(T)/cc-pVDZ					B = B3LYP/6-31G(2*df*,*p*)				
ξ Cutoff	MAE ϵ	ϵ_{max}	σ_ϵ	ϵ_{MAX}	η (%)	MAE ϵ	ϵ_{max}	σ_ϵ	ϵ_{MAX}	η (%)	MAE ϵ	ϵ_{max}	σ_ϵ	ϵ_{MAX}	η (%)
∞	0.11	0.54	0.49	9.34	0	0.22	0.74	0.45	3.96	0	0.16	0.59	0.36	3.42	0
0.20	0.099	0.45	0.40	7.12	0.6	0.22	0.73	0.45	3.96	0.5	0.16	0.58	0.36	3.42	7.7
0.18	0.099	0.45	0.40	7.12	0.7	0.22	0.73	0.45	3.96	0.5	0.16	0.58	0.36	3.42	7.8
0.16	0.093	0.40	0.37	7.12	0.8	0.22	0.73	0.45	3.96	0.6	0.16	0.58	0.36	3.42	7.9
0.14	0.093	0.40	0.37	7.12	0.8	0.22	0.73	0.45	3.96	0.9	0.16	0.58	0.35	3.42	8.1
0.12	0.081	0.33	0.25	3.67	1.0	0.22	0.72	0.45	3.96	1.0	0.16	0.58	0.35	3.42	8.4
0.10	0.080	0.32	0.25	3.67	1.3	0.21	0.72	0.44	3.96	1.4	0.16	0.56	0.35	3.42	8.8
0.08	0.079	0.32	0.25	3.67	1.9	0.21	0.73	0.44	3.96	1.8	0.15	0.55	0.35	3.42	9.3
0.06	0.070	0.27	0.21	3.55	4.3	0.21	0.71	0.44	3.96	3.1	0.15	0.55	0.35	3.42	11
0.04	0.059	0.22	0.18	2.78	11	0.19	0.62	0.39	2.87	8.2	0.15	0.53	0.34	3.42	14
0.02	0.047	0.17	0.15	1.81	33	0.13	0.40	0.28	2.56	29	0.12	0.39	0.28	2.77	33
0.01	0.036	0.14	0.13	1.81	76	0.10	0.29	0.23	2.66	63	0.08	0.25	0.20	2.77	71
0.005	0.022	0.10	0.11	1.81	135	0.07	0.24	0.20	2.68	121	0.06	0.20	0.18	2.77	132

[a]See footnotes to Tables 1 and 2 for definitions of the metrics. Level C = HF with the same basis used for Level B.

https://doi.org/10.1021/acs.jctc.4c01240
J. Chem. Theory Comput. 2024, 20, 10886−10898

The performance of CMA-2A for the entire data set of CCSD(T)/cc-pVTZ benchmark frequencies is illustrated in Figure 1, where the MAE is plotted vs η. It is abundantly clear from the curves that MP2/cc-pVTZ (blue) is once again superior to both CCSD(T)/cc-pVDZ (orange) and B3LYP/6-31G(2*df*,*p*) (green); hence, B = MP2/cc-pVTZ is the method of choice for both CMA-0A and CMA-2A. In this case an increase in cost of only 10% is sufficient to lower the MAE to a mere 0.06 cm^{-1}! The corresponding blue curve has an underlying ξ = 0.040 cutoff at η = 10%, and any further reduction of ξ improves the MAE very little. In contrast, the B = CCSD(T)/cc-pVDZ curve reveals a steady MAE reduction as η ranges up to 40% and as the ξ cutoff is diminished to 0.020. While not optimal, this choice of Level B still provides excellent final results once CMA-2A is applied. Although B = B3LYP/6-31G(2*df*,*p*) achieves much success in CMA-0A, the lag in further improvement via CMA-2A is somewhat disappointing, presumably because the C → B → A series is not systematic in this circumstance.

It is worth emphasizing that the maximum value of η in Figure 1 still corresponds to a very small fraction of off-diagonal elements included in $\mathbf{F}_{CMA}$. In order to reveal the asymptotic nature of CMA-2A, the MAE over the entire data set is plotted against the percent of included nonzero $\mathbf{F}_{CMA}$ off-diagonals in Figure 2. Of course, all the MAE curves must decay to zero when 100% of the off-diagonals are explicitly computed. For all intents and purposes, this asymptotic value is reached at the 90% mark on the abscissa, regardless of the choice of Level B. The key point of Figure 2 is that CMA-2A is truly a convergent method, and the smooth and rapid decay of the blue curve for B = MP2/cc-pVTZ shows how striking the performance can be. Statistics for the overall performance of CMA-2A on the benchmark frequencies are collected in Table 3, where MAE ϵ, σ_ϵ, ϵ_{max}, ϵ_{MAX}, and η are given as the ξ cutoff ranges from 0.005 to ∞ and as Level B is varied. These data allow the CMA-2A user to make a wise choice for ξ and Level B in accord with the accuracy goals and computational costs of the chemical application at hand. A salient feature of Table 3 is that the MAE, standard deviation, and largest residual per molecule all converge much more rapidly as the ξ cutoff is made more stringent when B = MP2/cc-pVTZ as compared to B = CCSD(T)/cc-pVDZ or B3LYP/6-31G(2*df*,*p*). This observation further amplifies the conclusions from Figures 1 and 2 on the selection of Level B. Because MAE ϵ, σ_ϵ, and ϵ_{max} are already much less than 1 cm^{-1} when no off-diagonal elements of $\mathbf{F}_{CMA}$ are accounted for (ξ cutoff = ∞), CMA-0A will be more than sufficient in a preponderance of applications. However, a very low probability remains that an isolated frequency error might occur on the order of ϵ_{MAX}. Informed by Table 3, the user can apply CMA-2A to eliminate this concern, while still maintaining vast savings in computing the harmonic frequencies of the molecule at higher level A. For example, application of CMA-2A with our recommended Level B = MP2/cc-pVTZ and a ξ cutoff of 0.02 yields truly outstanding accuracy while only increasing the average cost by 33% over CMA-0A. The convincing residual statistics in this case can be spelled out as follows: mean absolute error = 0.047 cm^{-1}, average maximum absolute residual per molecule = 0.17 cm^{-1}, standard deviation of residuals = 0.15 cm^{-1}, and global maximum absolute residual = 1.81 cm^{-1}.

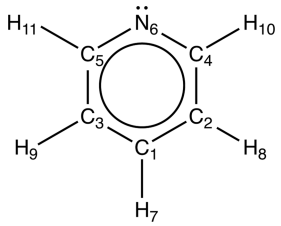


Figure 3. Enumerated carbon backbone of pyridine, corresponding to selected internal coordinates utilized for all CMA computations.

6. SOLUTION OF THE PYRIDINE PROBLEM

The unusual behavior of pyridine in CMA applications demands that this molecule be analyzed in greater detail as an instructive, isolated benchmark species rather than as an outlier that excessively skews the statistics of the G2 test set. Pyridine is depicted in Figure 3 with atomic labeling, and corresponding NICs for our CMA analyses are defined in Table 4. The ring coordinates are constructed with the 6-fold symmetry of a regular hexagon, but the CMA frequencies would be mathematically invariant if linear combinations were made only with the C_{2v} symmetry of pyridine itself.

Comprehensive results for pyridine frequencies are listed in Table 5, where residuals for all modes with respect to the CCSD(T)/cc-pVTZ reference values are provided for seven different treatments: CMA-0A with B = CCSD(T)/cc-pVDZ, MP2/cc-pVTZ, B3LYP/6-31G(2*df*,*p*), CCSD/cc-pVTZ, and MP4/cc-pVTZ, as well as CMA-2A with B = MP2/cc-pVTZ (ξ = 0.04) and CCSD(T)/cc-pVDZ (ξ = 0.015). First we focus on the CMA-0A outlier residuals greater than 1.5 cm^{-1} in magnitude that are highlighted in boldface in Table 5. With B = CCSD(T)/cc-pVDZ, five such residuals appear of modest size, (ϵ_8, ϵ_9, ϵ_{16}, ϵ_{17}, ϵ_{25}) = (−3.24, 4.61, −1.58, 2.46, 2.50) cm^{-1}. With B = MP2/cc-pVTZ, a clustered set of four much larger residuals crops up, (ϵ_{22}, ϵ_{23}, ϵ_{24}, ϵ_{25}) = (−6.46, −22.61, 27.60, 4.87) cm^{-1}, despite the fact this method is the clear overall champion for CMA applications within the G2 test set. One infers that both orbital basis set and electron correlation effects are unusually important for the normal modes of pyridine, because there is little overlap among the cases for which these two Level B methods struggle. Quite surprisingly, when B = B3LYP/6-31G(2*df*,*p*) all the CMA-0A residuals ($|\epsilon|$ < 0.8 cm^{-1}) are well below the outlier threshold, and thus this method happens to capture both the basis set and correlation effects present in the CCSD(T)/cc-pVTZ normal modes of pyridine. While CMA-0A with B = CCSD/cc-pVTZ generally performs better than with B = MP2/cc-pVTZ, the large outliers (ϵ_{24}, ϵ_{25}) = (−18.15, 19.79) cm^{-1} still remain, the latter of which represents a considerable deterioration. This occurrence suggests that the triple-excitation effects described by the (T) correlation correction are an essential component in the physical description of the normal modes of pyridine. While MP4 is a largely defunct method in modern computational chemistry, correlation terms within this theory inspired the (T) ansatz.[28] It is thus compelling that the two problematic outliers for B = CCSD/cc-pVTZ are greatly diminished to (ϵ_{24}, ϵ_{25}) = (−1.56, 2.00) cm^{-1} when B = MP4/cc-pVTZ is invoked.

The outlier normal modes of pyridine are quantitatively characterized by the total energy distributions (TEDs)[49−51] provided in Table 6, where the classical vibrational energy of each mode k is decomposed into the predominant percentage

https://doi.org/10.1021/acs.jctc.4c01240
J. Chem. Theory Comput. 2024, 20, 10886−10898

Table 4. Pyridine Natural Internal Coordinates for CMA Applications

Description	Unnormalized natural internal coordinate[a]
ring breathing	$S_1(a_1) = r(5,3) + r(3,1) + r(1,2) + r(2,4) + r(4,6) + r(6,5)$
ring stretching def.	$S_2(b_2) = r(5,3) - r(3,1) + r(1,2) - r(2,4) + r(4,6) - r(6,5)$
ring stretching def.	$S_3(a_1) = 2r(5,3) - r(3,1) - r(1,2) + 2r(2,4) - r(4,6) - r(6,5)$
ring stretching def.	$S_4(b_2) = 2r(5,3) + r(3,1) - r(1,2) - 2r(2,4) - r(4,6) + r(6,5)$
ring stretching def.	$S_5(a_1) = r(3,1) + r(1,2) - r(4,6) - r(6,5)$
ring stretching def.	$S_6(b_2) = r(3,1) - r(1,2) + r(4,6) - r(6,5)$
sym. CH stretch	$S_7(a_1) = r(5,11) + r(4,10)$
antisym. CH stretch	$S_8(b_2) = r(5,11) - r(4,10)$
sym. CH stretch	$S_9(a_1) = r(3,9) + r(2,8)$
antisym. CH stretch	$S_{10}(b_2) = r(3,9) - r(2,8)$
CH stretch	$S_{11}(a_1) = r(1,7)$
ring stellation	$S_{12}(a_1) = \theta(4,6,5) - \theta(6,5,3) + \theta(5,3,1) - \theta(3,1,2) + \theta(1,2,4) - \theta(2,4,6)$
ring rectangulation	$S_{13}(a_1) = 2\theta(4,6,5) - \theta(6,5,3) - \theta(5,3,1) + 2\theta(3,1,2) - \theta(1,2,4) - \theta(2,4,6)$
ring shearing	$S_{14}(b_2) = \theta(6,5,3) - \theta(5,3,1) + \theta(1,2,4) - \theta(2,4,6)$
sym. i.p. CH rock	$S_{15}(a_1) = \theta(11,5,6) - \theta(11,5,3) + \theta(10,4,6) - \theta(10,4,2)$
antisym. i.p. CH rock	$S_{16}(b_2) = \theta(11,5,6) - \theta(11,5,3) - \theta(10,4,6) + \theta(10,4,2)$
sym. i.p. CH rock	$S_{17}(a_1) = \theta(9,3,5) - \theta(9,3,1) + \theta(8,2,4) - \theta(8,2,1)$
antisym. i.p. CH rock	$S_{18}(b_2) = \theta(9,3,5) - \theta(9,3,1) - \theta(8,2,4) + \theta(8,2,1)$
i.p. CH rock	$S_{19}(b_2) = \theta(7,1,2) - \theta(7,1,3)$
chair ring pucker	$S_{20}(b_1) = \tau(6,5,3,1) + \tau(3,1,2,4) - \tau(5,3,1,2) - \tau(1,2,4,6) + \tau(2,4,6,5) - \tau(4,6,5,3)$
boat ring pucker	$S_{21}(b_1) = \tau(3,1,2,4) - \tau(5,3,1,2) - \tau(2,4,6,5) + \tau(4,6,5,3)$
ring twist	$S_{22}(a_2) = 2\tau(6,5,3,1) - \tau(5,3,1,2) - \tau(3,1,2,4) + 2\tau(1,2,4,6) - \tau(2,4,6,5) - \tau(4,6,5,3)$
sym. o.o.p. CH wag	$S_{23}(b_1) = \gamma(11,5,3,6) + \gamma(10,4,6,2)$
antisym. o.o.p. CH wag	$S_{24}(a_2) = \gamma(11,5,3,6) - \gamma(10,4,6,2)$
sym. o.o.p. CH wag	$S_{25}(b_1) = \gamma(9,3,1,5) + \gamma(8,2,4,1)$
antisym. o.o.p. CH wag	$S_{26}(a_2) = \gamma(9,3,1,5) - \gamma(8,2,4,1)$
o.o.p. CH wag	$S_{27}(b_1) = \gamma(7,1,2,3)$

[a] $r(i,j) = i-j$ bond distance; $\theta(i,j,k) = i-j-k$ bond angle; $\tau(i,j,k,l)$ = dihedral angle between $i-j-k$ and $j-k-l$ plane; $\gamma(i,j,k,l)$ = signed angle of $i-j$ bond out of the $k-j-l$ plane.

contributions n attributable to the NICs S_i. When B = CCSD(T)/cc-pVDZ, all of the outliers with $|\epsilon| > 1.5$ cm^{-1} are removed by merely adding three off-diagonal force constants that couple the (ω_8, ω_9), (ω_{16}, ω_{17}), and (ω_{24}, ω_{25}) pairs, in accord with our earlier CMA-1A analysis of pyridine.[1] Table 6 shows that these three pairs mostly involve strong remixing of [$S_1(a_1)$, $S_{12}(a_1)$], [$S_{20}(b_1)$, $S_{25}(b_1)$], and [$S_2(b_2)$, $S_{18}(b_2)$, $S_{19}(b_2)$], respectively, which Table 4 in turn identifies as the vibrations (ring breathing, ring stellation), [chair ring pucker, out-of-plane (C_2-H_8) + (C_3-H_9) wag], and [ring stretch deformation, (C_2-H_8) − (C_3-H_9) in-plane rock, (C_1-H_7) in-plane rock]. The cluster of 4 outliers for B = MP2/cc-pVTZ does not exhibit simple remixing vis-à-vis the CCSD(T)/cc-pVTZ target modes. Once again $S_2(b_2)$, $S_{18}(b_2)$, and $S_{19}(b_2)$ are culprits, but now the fray is joined by an alternative ring stretching deformation $S_4(b_2)$ and the (C_4-H_{10}) − (C_5-H_{11}) in-plane rock $S_{16}(b_2)$. The two large outliers for B = CCSD/cc-pVTZ have similar origins as in the MP2 case, but the vibrational interactions are less intricate and mostly restricted to the [$S_2(b_2)$, $S_{18}(b_2)$, $S_{19}(b_2)$] set; accordingly, a single (ω_{24}, ω_{25}) coupling within CMA-1A(1) eliminates these outliers. Collectively, the TED analyses in Table 6 suggest that CMA-0A might generally encounter difficulties with aromatic ring vibrations that couple antisymmetric ring stretching deformations with CH in-plane rocks of the same symmetry. Indeed, our work has also found some moderate residuals up to 5 cm^{-1} in size for these types of vibrations in benzene, pyrrole, and furan.

While pyridine is an abnormal case for which CMA-0A is not fully adequate for some isolated modes, this challenging test molecule is totally vanquished by CMA-2A employing either B = CCSD(T)/cc-pVDZ or B = MP2/cc-pVTZ, as shown in Table 5. With our recommended B = MP2/cc-pVTZ and a ξ cutoff of 0.04, only 6 off-diagonals are chosen, and the cost increase is merely 22%. Nevertheless, the maximum absolute residual is reduced to 1.30 cm^{-1}, while the MAE and standard deviation are 0.22 and 0.39 cm^{-1}, respectively. For B = CCSD(T)/cc-pVDZ, a ξ cutoff of 0.015 is necessary to eliminate all outliers, resulting in 20 chosen off-diagonals and a 74% increase in cost; however, in the end the statistics achieved for the residuals are even better than in the MP2/cc-pVTZ case.

7. APPLICATION TO 1-(1*H*-PYRROL-3-YL)ETHANOL

The methodological advances achieved in this study were put to the test for the vibrations of a large molecule containing several challenging motifs, including a heterocyclic aromatic ring, aromatic N−H and C−H bonds, aliphatic O−H and C−H bonds, methyl and hydroxyl internal rotations, and torsions of entire monomer groups about a central C−C bond mediated by noncovalent interactions. In particular, the CMA-0A, CMA-1A(1), and CMA-2A methods with higher level A = CCSD(T)/cc-pVTZ and lower-level B = MP2/cc-pVTZ were applied to the 1-(1*H*-pyrrol-3-yl)ethanol conformer shown in Figure 4. This molecule lies outside the G2 test set, and very little is known about its vibrational spectrum. Benchmark CCSD(T)/cc-pVTZ harmonic frequencies and corresponding CMA residuals are reported in Table 7. The NICs were constructed by starting with the full set of customary coordinates for the individual P = pyrrole and E

https://doi.org/10.1021/acs.jctc.4c01240
J. Chem. Theory Comput. 2024, 20, 10886−10898

Table 5. CMA-0A and CMA-2A Residuals (ϵ, cm^{-1}) for CCSD(T)/cc-pVTZ Harmonic Frequencies (ω_{ref}, cm^{-1}) of Pyridine Obtained with Various Level B Methods

			ϵ[CMA-0A]					ϵ[CMA-2A]	
Mode	Description	ω_{ref}	B = CCSD(T)/cc-pVDZ	B = MP2/cc-pVTZ	B = B3LYP/6-31G(2*df*,*p*)	B = CCSD/cc-pVTZ	B = MP4/cc-pVTZ	B = CCSD(T)/cc-pVDZ[a]	B = MP2/cc-pVTZ[b]
$\omega_1(a_1)$	sym. CH str.	3212.85	−0.01	−0.05	−0.03	0.00	0.00	−0.01	−0.06
$\omega_2(a_1)$	asym. CH str.	3187.57	−0.01	0.01	−0.08	−0.01	0.00	0.03	0.06
$\omega_3(a_1)$	asym. CH str.	3169.83	0.01	0.05	0.08	0.01	0.01	−0.03	0.02
$\omega_4(a_1)$	ring str. def.	1630.29	−0.47	−0.21	−0.02	−0.01	−0.06	−0.01	−0.18
$\omega_5(a_1)$	sym. CH rock	1510.12	−0.48	0.10	−0.05	0.01	0.04	−0.15	0.10
$\omega_6(a_1)$	asym. CH rock	1236.72	0.50	0.08	0.03	0.00	0.02	−0.02	0.08
$\omega_7(a_1)$	sym. CH rock	1087.86	−0.37	−0.07	0.00	0.00	−0.06	−0.89	−0.07
$\omega_8(a_1)$	ring stellation	1043.14	**−3.24**	0.00	−0.16	0.00	−0.02	1.06	0.00
$\omega_9(a_1)$	ring breathing	1001.24	**4.61**	0.16	0.30	0.01	0.10	0.14	0.17
$\omega_{10}(a_1)$	ring rectangle	603.31	0.10	0.02	0.08	0.01	0.01	0.03	0.02
$\omega_{11}(a_2)$	asym. CH wag	995.11	−0.01	−0.51	−0.04	−0.01	−0.11	−0.01	−0.50
$\omega_{12}(a_2)$	asym. CH wag	890.79	0.00	0.56	0.02	0.01	0.12	0.00	0.56
$\omega_{13}(a_2)$	ring twist	378.44	0.01	0.01	0.06	0.01	0.02	0.01	0.01
$\omega_{14}(b_1)$	asym. CH wag	996.87	−0.69	−0.41	−0.11	−0.06	−0.06	−0.15	−0.16
$\omega_{15}(b_1)$	sym. CH wag	953.64	0.02	0.12	−0.02	0.01	0.03	0.12	0.12
$\omega_{16}(b_1)$	ring chair	753.17	**−1.58**	0.02	0.03	−0.02	−0.14	0.00	−0.30
$\omega_{17}(b_1)$	sym. CH wag	711.89	**2.46**	0.36	0.12	0.08	0.19	0.03	0.36
$\omega_{18}(b_1)$	ring boat	409.22	0.21	0.01	0.03	0.00	0.00	0.03	0.01
$\omega_{19}(b_2)$	asym. CH str.	3204.49	−0.01	−0.03	−0.02	0.00	−0.01	0.00	−0.02
$\omega_{20}(b_2)$	asym. CH str.	3168.42	0.00	0.03	−0.04	0.00	0.01	−0.01	0.00
$\omega_{21}(b_2)$	ring str. def.	1618.43	−0.50	−0.46	−0.01	−0.04	−0.03	−0.25	−0.48
$\omega_{22}(b_2)$	ring str. def.	1464.66	−1.02	**−6.46**	−0.10	−0.17	−0.18	0.17	0.06
$\omega_{23}(b_2)$	sym. CH wag	1379.21	−1.34	**−22.61**	−0.12	−0.04	−0.06	0.04	0.29
$\omega_{24}(b_2)$	ring str. def.	1266.95	0.38	**27.60**	−0.37	**−18.15**	**−1.56**	−0.01	−1.05
$\omega_{25}(b_2)$	asym. CH wag	1158.82	**2.50**	**4.87**	0.75	**19.79**	**2.00**	0.02	1.30
$\omega_{26}(b_2)$	ring str. def.	1071.35	0.65	0.07	0.07	0.05	0.02	0.09	0.07
$\omega_{27}(b_2)$	ring shearing	656.98	0.11	0.01	0.02	0.01	0.00	0.03	0.01

[a] ξ cutoff = 0.015, η = 74%, and n = 20 off-diagonals included. [b] ξ cutoff = 0.04, η = 22%, and n = 6 off-diagonals included.

= ethanol monomers. Conceptually, the hydrogens on C_5 and C_6 were then removed while keeping the bond vectors fixed, whence the entire molecule was assembled by adding the C_5−C_6 distance and the intermonomer torsion angles about this linkage. The resulting 45 NICs for 1-(1*H*-pyrrol-3-yl)ethanol are provided in Table 8, and the Cartesian coordinates of all atoms are given in the Supporting Information.

For all but two frequencies, the performance of CMA-0A is excellent, achieving a MAE and maximum absolute residual (ϵ_{max}) of only 0.24 and 1.61 cm^{-1}, respectively, for 43 of the 45 normal modes. The two prominent outliers (ϵ_{37}, ϵ_{39}) = (10.9, −13.1) cm^{-1} arise from (ω_{37}, ω_{39}) = (480.6, 443.0) cm^{-1}, whose largest TED components occur for the (C_7−C_6−O_8 bend, N−H out-of-plane wag) coordinates. If CMA-1A(1) is applied with a single (ω_{37}, ω_{39}) coupling, then the (ϵ_{37}, ϵ_{39}) residuals are dramatically reduced to (0.04, 1.56) cm^{-1}. What is remarkable about this solution is that the largest contributors for the two modes occur on different monomers, so that the coupling occurs at long-range across the C_5−C_6 linkage. A likely explanation is that the coupling involves noncovalent interactions of the aromatic π cloud with the polar O−H group. In any event, the system clearly provides a rigorous test for CMA methods. When CMA-2A is invoked with ξ = 0.04 (η = 27%, n = 12 couplings), the automatic selection process indeed picks up the (ω_{37}, ω_{39}) coupling, so that the MAE becomes a mere 0.22 cm^{-1} for the entire system of 45 vibrations. The residuals of largest size for CMA-2A(ξ = 0.04) are (ϵ_{11}, ϵ_{39}) = (−1.56, 1.56) cm^{-1}, which barely qualify as

https://doi.org/10.1021/acs.jctc.4c01240
J. Chem. Theory Comput. 2024, 20, 10886−10898

Table 6. Total Energy Distributions (TEDs) for the Outlier Modes of Pyridine at Various Levels of Theory

Mode	Frequency (cm^{-1})	TED[a]
	CCSD(T)/cc-pVTZ	
$\omega_8(a_1)$	1043.14	$S_{12}(51) + S_1(33) + S_5(13)$
$\omega_9(a_1)$	1001.24	$S_1(65) - S_{12}(32)$
$\omega_{16}(b_1)$	753.17	$S_{20}(68) - S_{25}(15) - S_{27}(13)$
$\omega_{17}(b_1)$	711.89	$S_{20}(53) + S_{25}(27) + S_{27}(10) + S_{23}(9)$
$\omega_{22}(b_2)$	1464.66	$S_4(32) - S_{18}(25) + S_{19}(20) + S_{16}(15)$
$\omega_{23}(b_2)$	1379.21	$S_{16}(67) + S_{18}(18) - S_{19}(13)$
$\omega_{24}(b_2)$	1266.95	$S_2(78) + S_6(8) - S_{19}(8) - S_{18}(7)$
$\omega_{25}(b_2)$	1158.82	$S_{18}(36) + S_{19}(35) + S_2(20) - S_6(8)$
	CMA-0A[CCSD(T)/cc-pVTZ, CCSD(T)/cc-pVDZ]	
$\omega_8(a_1)$	1039.90	$S_1(64) + S_{12}(30)$
$\omega_9(a_1)$	1005.85	$S_{12}(61) - S_1(34)$
$\omega_{16}(b_1)$	751.59	$S_{20}(46) - S_{25}(24) - S_{27}(20) - S_{23}(11)$
$\omega_{17}(b_1)$	714.35	$S_{20}(75) + S_{25}(20)$
$\omega_{25}(b_2)$	1161.32	$S_{18}(41) + S_{19}(40) + S_2(9) - S_6(8)$
	CMA-0A[CCSD(T)/cc-pVTZ, MP2/cc-pVTZ]	
$\omega_{22}(b_2)$	1458.20	$S_4(29) - S_{18}(23) + S_{19}(17) + S_{16}(17)$
$\omega_{23}(b_2)$	1356.60	$S_{16}(56) + S_{18}(17) - S_2(15) - S_6(6)$
$\omega_{24}(b_2)$	1294.55	$S_2(70) - S_{19}(13) + S_{16}(9)$
$\omega_{25}(b_2)$	1163.69	$S_{18}(42) + S_{19}(41) - S_6(10) + S_2(6)$
	CMA-0A[CCSD(T)/cc-pVTZ, CCSD/cc-pVTZ]	
$\omega_{24}(b_2)$	1248.80	$S_2(38) - S_{18}(26) - S_{19}(25) + S_6(12)$
$\omega_{25}(b_2)$	1178.61	$S_2(61) + S_{19}(18) + S_{18}(18)$

[a]For each normal mode k, the percentage n of the vibrational energy attributable to internal coordinate i is listed as $S_i(n)$, as calculated from eq 32 of ref 49. The signs preceding these entries denote the relative phases in the normal-mode eigenvectors.

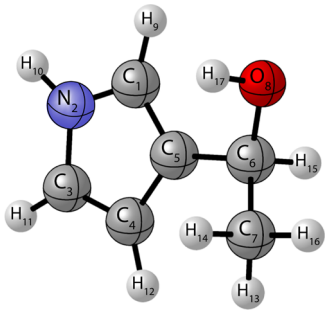


Figure 4. Enumerated atomic structure of 1-(1*H*-pyrrol-3-yl)ethanol, corresponding to selected internal coordinates utilized for all CMA computations.

outliers by our adopted CMA-2A threshold of 1.5 cm^{-1}. While this level of performance is impressive, the application of CMA-2A with $\xi = 0.02$ ($\eta = 140\%$, $n = 63$ couplings) removes any hint of an outlier. Specifically, the (MAE, ϵ_{max}) metrics for this final run are (0.15, 1.2) cm^{-1}.

All computations for 1-(1*H*-pyrrol-3-yl)ethanol were performed using 6 cores with 60 GB of memory on an Intel Xeon Gold 6130 processor in conjunction with a Luster file system. The CMA-0A, CMA-2A($\xi = 0.04$), and CMA-2A($\xi = 0.02$) results required a mere 4.3%, 5.6%, and 10.5%, respectively, of the CPU time for the complete CCSD(T)/cc-pVTZ harmonic frequency computation. In summary, CMA-2A gives a commanding performance for this test molecule. CMA-2A($\xi = 0.04$) cuts the CPU time by a factor of 18 while evaluating all frequencies to better than 1.6 cm^{-1}, whereas CMA-2A($\xi = 0.02$) provides additional assurance against outliers yet still cuts the original cost by almost a factor of 10.

8. SUMMARY

The Concordant Mode Approach (CMA) is a general protocol by which harmonic vibrational frequencies at a higher level theory A can be computed using a normal mode basis generated by a lower-level theory B, with cost reductions approaching an order of magnitude for larger molecules. Our introductory study[1] on the G2 test set showed that CMA-0A can produce vibrational frequencies with at least a 99% probability of being accurate to within 1.5 cm^{-1} but with a 0.2−0.4% chance of outliers greater than 2.5 cm^{-1}. The current research has focused on creating new CMA methods that rapidly and systematically converge to the exact Level A frequencies while also eliminating outliers. Overall, MP2/cc-pVTZ is found to be the best Level B when targeting CCSD(T)/cc-pVTZ frequencies. Our results show the importance of going beyond HF to incorporate electron correlation at Level B, but matching the quality of the Level A orbital basis set becomes paramount thereafter. Another key conclusion for Level B is that density-fitted (df) MP2 performs just as well as conventional MP2; hence, the reduced scaling of df-MP2 provides a wealth of opportunities for CMA applications on large molecules. The CMA-2A procedure developed here is demonstrated to be a robust, convergent method (Figure 2) that selects which off-diagonal force field elements to explicitly evaluate at Level A based on dimensionless ξ parameters that can be evaluated at a Level C with essentially no additional computational cost. Our current recommendation for CMA-2A applied to CCSD(T)/cc-pV*X*Z frequencies is to choose B = MP2/cc-pV*X*Z (or its df-MP2 variant) and C = HF/cc-pV*X*Z with ξ in the 0.02−0.04 range depending on the required certainty of the results. With $\xi = 0.02$ the CMA-2A residuals exhibit extremely small (MAE, σ_ϵ, ϵ_{max}) statistics of (0.05, 0.15, 0.17) cm^{-1}, and the maximum absolute discrepancy (ϵ_{MAX}) over the entire database is only 1.8 cm^{-1}, all achieved with only a 33% increase (η) in average cost over CMA-0A. Moreover, the troublesome pyridine and challenging 1-(1*H*-pyrrol-3-yl)-ethanol vibrational problems can be fully solved by CMA-2A with the less stringent threshold $\xi = 0.04$, resulting in (ϵ_{MAX}, η) = (1.3 cm^{-1}, 22%) and (1.6 cm^{-1}, 27%) for these two molecules, respectively. The hierarchy CMA-*N* is now in place for $N = 0$, 1, 2, and the thoroughgoing success of these methods promises continuing improvements in cost and accuracy, applications to ever-larger molecules, and diversification into intermolecular vibrations. Other advances are also readily envisioned for which CMA-*N* can provide an ideal solution, such as the inclusion of core electron correlation and complete basis set (CBS) extrapolation in the determination of vibrational frequencies for larger molecules.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.4c01240.

CMA-2A (RMSD, ϵ_{max}) versus (η, percent off-diagonals in $\mathbf{F}_{CMA}$) plots. Cartesian coordinates for the CCSD(T)/cc-pVTZ structure of 1-(1*H*-pyrrol-3-yl)ethanol and all molecules of the G2 test set. 1501 pure, CMA-0A, and CMA-2A frequencies at $\xi = 0.02$ with A = CCSD(T)/

https://doi.org/10.1021/acs.jctc.4c01240
J. Chem. Theory Comput. 2024, 20, 10886−10898

Table 7. CMA-(0-2)A Residuals (ϵ, cm^{-1}) for CCSD(T)/cc-pVTZ Harmonic Frequencies (ω_{ref}, cm^{-1}) of 1-(1*H*-Pyrrol-3-yl)ethanol Obtained with B = MP2/cc-pVTZ

			ω_{ref}	ϵ[CMA-0A]	ϵ[CMA-1A(1)]	ϵ[CMA-2A]	ϵ[CMA-2A]
		Monomer[a]	CCSD(T)/cc-pVTZ	B = MP2/cc-pVTZ	B = MP2/cc-pVTZ	B = MP2/cc-pVTZ[b]	B = MP2/cc-pVTZ[c]
ω_1	O−H stretch	E	3808.93	0.00	0.00	0.00	0.00
ω_2	N−H stretch	P	3696.76	0.00	0.00	0.00	0.00
ω_3	C_1−H_9 stretch	P	3276.98	0.03	0.03	0.03	0.03
ω_4	C−H sym. stretch	P	3271.79	−0.01	−0.01	−0.01	−0.01
ω_5	C−H antisym. stretch	P	3248.80	0.04	0.04	0.04	0.04
ω_6	CH_3 antisym. stretch	E	3124.60	−0.04	−0.04	−0.04	−0.04
ω_7	CH_3 sym. stretch	E	3108.32	−0.07	−0.07	−0.07	−0.07
ω_8	C_6−H_{15} stretch	E	3047.93	−0.00	0.00	0.00	0.00
ω_9	CH_3 sym. stretch	E	3031.56	0.05	0.05	0.05	0.05
ω_{10}	ring stretching def.	P	1604.19	−0.70	−0.70	−0.25	−0.15
ω_{11}	ring stretching def.	P	1521.94	**−1.61**	**−1.61**	**−1.56**	−0.11
ω_{12}	CH_3 def.	E	1498.72	−0.10	−0.10	−0.10	−0.10
ω_{13}	CH_3 scissor	E	1495.56	−0.09	−0.09	−0.09	−0.09
ω_{14}	ring stretching def.	P	1475.70	0.61	0.61	0.76	−0.09
ω_{15}	H_{15}−C_6−C_5 wag	E	1429.80	−0.23	−0.23	−0.00	−0.02
ω_{16}	ring stretching def.	P	1408.44	0.04	0.04	0.22	0.33
ω_{17}	CH_3 umbrella	E	1401.83	0.01	0.01	0.01	0.01
ω_{18}	H_{15}−C_6−C_5 rock	E	1365.39	0.33	0.33	0.33	0.34
ω_{19}	C−O−H bend	E	1306.54	−0.29	−0.29	−0.29	−0.27
ω_{20}	C_1−H_9 rock	P	1276.10	0.55	0.55	0.32	0.40
ω_{21}	sym. C−H rock	P	1251.64	0.17	0.17	0.17	0.09
ω_{22}	N−H rock	P	1156.61	0.72	0.72	0.15	0.02
ω_{23}	C−O stretch	E	1125.73	−0.01	−0.01	−0.30	−0.36
ω_{24}	ring breathing	P	1108.99	−0.10	−0.10	−0.10	−0.10
ω_{25}	antisym. C−H rock	P	1088.87	0.67	0.67	0.37	0.54
ω_{26}	CH_3 rock	E	1067.33	0.48	0.47	0.48	0.23
ω_{27}	CH_3 wag	E	1034.92	0.14	0.14	0.14	0.13
ω_{28}	ring shearing	P	958.88	0.06	0.06	0.06	0.06
ω_{29}	ring bending	P	898.10	−0.01	−0.01	−0.01	−0.01
ω_{30}	C_6−C_7 stretch	E	892.03	0.11	0.11	0.11	0.10
ω_{31}	antisym. C−H wag	P	830.69	−0.47	−0.47	−0.21	−0.12
ω_{32}	sym. C−H wag	P	789.78	−0.07	−0.07	−0.07	0.10
ω_{33}	C_5−C_6 E wag	P ∩ E	701.34	−0.42	−0.42	−0.42	−0.03
ω_{34}	antisym. C−H wag	P	696.28	0.52	0.52	0.29	−0.08
ω_{35}	ring puckering	P	635.44	0.20	0.20	0.12	0.06
ω_{36}	ring twisting	P	614.14	0.10	0.10	0.10	0.07
ω_{37}	C_7−C_6−O_8 bend	E	480.63	**−10.90**	0.04	0.04	−0.02
ω_{38}	C_5−C_6 E rock	P ∩ E	457.57	−1.05	−1.05	−1.05	−1.05
ω_{39}	N−H wag	P	442.97	**13.10**	**1.56**	**1.56**	1.20
ω_{40}	C_5−C_6 stretch	P ∩ E	357.34	0.06	0.06	0.06	0.05
ω_{41}	O−H torsion	E	320.88	0.04	0.04	0.04	0.04
ω_{42}	CH_3 torsion	E	261.50	0.02	0.02	0.02	0.02
ω_{43}	C_6−C_5 P rock	P ∩ E	196.99	0.02	0.02	0.02	0.01
ω_{44}	C_6−C_5 P wag	P ∩ E	172.27	0.04	0.04	0.04	0.02
ω_{45}	P−E torsion	P ∩ E	47.91	0.03	0.03	0.03	0.02

[a](P, E) = (pyrrole, ethanol). [b]ξ cutoff = 0.04, η = 26.7%, and n = 12 off-diagonals included. [c]ξ cutoff = 0.02, η = 140.0%, and n = 63 off-diagonals included.

cc-pVTZ, B = [MP2/cc-pVTZ, CCSD(T)/cc-pVDZ, B3LYP/6-31G(2*df*,*p*)], and C = [HF/cc-pVTZ, HF/cc-pVDZ, HF/6-31G(2*df*,*p*)] (PDF)

AUTHOR INFORMATION

Corresponding Author

Henry F. Schaefer III − *Center for Computational Quantum Chemistry and Department of Chemistry, University of Georgia, Athens, Georgia 30602, United States;* orcid.org/0000-0003-0252-2083; Email: qc@uga.edu

Authors

Nathaniel L. Kitzmiller − *Center for Computational Quantum Chemistry and Department of Chemistry, University of Georgia, Athens, Georgia 30602, United States; Indiana Wesleyan University, Marion, Indiana 46953, United States;* orcid.org/0000-0002-0618-6497

Mitchell E. Lahm − *Center for Computational Quantum Chemistry and Department of Chemistry, University of Georgia, Athens, Georgia 30602, United States;* orcid.org/0000-0001-8908-8157

https://doi.org/10.1021/acs.jctc.4c01240
J. Chem. Theory Comput. 2024, 20, 10886−10898

Table 8. 1-(1*H*-Pyrrol-3-yl)ethanol Natural Internal Coordinates

Description	Monomer[a]	Unnormalized natural internal coordinate[b]
ring breathing	P	$S_1 = r(5,4) + r(1,5) + r(4,3) + r(2,1) + r(2,3)$
ring stretching def.	P	$S_2 = r(5,4) + B\ r(1,5) + B\ r(4,3) + A\ r(2,1) + A\ r(2,3)$
ring stretching def.	P	$S_3 = D\ r(1,5) - D\ r(4,3) + C\ r(2,1) - C\ r(2,3)$
ring stretching def.	P	$S_4 = r(5,4) + A\ r(1,5) + A\ r(4,3) + B\ r(2,1) + B\ r(2,3)$
ring stretching def.	P	$S_5 = C\ r(1,5) - C\ r(4,3) - D\ r(2,1) + D\ r(2,3)$
C−C stretch	P ∩ E	$S_6 = r(5,6)$
C−C stretch	E	$S_7 = r(6,7)$
C−O stretch	E	$S_8 = r(6,8)$
C−H stretch	P	$S_9 = r(1,9)$
N−H stretch	P	$S_{10} = r(2,10)$
C−H stretch	P	$S_{11} = r(3,11)$
C−H stretch	P	$S_{12} = r(4,12)$
C−H stretch	E	$S_{13} = r(6,15)$
O−H stretch	E	$S_{14} = r(8,17)$
CH_3 sym. stretch	E	$S_{15} = r(7,13) + r(7,14) + r(7,16)$
CH_3 sym. stretch	E	$S_{16} = 2r(7,13) - r(7,14) - r(7,16)$
CH_3 antisym. stretch	E	$S_{17} = r(7,14) - r(7,16)$
ring bending	P	$S_{18} = \theta(1,2,3) + A\ \theta(2,1,5) + A\ \theta(2,3,4) + B\ \theta(1,5,4) + B\ \theta(3,4,5)$
ring shearing	P	$S_{19} = (A - B)\ \theta(2,1,5) - (A - B)\ \theta(2,3,4) + (1 - A)\ \theta(1,5,4) - (1 - A)\ \theta(3,4,5)$
C_6−C_5 P rock	P ∩ E	$S_{20} = \theta(6,5,4) - \theta(6,5,1)$
C_7−C_6−O bend	E	$S_{21} = \theta(8,6,7)$
C_5−C_6 E wag	P ∩ E	$S_{22} = \theta(5,6,7) + \theta(5,6,8)$
C_5−C_6 E rock	P ∩ E	$S_{23} = \theta(5,6,7) - \theta(5,6,8)$
C−H rock	P	$S_{24} = \theta(9,1,2) - \theta(9,1,5)$
N−H rock	P	$S_{25} = \theta(10,2,3) - \theta(10,2,1)$
C−H rock	P	$S_{26} = \theta(11,3,4) - \theta(11,3,2)$
C−H rock	P	$S_{27} = \theta(12,4,5) - \theta(12,4,3)$
C−O−H bend	E	$S_{28} = \theta(17,8,6)$
H−C−C rock	E	$S_{29} = \theta(15,6,7) + \theta(15,6,8)$
H−C−C wag	E	$S_{30} = \theta(15,6,7) - \theta(15,6,8)$
CH_3 umbrella	E	$S_{31} = \theta(13,7,6) + \theta(14,7,6) + \theta(16,7,6) - \theta(14,7,16) - \theta(13,7,14) - \theta(13,7,16)$
CH_3 rock	E	$S_{32} = 2\theta(13,7,6) - \theta(14,7,6) - \theta(16,7,6)$
CH_3 wag	E	$S_{33} = \theta(14,7,6) - \theta(16,7,6)$
CH_3 scissor	E	$S_{34} = 2\theta(14,7,16) - \theta(13,7,14) - \theta(13,7,16)$
CH_3 def.	E	$S_{35} = \theta(13,7,14) - \theta(13,7,16)$
ring twisting	P	$S_{36} = \tau(1,5,4,3) + B\ \tau(4,3,2,1) + B\ \tau(3,2,1,5) + A\ \tau(5,4,3,2) + A\ \tau(2,1,5,4)$
ring puckering	P	$S_{37} = (1-A)\ \tau(4,3,2,1) - (1-A)\ \tau(3,2,1,5) - (B-A)\ \tau(5,4,3,2) + (B-A)\ \tau(2,1,5,4)$
CH_3 torsion	E	$S_{38} = \tau(13,7,6,8) + \tau(14,7,6,8) + \tau(16,7,6,8)$
O−H torsion	E	$S_{39} = \tau(17,8,6,7)$
P−E torsion	P ∩ E	$S_{40} = \tau(1,5,6,7) + \tau(1,5,6,8) + \tau(4,5,6,7) + \tau(4,5,6,8)$
C_6−C_5 P wag	P ∩ E	$S_{41} = \gamma(6,5,1,4)$
C−H wag	P	$S_{42} = \gamma(9,1,2,5)$
N−H wag	P	$S_{43} = \gamma(10,2,3,1)$
C−H wag	P	$S_{44} = \gamma(11,3,4,2)$
C−H wag	P	$S_{45} = \gamma(12,4,5,3)$

[a](P, E) = (pyrrole, ethanol). [b](A, B, C, D) = $\left[\cos\left(\frac{4}{5}\pi\right), \cos\left(\frac{2}{5}\pi\right), \sin\left(\frac{4}{5}\pi\right), \sin\left(\frac{2}{5}\pi\right)\right]$.

Laura N. Olive Dornshuld − *Center for Computational Quantum Chemistry and Department of Chemistry, University of Georgia, Athens, Georgia 30602, United States*

Jincan Jin − *Center for Computational Quantum Chemistry and Department of Chemistry, University of Georgia, Athens, Georgia 30602, United States;* orcid.org/0000-0002-6123-1115

Wesley D. Allen − *Center for Computational Quantum Chemistry and Department of Chemistry, University of Georgia, Athens, Georgia 30602, United States; Allen Heritage Foundation, Dickson, Tennessee 37055, United States;* orcid.org/0000-0002-4288-2297

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.4c01240

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported by the U.S. Department of Energy, Basic Energy Sciences, Division of Chemistry, Computational and Theoretical Chemistry (CTC) Program, under Contract DE-SC0018412. L.N.O.D. acknowledges the NSF for a Graduate Research Fellowship (NSF Award 1842396).

REFERENCES

(1) Lahm, M. E.; Kitzmiller, N. L.; Mull, H. F.; Allen, W. D.; Schaefer, H. F. Concordant Mode Approach for Molecular Vibrations. *J. Am. Chem. Soc.* **2022**, *144*, 23271−23274.

(2) Kállay, M.; Gauss, J. Analytic Second Derivatives for General Coupled-Cluster and Configuration-Interaction Models. *J. Chem. Phys.* **2004**, *120*, 6841−6848.

(3) Szalay, P. G.; Gauss, J.; Stanton, J. F. Analytic UHF-CCSD(T) Second Derivatives: Implementation and Application to the Calculation of the Vibration-Rotation Interaction Constants of NCO and NCS. *Theor. Chem. Acc.* **1998**, *100*, 5−11.

(4) Wang, F.; Gauss, J. Analytic Second Derivatives in Closed-Shell Coupled-Cluster Theory with Spin-Orbit Coupling. *J. Chem. Phys.* **2009**, *131*, No. 164113.

(5) Gatti, F.; Iung, C.; Menou, M.; Justum, Y.; Nauts, A.; Chapuisat, X. Vector Parametrization of the *N*-atom Problem in Quantum Mechanics. I. Jacobi Vectors. *J. Chem. Phys.* **1998**, *108*, 8804−8820.

(6) Gatti, F.; Muñoz, C.; Iung, C. A General Expression of the Exact Kinetic Energy Operator in Polyspherical Coordinates. *J. Chem. Phys.* **2001**, *114*, 8275−8281.

(7) Gatti, F. Vector Parametrization of the *N*-atom Problem in Quantum Mechanics. III. Separation into Two Subsystems: Application to NH_3. *J. Chem. Phys.* **1999**, *111*, 7225−7235.

(8) Leforestier, C.; Viel, A.; Gatti, F.; Muñoz, C.; Iung, C. The Jacobi-Wilson Method: A New Approach to the Description of Polyatomic Molecules. *J. Chem. Phys.* **2001**, *114*, 2099−2105.

(9) Radau, R. Sur Une Transformation Des Équations Différentielles de La Dynamique. *Ann. Sci. École Norm. Sup.* **1868**, *5*, 311−375.

(10) Smith, F. T. Modified Heliocentric Coordinates for Particle Dynamics. *Phys. Rev. Lett.* **1980**, *45*, 1157−1160.

(11) Wang, X.-G.; Carrington, T., Jr. Vibrational Energy Levels of CH_5^+. *J. Chem. Phys.* **2008**, *129*, No. 234102.

(12) Chapuisat, X. Exact Quantum Molecular Hamiltonians: II. On the Choice of the Moving Frame of Reference. The Principal Axis System. *Mol. Phys.* **1991**, *72*, 1233−1265.

(13) Horn, T. R.; Gerber, R. B.; Valentini, J. J.; Ratner, M. A. Vibrational States and Structure of Ar_3: The Role of Three-Body Forces. *J. Chem. Phys.* **1991**, *94*, 6728−6736.

(14) Chapuisat, X.; Iung, C. Vector Parametrization of the *N*-body Problem in Quantum Mechanics: Polyspherical Coordinates. *Phys. Rev. A* **1992**, *45*, 6217−6235.

(15) Horn, T. R.; Gerber, R. B.; Ratner, M. A. Vibrational States of Very Floppy Clusters: Approximate Separability and the Choice of Good Curvilinear Coordinates for $XeHe_2$, I_2He. *J. Chem. Phys.* **1989**, *91*, 1813−1823.

(16) Gibson, L. L.; Roth, R. M.; Ratner, M. A.; Gerber, R. B. The Semiclassical Self-Consistent Field Method for Polyatomic Vibrations: Use of Hyperspherical Coordinates for H_2O and CO_2. *J. Chem. Phys.* **1986**, *85*, 3425−3431.

(17) Mladenović, M. Rovibrational Hamiltonians for General Polyatomic Molecules in Spherical Polar Parametrization. I. Orthogonal Representations. *J. Chem. Phys.* **2000**, *112*, 1070−1081.

(18) Ndong, M.; Joubert-Doriol, L.; Meyer, H.-D.; Nauts, A.; Gatti, F.; Lauvergnat, D. Automatic Computer Procedure for Generating Exact and Analytical Kinetic Energy Operators Based on the Polyspherical Approach. *J. Chem. Phys.* **2012**, *136*, No. 034107.

(19) Ndong, M.; Nauts, A.; Joubert-Doriol, L.; Meyer, H.-D.; Gatti, F.; Lauvergnat, D. Automatic Computer Procedure for Generating Exact and Analytical Kinetic Energy Operators Based on the Polyspherical Approach: General Formulation and Removal of Singularities. *J. Chem. Phys.* **2013**, *139*, No. 204107.

(20) Sadri, K.; Lauvergnat, D.; Gatti, F.; Meyer, H.-D. Numeric Kinetic Energy Operators for Molecules in Polyspherical Coordinates. *J. Chem. Phys.* **2012**, *136*, No. 234112.

(21) Sadri, K.; Lauvergnat, D.; Gatti, F.; Meyer, H.-D. Rovibrational Spectroscopy Using a Kinetic Energy Operator in Eckart Frame and the Multi-Configuration Time-Dependent Hartree (MCTDH) Approach. *J. Chem. Phys.* **2014**, *141*, No. 114101.

(22) Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. Systematic Ab Initio Gradient Calculation of Molecular Geometries, Force Constants, and Dipole Moment Derivatives. *J. Am. Chem. Soc.* **1979**, *101*, 2550−2560.

(23) Fogarasi, G.; Zhou, X.; Taylor, P. W.; Pulay, P. The Calculation of Ab Initio Molecular Geometries: Efficient Optimization by Natural Internal Coordinates and Empirical Correction by Offset Forces. *J. Am. Chem. Soc.* **1992**, *114*, 8191−8201.

(24) Yagi, K.; Keçeli, M.; Hirata, S. Optimized Coordinates for Anharmonic Vibrational Structure Theories. *J. Chem. Phys.* **2012**, *137*, No. 204118.

(25) Zimmerman, P. M.; Smereka, P. Optimizing Vibrational Coordinates To Modulate Intermode Coupling. *J. Chem. Theory Comput.* **2016**, *12*, 1883−1891.

(26) Oenen, K.; Dinu, D. F.; Liedl, K. R. Determining Internal Coordinate Sets for Optimal Representation of Molecular Vibration. *J. Chem. Phys.* **2024**, *160*, No. 014104.

(27) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. Assessment of Gaussian-2 and Density Functional Theories for the Computation of Enthalpies of Formation. *J. Chem. Phys.* **1997**, *106*, 1063−1079.

(28) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A Fifth-Order Perturbation Comparison of Electron Correlation Theories. *Chem. Phys. Lett.* **1989**, *157*, 479−483.

(29) Bartlett, R. J.; Watts, J. D.; Kucharski, S. A.; Noga, J. Non-Iterative Fifth-Order Triple and Quadruple Excitation Energy Corrections in Correlated Methods. *Chem. Phys. Lett.* **1990**, *165*, 513−522.

(30) Stanton, J. F. Why CCSD(T) Works: A Different Perspective. *Chem. Phys. Lett.* **1997**, *281*, 130−134.

(31) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098−3100.

(32) Lee, T. J.; Taylor, P. R. A Diagnostic for Determining the Quality of Single-Reference Electron Correlation Methods. *Int. J. Quantum Chem.* **2009**, *36*, 199−207.

(33) Nielsen, I. M.; Janssen, C. L. Double-Substitution-Based Diagnostics for Coupled-Cluster and Møller-Plesset Perturbation Theory. *Chem. Phys. Lett.* **1999**, *310*, 568−576.

(34) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257−2261.

(35) Hariharan, P. C.; Pople, J. A. The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies. *Theor. Chem. Acta* **1973**, *28*, 213−222.

(36) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *Ab Initio* Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem. A* **1994**, *98*, 11623−11627.

(37) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618.

(38) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. Use of Approximate Integrals in *Ab Initio* Theory. An Application in MP2 Energy Calculations. *Chem. Phys. Lett.* **1993**, *208*, 359−363.

(39) Distasio, R. A., Jr.; Steele, R. P.; Rhee, Y. M.; Shao, Y.; Head-Gordon, M. An Improved Algorithm for Analytical Gradient Evaluation in Resolution-of-the-Identity Second-Order Møller-Plesset Perturbation Theory: Application to Alanine Tetrapeptide Conformational Analysis. *J. Comput. Chem.* **2007**, *28*, 839−856.

(40) Dunning, T. H., Jr. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.

(41) Dunning, T. H., Jr.; Peterson, K. A.; Wilson, A. K. Gaussian Basis Sets for Use in Correlated Molecular Calculations. X. The Atoms Aluminum through Argon Revisited. *J. Chem. Phys.* **2001**, *114*, 9244−9253.

(42) Peterson, K. A.; Dunning, T. H., Jr. Accurate Correlation Consistent Basis Sets for Molecular Core-Valence Correlation Effects:

https://doi.org/10.1021/acs.jctc.4c01240
J. Chem. Theory Comput. 2024, 20, 10886−10898

The Second Row Atoms Al-Ar, and the First Row Atoms B-Ne Revisited. *J. Chem. Phys.* **2002**, *117*, 10548−10560.
(43) Wilson, E. B., Jr.; Decius, J. C.; Cross, P. C. *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*; Dover Publications; Revised ed., 1980.
(44) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. et al. MOLPRO, Version 2010.1, A Package of *Ab Initio* Programs. 2010.
(45) Smith, D. G. A.; Burns, L. A.; Simmonett, A. C.; Parrish, R. M.; Schieber, M. C.; Galvelis, R.; Kraus, P.; Kruse, H.; Di Remigio, R.; Alenaizan, A.; et al. PSI4 1.4: Open-Source Software for High-Throughput Quantum Chemistry. *J. Chem. Phys.* **2020**, *152*, No. 184108.
(46) Allen, W. D.; Császár, A. G. On the *Ab Initio* Determination of Higher-Order Force Constants at Nonstationary Reference Geometries. *J. Chem. Phys.* **1993**, *98*, 2983−3015.
(47) Pulay, P. *Ab Initio Calculation of Force Constants and Equilibrium Geometries in Polyatomic Molecules. 1. Theory*; Springer, Berlin, Heidelberg, 1969.
(48) Stanton, J. F.; Lipscomb, W. N.; Magers, D. H.; Bartlett, R. J. Highly Correlated Single-Reference Studies of the O_3 Potential Surface. I. Effects of High Order Excitations on the Equilibrium Structure and Harmonic Force Field of Ozone. *J. Chem. Phys.* **1989**, *90*, 1077−1082.
(49) Allen, W. D.; Császár, A. G.; Horner, D. A. The Puckering Inversion Barrier and Vibrational Spectrum of Cyclopentene. A Scaled Quantum Mechanical Force Field Algorithm. *J. Am. Chem. Soc.* **1992**, *114*, 6834−6849.
(50) Pulay, P.; Török, F. Über die Parameterdarstellung der KraftkonstantenMatrix. I. *Acta Chim. Hung.* **1965**, *44*, 287.
(51) Keresztury, G.; Jalsovszky, G. An Alternative Calculation of the Vibrational Potential Energy Distribution. *J. Mol. Struct.* **1971**, *10*, 304−305.

https://doi.org/10.1021/acs.jctc.4c01240
J. Chem. Theory Comput. 2024, 20, 10886−10898