



A revised formulation of the generalized subsystem vibrational analysis (GSVA)

Yunwen Tao¹ · Wenli Zou² · Sadisha Nanayakkara¹ · Marek Freindorf¹ · Elfi Kraka¹

Received: 9 June 2020 / Accepted: 8 February 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

In this work, a simplified formulation of our recently developed generalized subsystem vibrational analysis (GSVA) for obtaining intrinsic fragmental vibrations (J Chem Theory Comput 14:2558, 2018) is presented. In contrast to the earlier implementation, which requires the explicit definition of a non-redundant set of internal coordinate parameters to be constructed for the subsystem, the new implementation circumvents this process by employing massless Eckart conditions to the subsystem fragment paired with a Gram–Schmidt orthogonalization to span the same internal vibration space indirectly. This revised version of GSVA (rev-GSVA) can be applied to equilibrium structure as well as transition state structure, and it has been incorporated into the open-source package UniMoVib (<https://github.com/zorkzou/UniMoVib>).

Keywords Vibrational spectroscopy · Normal mode analysis · Subsystem · Harmonic approximation · GSVA · UniMoVib

Molecular vibrations described by normal modes are in general delocalized over the molecular structure of interest [1]. This delocalization hampers the direct comparison between the vibrations of a molecule in different environments (e.g., gas phase and solution) or when this molecule binds to a host compound. Over the past two decades, enormous effort has been put into finding “localized” normal vibrational modes (which can also be called “localized modals” in analogue to “localized orbitals” [2]), belonging to a molecule within a noncovalently bonded complex or a fragment of one molecule, including partial Hessian diagonalization [3], partial Hessian vibrational analysis (PHVA) [4, 5], mobile block Hessian (MBH) [6–10], vibrational subsystem analysis (VSA) [11–13], local Hessian transformation [14], just to name a few [15]. However, all these approaches share a common deficiency—the unphysical partitioning of the full Hessian matrix, which causes information loss about the interaction between the subsystem and its environment.

Recently, we proposed the generalized subsystem vibrational analysis (GSVA) as a new solution to obtain intrinsic fragmental vibrations [16, 17]. The key feature of GSVA compared to its predecessors lies in avoiding the partitioning of the full Hessian matrix. Instead, GSVA extracts for the subsystem a unique *effective Hessian matrix* $\mathbf{F}_{\text{sub}}^x$

$$\mathbf{F}_{\text{sub}}^x = \mathbf{B}_{\text{sub}}^{\dagger} (\mathbf{B}' (\mathbf{F}^x)^+ \mathbf{B}'^{\dagger})^{-1} \mathbf{B}'_{\text{sub}} \quad (1)$$

where \mathbf{F}^x is the full Hessian matrix expressed in Cartesian coordinates of dimension $(3N \times 3N)$ for the whole molecular system being composed of N atoms including n atoms of the target subsystem and $N - n$ atoms of the environment. The Wilson \mathbf{B} -matrices [1] \mathbf{B}' and \mathbf{B}'_{sub} define a non-redundant set of $(3n - k_{\text{sub}})$ internal coordinates for the subsystem fragment in rows with full $3N$ columns and truncated $3n$ columns (excluding the environment atoms), respectively. k_{sub} is the total number of rotations and translations for the subsystem being 5 or 6 depending on whether the subsystem geometry is linear or nonlinear. $(\mathbf{F}^x)^+$ is the Moore–Penrose inverse [18] of \mathbf{F}^x , which is singular. $\mathbf{F}_{\text{sub}}^x$ on the left-hand side is a symmetric matrix of dimension $(3n \times 3n)$, and it has exactly k_{sub} zero eigenvalues. The \dagger superscript denotes matrix transpose.

With the effective Hessian matrix $\mathbf{F}_{\text{sub}}^x$ expressed in Cartesian coordinates, the conventional normal mode analysis (NMA) machinery [19, 20] which is widely implemented in

✉ Elfi Kraka
ekraka@gmail.com

¹ Department of Chemistry, Southern Methodist University, 3215 Daniel Ave, Dallas, TX 75275-0314, USA

² Institute of Modern Physics, Northwest University, and Shaanxi Key Laboratory for Theoretical Physics Frontiers, Xi'an 710127, Shaanxi, People's Republic of China

most quantum chemical packages can be employed to calculate for the subsystem a new type of localized normal modes, which we coined *intrinsic fragmental vibrations*. The reason why these normal vibrations are called “intrinsic” is due to the fact that the effective Hessian matrix $\mathbf{F}_{\text{sub}}^x$ retains the curvature of the potential energy surface (PES) in the direction defined by any internal coordinate within the subsystem [16, 17]. In other words, the subsystem fragment “feels” exactly the same curvature of the PES as the whole system being described with the full Hessian matrix \mathbf{F}^x . This property of $\mathbf{F}_{\text{sub}}^x$ endows our GSVA method with a solid physical basis [16].

As discussed in our earlier work on the original GSVA implementation [16], GSVA requires a complete and non-redundant set of internal coordinate parameters, and its Wilson- \mathbf{B} matrices (see Eq. 1) to span the internal vibration space of the subsystem. However, the construction of the non-redundant parameter set is nontrivial and it needs either judicious selection of parameters manually with expert knowledge, or a dedicated algorithm which automatically selects the non-redundant parameter set from a series of redundant set of parameters in a trial-and-error manner.

In this work, we propose an alternative formulation of GSVA, which can save the effort of constructing the non-redundant parameter set for the subsystem. The new implementation replaces the Wilson- \mathbf{B} matrices in Eq. 1 with a different matrix, which also spans the internal vibrational space of the subsystem via the following procedure.

First, we apply to the subsystem fragment with its Cartesian coordinates collected in a $3n \times 1$ column vector \mathbf{R}_{cart} the massless (assuming all atomic masses are identical) Eckart conditions [21, 22] to generate a set of five or six translational and rotational vectors which are orthonormal to each other; see Eq. 2:

$$\mathbf{R}_{tr.+ro.} = \{\mathbf{r}_1, \dots, \mathbf{r}_i, \dots, \mathbf{r}_k\} \quad (2)$$

where k equals 5 or 6 depending on whether the subsystem is linear or not and \mathbf{r}_i is a column vector of length $3n$.

Next, a Gram–Schmidt orthonormalization is conducted on $\mathbf{R}_{tr.+ro.}$ to generate $n_{\text{vib}} = 3n - k$ remaining vectors collected in \mathbf{V}

$$\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_j, \dots, \mathbf{v}_{n_{\text{vib}}}\} \quad (3)$$

where \mathbf{v}_j is a column vector of length $3n$. It has to be noted that matrix \mathbf{V} is equivalent to matrix $\mathbf{B}_{\text{sub}}^{T\dagger}$ in spanning the internal coordinate/vibration space of the subsystem. In order to obtain the equivalent matrix of $\mathbf{B}^{T\dagger}$, we pad each column vector \mathbf{v}_j with $3(N - n)$ zeros associated with environmental atoms, resulting in the matrix \mathbf{V}_{full} with the dimension $(3N \times n_{\text{vib}})$.

In this way, the effective Hessian matrix $\mathbf{F}_{\text{sub}}^x$ for the subsystem can be written as

$$\mathbf{F}_{\text{sub}}^x = \mathbf{V}(\mathbf{V}_{\text{full}}^\dagger(\mathbf{F}^x) + \mathbf{V}_{\text{full}})^{-1}\mathbf{V}^\dagger. \quad (4)$$

Then, the conventional NMA machinery is applied to obtain the intrinsic fragmental vibrations for the subsystem as in our earlier formulation [16]. The new formulation has two major advantages for practical implementation. (1) It avoids the complicated process of finding the complete and non-redundant internal coordinate parameter set for the subsystem; (2) the code for finding translation/rotation vectors from the Eckart conditions and conducting Gram–Schmidt orthonormalization in a modern quantum chemical package can be reused, which facilitates implementing GSVA. In order to distinguish from the original implementation of GSVA, this revised implementation was named rev-GSVA.

As a showcase example, we have employed rev-GSVA to calculate the intrinsic fragmental vibrations of the methane (CH_4) molecule in (1) methane-intercalated $\text{B}_{36}\text{N}_{36}$ complex (Fig. 1a), (2) methane-intercalated C_{60} structure (Fig. 1b) [23] and (3) gas phase as reference. Unlike the methane-intercalated C_{60} complex, the methane-intercalated $\text{B}_{36}\text{N}_{36}$ system [24] has not been synthesized experimentally so far, and it is interesting to compare the intrinsic fragmental vibrations of the methane molecule in $\text{B}_{36}\text{N}_{36}$ and C_{60} in order to explore the different encapsulation effect. These three molecular systems were modeled at the M06-2X/6-31G(d,p) level [25–27] with Grimme’s D3(0) dispersion correction [28] using the Gaussian 16 package [29].

The results in Table 1 show that the methane molecules encapsulated inside the two cages retain T_d symmetry, as the reference methane molecule in gas phase. The non-degenerate A_1 vibration describes the symmetric stretching of four C–H bonds. The doubly degenerate E modes specify the relative turnstile twisting motions of two H–C–H fragments. The triply degenerate T_2 modes (1–3) with lower frequencies denote the bending of methane, while the other triply degenerate T_2 modes (1’–3’) with higher frequencies are antisymmetric stretching motions of four C–H bonds. We found the

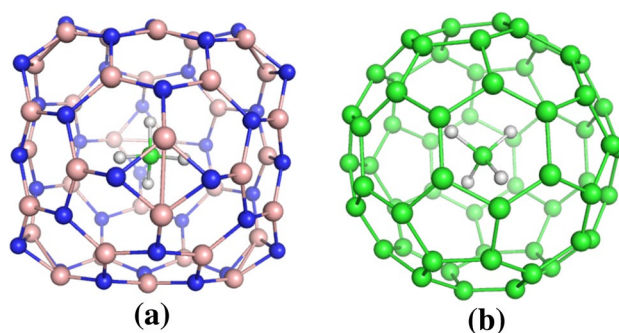


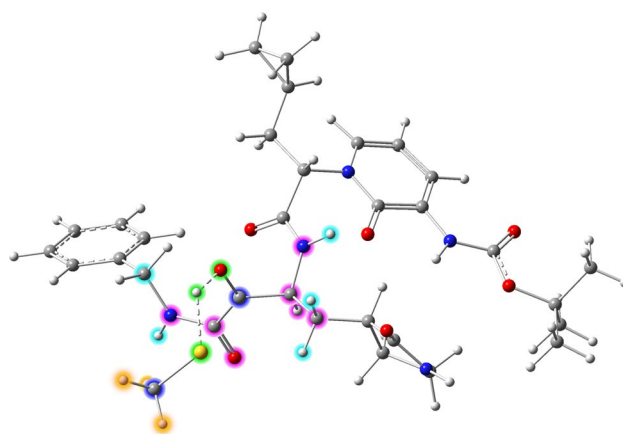
Fig. 1 Structure of **a** methane molecule encapsulated in $\text{B}_{36}\text{N}_{36}$ cage with T_d symmetry and **b** methane encapsulated in fullerene (C_{60}) with T symmetry

Table 1 Normal mode frequencies (in cm^{-1}) of CH_4 determined with rev-GSVA in different environments

Symmetry(No.)	$\text{CH}_4@B_{36}N_{36}$	$\text{CH}_4@C_{60}$	Gas phase
A_1	3051	3150	3080
$E(1)$	1560	1583	1584
$E(2)$	1560	1583	1584
$T_2(1)$	1326	1328	1356
$T_2(2)$	1326	1328	1356
$T_2(3)$	1326	1328	1356
$T_2(1')$	3167	3250	3205
$T_2(2')$	3167	3250	3205
$T_2(3')$	3167	3250	3205

largest deviation in the vibrational frequencies relative to methane in gas phase for the A_1 mode and the $T_2(1'-3')$ modes which are concerned with the C–H bond stretching. For methane molecule contained in the $B_{39}N_{39}$ cage, above two vibrations are redshifted by 29 and 38 cm^{-1} respectively compared to gas phase. However, these two vibrations of methane in fullerene are blueshifted by 70 and 45 cm^{-1} , respectively. This means the fullerene cage could strengthen the C–H bonds of methane, while the $B_{36}N_{36}$ weakens the C–H bonds of the contained methane molecule. One might argue that comparing the normal mode frequencies of the whole system (the way spectroscopists usually adopt) could lead to similar conclusion because the methane molecule is well separated from the cage structure; however, one needs to note that only the (rev-)GSVA method could provide the localized normal modes and frequencies which can be legitimately comparable across different systems containing the same target subsystem.

In addition to local minima on the PES, we have also tested a first-order saddle point (i.e., transition state, TS) structure. As it has been proven in our earlier work [16] that GSVA retains the curvature of the PES, it is of interest to explore whether (rev-)GSVA can retain the imaginary vibrational mode specifying the bond breaking/forming in the subsystem. In this pilot study, we investigated the TS of the chemical reaction involving a potential α -ketoamide inhibitor of the SARS-CoV-2 main protease (Mpro), which is assumed to inhibit the activity of SARS-CoV-2 virus by blocking viral replication [30]. According to a recent X-ray structure of the α -ketoamide SARS-CoV-2 main protease (Mpro) complex [30], ketoamide and enzyme are linked via a cysteine side chain of the enzyme. According to the suggested catalytic mechanism, the chemical reaction starts with a nucleophilic attack of the sulfur atom onto a C=O carbon atom of α -ketoamide moiety, which is followed by proton transfer from the –SH group to a nearby oxygen atom of the inhibitor, as shown in Fig. 2. Work is in progress to model the reaction in the enzyme.

**Fig. 2** Transition-state structure of proton transfer from methanethiol to an α -ketoamide inhibitor [30]. The methanethiol group is a simplified model of cysteine in SARS-CoV-2 main protease. The minimal 3-atom subsystem is highlighted in green color. The 5-atom subsystem is highlighted with green and blue. The 8-atom subsystem includes the 5-atom subsystem and atoms highlighted with orange. The 15-atom subsystem includes the 8-atom subsystem and atoms highlighted in purple. The 20-atom subsystem includes the 15-atom subsystem and atoms highlighted with cyan. This system was modeled at B3LYP/6-31G(d,p) level in Gaussian 16

We started with a minimal subsystem of 3 atoms containing the proton and its donor/acceptor atoms, and rev-GSVA was applied to calculate the corresponding intrinsic fragmental vibrations. Surprisingly, no imaginary frequency exists for this small 3-atom subsystem. This is probably due to the fact that the reaction center should also include the carbon atom to draw a more complete picture about bond forming and breaking. However, one imaginary frequency starts to emerge when more surrounding atoms are included into the subsystem (see Table 2) and the imaginary frequency value quickly converges to that of the full system when the subsystem contains 20 atoms. This result indicates that if the normal vibration is localized in a particular part of the molecule (e.g., bond breaking/forming or C=O bond stretching), (rev-)GSVA

Table 2 Imaginary frequency (in cm^{-1}) within the subsystem of different sizes for the transition state of proton transfer reaction

Subsystem size (n)	Imaginary frequency (ω_i)
3	n/a
5	-789
8	-778
15	-729
20	-727
Full system	-725

is expected to reproduce this vibration using a subsystem containing this segment and a few surrounding atoms. This valuable feature of (rev-)GSVA can lead to important first insights into the role of surrounding atoms for the reaction mechanism by analyzing the TS, before starting a more complex reaction path following procedure.

We have implemented the new formulation of GSVA (rev-GSVA) introduced in this work into the open-source package UniMoVib [20] for interested readers to use in their own research.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00214-021-02727-y>.

Acknowledgements This work was financially supported by the National Science Foundation Grants CHE 1464906. W. Z. was supported by the National Natural Science Foundation of China (Grant No. 21673175) and the Double First-Class University Construction Project of Northwest University. We thank SMU for providing super-computing resources.

References

- Wilson EB, Decius JC, Cross PC (2012) *Molecular vibrations: the theory of infrared and raman vibrational spectra*. Dover Publications, Mineola
- Thompson TC, Truhlar DG (1980) SCF CI calculations for vibrational eigenvalues and wavefunctions of systems exhibiting fermi resonance. *Chem Phys Lett* 75(1):87–90
- Head JD (1997) Computation of vibrational frequencies for adsorbates on surfaces. *Int J Quantum Chem* 65(5):827–838
- Li H, Jensen JH (2002) Partial Hessian vibrational analysis: the localization of the molecular vibrational energy and entropy. *Theor Chem Acc* 107(4):211–219
- Besley NA, Bryan JA (2008) Partial Hessian vibrational analysis of organic molecules adsorbed on Si(100). *J Phys Chem C* 112(11):4308–4314
- Ghysels A, Van Neck D, Van Speybroeck V, Verstraelen T, Waroquier M (2007a) Vibrational modes in partially optimized molecular systems. *J Chem Phys* 126(22):224102
- Ghysels A, Van Neck D, Waroquier M (2007b) Cartesian formulation of the mobile block Hessian approach to vibrational analysis in partially optimized systems. *J Chem Phys* 127(16):164108
- Ghysels A, Van Neck D, Brooks BR, Van Speybroeck V, Waroquier M (2009a) Normal modes for large molecules with arbitrary link constraints in the mobile block Hessian approach. *J Chem Phys* 130(8):084107
- Ghysels A, Van Speybroeck V, Pauwels E, Van Neck D, Brooks BR, Waroquier M (2009b) Mobile block Hessian approach with adjoined blocks: an efficient approach for the calculation of frequencies in macromolecules. *J Chem Theory Comput* 5(5):1203–1215
- Ghysels A, Van Speybroeck V, Verstraelen T, Van Neck D, Waroquier M (2008) Calculating reaction rates with partial Hessians: validation of the mobile block Hessian approach. *J Chem Theory Comput* 4(4):614–625
- Woodcock HL, Zheng W, Ghysels A, Shao Y, Kong J, Brooks BR (2008) Vibrational subsystem analysis: a method for probing free energies and correlations in the harmonic limit. *J Chem Phys* 129(21):12B603
- Ghysels A, Van Speybroeck V, Pauwels E, Catak S, Brooks BR, Van Neck D, Waroquier M (2010a) Comparative study of various normal mode analysis techniques based on partial Hessians. *J Comput Chem* 31(5):994–1007
- Ghysels A, Van Speybroeck V, Pauwels E, Catak S, Brooks BR, Van Neck D, Waroquier M (2010b) Comparative study of various normal mode analysis techniques based on partial Hessians. *J Comput Chem* 31(5):994–1007
- Huix-Rotllant M, Ferré N (2016) An effective procedure for analyzing molecular vibrations in terms of local fragment modes. *J Chem Theory Comput* 12(10):4768–4777
- Jacob CR, Reiher M (2009) Localizing normal modes in large molecules. *J Chem Phys* 130(8):084106
- Tao Y, Tian C, Verma N, Zou W, Wang C, Cremer D, Kraka E (2018) Recovering intrinsic fragmental vibrations using the generalized subsystem vibrational analysis. *J Chem Theory Comput* 14(5):2558–2569
- Kraka E, Zou W, Tao Y (2020) Decoding chemical information from vibrational spectroscopy data: local vibrational mode theory. *WIREs Comput Mol Sci* 10:e1480
- Penrose R (1955) A generalized inverse for matrices. *Proc Camb Philos Soc* 51(3):406–413
- Ochterski J (1999) White paper: vibrational analysis in Gaussian. Gaussian Inc, Wallingford
- Zou W (2020) UniMoVib: a unified interface for molecular harmonic vibrational frequency calculations. <https://github.com/zorkzou/UniMoVib>. Accessed 1 May 2020
- Eckart C (1935) Some studies concerning rotating axes and polyatomic molecules. *Phys Rev* 47(7):552–558
- Sayvetz A (1939) The kinetic energy of polyatomic molecules. *J Chem Phys* 7(6):383–389
- Morosin B, Assink R, Dunn R, Massis T, Schirber J, Kwei G (1997) Methane-intercalated C_{60} : preparation, orientational ordering, and structure. *Phys Rev B* 56(21):13611–13614
- Zope RR, Baruah T, Pederson MR, Dunlap BI (2005) Theoretical infrared, Raman, and optical spectra of the $B_{36}N_{36}$ cage. *Phys Rev A* 71(2):025201
- Zhao Y, Truhlar DG (2008) The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, non-covalent interactions, excited states, and transition elements: two new functionals and systematic testing of four m06-class functionals and 12 other functionals. *Theor Chem Acc* 120(1):215–241
- Dill JD, Pople JA (1975) Self-consistent molecular orbital methods. XV. Extended Gaussian-type basis sets for lithium, beryllium, and boron. *J Chem Phys* 62(7):2921–2923
- Hehre WJ, Ditchfield R, Pople JA (1972) 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* 56(5):2257–2261
- Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J Chem Phys* 132(15):154104
- Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Petersson GA, Nakatsuji H, Li X, Caricato M, Marenich AV, Bloino J, Janesko BG, Gomperts R, Mennucci B, Hratchian HP, Ortiz JV, Izmaylov AF, Sonnenberg JL, Williams-Young D, Ding F, Lipparini F, Egidi F, Goings J, Peng B, Petrone A, Henderson T, Ranasinghe D, Zakrzewski VG, Gao J, Rega N, Zheng G, Liang W, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Throssell K, Montgomery JA Jr, Peralta JE, Ogliaro F, Bearpark MJ, Heyd JJ, Brothers EN, Kudin KN, Staroverov VN, Keith TA, Kobayashi R, Normand J, Raghavachari K, Rendell AP, Burant JC, Iyengar SS, Tomasi J, Cossi M, Millam JM, Klene M, Adamo C, Cammi R, Ochterski

- JW, Martin RL, Morokuma K, Farkas O, Foresman JB, Fox DJ (2016) Gaussian 16 revision B.01. Gaussian Inc., Wallingford
30. Zhang L, Lin D, Sun X, Curth U, Drosten C, Sauerhering L, Becker S, Rox K, Hilgenfeld R (2020) Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science* 368(6489):409–412

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.