



Indices to evaluate the reliability of coarse-grained representations of mixed inter/intramolecular vibrations

Makoto Isogai¹ · Hirohiko Houjou¹

Received: 16 February 2018 / Accepted: 20 July 2018 / Published online: 2 August 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

We propose some methods for quantifying the reliability of coarse-grained representations of displacement vectors of normal mode vibrations. In the framework of our basic theory, the original displacement vectors are projected onto a lower-dimensional (i.e., a coarse-grained) space. Four types of functions denoted fidelity indices were introduced as measures of the similarity of the original to the restored displacement vectors. These indices were applied to several hydrogen-bonded homodimers, and the behavior of each index was examined. We found that a coarse-grained representation with high reliability resulted in the accurate restoration of properties such as eigenfrequency, modal mass, and modal stiffness.

Keywords Molecular vibration · Coarse graining · Molecular assembly · Normal mode analysis · GF matrix method

Introduction

Among the vibrational modes of molecular assemblies such as dimers, clusters, and crystals, low-frequency (LF) vibrational modes are of considerable importance in terahertz absorption [1–5], Raman scattering [6–9], and resonant two-photon ionization [10–12]. These methods provide a great deal of information about intermolecular forces concealed in translational and librational motions of molecules. If the constituent molecules are sufficiently rigid, the LF vibrations can be represented as the relative displacements of rigid bodies and are virtually separate from high-frequency intramolecular vibration modes [13, 14]. In molecules containing flexible moieties, some intramolecular vibrations may couple with intermolecular motions [15]. The contributions of inter/intramolecular components to LF vibrations are of considerable importance in crystal structure prediction [16, 17]. Several methods of modal analysis that explicitly account for coupling effects among inter/intramolecular vibrations have been proposed [3, 5, 18–20]. However, strict atomistic models often face

problems relating to the computational cost of the associated calculations.

Previously, we proposed a universal formulation for coarse-graining LF vibrational modes that incorporated contributions from inter/intramolecular vibrational coupling [21–23]. In the framework of our formulation, an atomic displacement vector ($3N$ dimensions for a molecule containing N atoms) is represented by a linear expansion in a lower-dimensional space (a coarse-grained space) spanned by a basis set corresponding to inter/intramolecular motions, namely three translations, three rotations, and some (a maximum of $3N - 6$) intramolecular vibrations per nonlinear molecule. The coefficient of linear expansion can be regarded as a component in an internal coordinate system that determines its own representation of the stiffness constants and inertial loads of the constituents. In principle, we can restore information about the original vibrations, namely the atomic displacement, modal stiffness, modal mass, and eigenfrequency. However, if the basis set of the coarse-graining space is inadequately selected, some essential details will be lost, so the results of that calculation may be erroneous [22]. In other words, the quality of the basis set can be evaluated by measuring the similarity of the restored data to the original data. In this study, we propose four such indices for quantifying the reliability of coarse-grained representations of inter/intramolecular vibrations. We apply these measures to several hydrogen-bonded dimers

✉ Hirohiko Houjou
houjou@iis.u-tokyo.ac.jp

¹ Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan

and compare their characteristics. We then consider the best index to use according to the type of information to be restored.

Theory

The normal mode motion of a molecular assembly composed of N atoms is represented as a displacement vector \mathbf{c}_i with $3N$ dimensions, where i identifies either the zero-frequency modes (three translations $i = 1, 2, 3$; three rotations $i = 4, 5, 6$) or “true” vibrations ($i = 7, 8, \dots, 3N$). A full-displacement matrix \mathbf{C} contains \mathbf{c}_i as the i th column vector. By definition, \mathbf{c}_i is normalized but not necessarily orthogonal to another column vector. Here, we select a non-square ($3N \times n$) matrix $\tilde{\mathbf{C}}$ that contains n columns (implicitly including zero-frequency modes and LF vibrational modes) from \mathbf{C} , where n is the number of LF-mode vectors to be coarse-grained and equals the dimension of the coarse-grained space.

$\tilde{\mathbf{C}}$ is converted to a mass-weighted displacement (MWD) matrix $\tilde{\mathbf{W}}$ by a diagonal matrix \mathbf{M} containing atomic masses. In Eq. 1, $\tilde{\mathbf{w}}_i$, the i th column vector of $\tilde{\mathbf{W}}$, is normalized using matrix \mathbf{L} , the square root of the modal mass matrix \mathbf{L}^2 .

$$\tilde{\mathbf{W}} = \mathbf{M}^{1/2} \tilde{\mathbf{C}} \mathbf{L}^{-1} \quad (1)$$

$$\mathbf{L}^2 = \tilde{\mathbf{C}}^T \mathbf{M} \tilde{\mathbf{C}} \quad (2)$$

To fulfill Eq. 3, a coarse-grained displacement matrix Ξ is defined that involves $\tilde{\mathbf{B}}$, a non-square ($3N \times n$) coarse-graining matrix detailed in previous papers [21, 22]. In the present work, some improvements are made in the definition of $\tilde{\mathbf{B}}$ (see the “Appendix”).

$$\Xi = \Gamma \tilde{\mathbf{B}}^T \mathbf{M} \tilde{\mathbf{C}} \quad (3)$$

Matrix Γ is the inverse of the inertial load matrix Γ^{-1} , which is named by analogy with the matrix \mathbf{G}^{-1} in the GF method [24].

$$\Gamma^{-1} = \tilde{\mathbf{B}}^T \mathbf{M} \tilde{\mathbf{B}} \quad (4)$$

Using the inertial load matrix Γ^{-1} in the given coarse-grained coordinate system, Ξ is converted to the coarse-grained MWD matrix \mathbf{U} . In Eq. 4, each column vector of \mathbf{U} is normalized using matrix Λ , the square root of the modal mass matrix Λ^2 . We then have the following representations for the given coordinate system (which are somewhat similar to those for the Cartesian coordinate system, Eqs. 1 and 2):

$$\mathbf{U} = \Gamma^{-1/2} \Xi \Lambda^{-1} \quad (5)$$

$$\Lambda^2 = \Xi^T \Gamma^{-1} \Xi. \quad (6)$$

The modal mass matrix Λ^2 is largely identical to \mathbf{L}^2 when the transformer $\tilde{\mathbf{B}}$ is chosen appropriately. The MWD matrix in the original coordinate system is then approximately restored to the full-dimensional mass-weighted displacements by the following equation:

$$\bar{\mathbf{W}} = \mathbf{M}^{1/2} \tilde{\mathbf{B}} \Gamma^{1/2} \mathbf{U}. \quad (7)$$

The restored MWD matrix $\bar{\mathbf{W}}$ contains $\bar{\mathbf{w}}_i$ as the i th column, which is normalized but not necessarily orthogonal to another column.

We define a correlation matrix \mathbf{R} as the product of $\tilde{\mathbf{W}}^T$ and $\bar{\mathbf{W}}$; therefore, its component R_{ij} is the inner product of $\tilde{\mathbf{w}}_i$ and $\bar{\mathbf{w}}_j$.

$$\mathbf{R} = \tilde{\mathbf{W}}^T \bar{\mathbf{W}} \quad (8)$$

By definition, \mathbf{R} will be a unit matrix if $\bar{\mathbf{W}}$ is identical to $\tilde{\mathbf{W}}$. To quantify the similarity between $\tilde{\mathbf{W}}$ and $\bar{\mathbf{W}}$, we define four types of fidelity indices (F_1 to F_4) as follows:

$$F_1 = \frac{1}{n} \text{tr}(\mathbf{R}) \quad (9)$$

$$F_2 = \left(\prod_{i=1}^n R_{ii} \right)^{\frac{1}{n}} \quad (10)$$

$$F_3 = \prod_{i=1}^n R_{ii} \quad (11)$$

$$F_4 = \det(\mathbf{R}) \quad (12)$$

According to the original formulation, the reverse transformation of Ω (an eigenfrequency matrix) using \mathbf{U} and Γ gives the stiffness matrix Φ . The element Φ_{ij} in the stiffness matrix is a force constant with respect to the given coarse-grained internal coordinate system:

$$\Phi = \Gamma^{1/2} \mathbf{U} \Omega^2 \mathbf{U}^T \Gamma^{1/2} \quad (13)$$

The matrix Φ also includes an error originating from the selection of the coarse-graining matrix $\tilde{\mathbf{B}}$. Therefore, the diagonalization of $\Gamma^{-1/2} \Phi \Gamma^{-1/2}$ does not need to yield eigenfrequencies identical to the original ones. Here, we define the reproduced eigenfrequency matrix as $\bar{\Omega}$, which fulfills the following eigenvalue equation:

$$\Gamma^{1/2} \Phi \Gamma^{1/2} \bar{\mathbf{U}} = \bar{\Omega}^2 \bar{\mathbf{U}} \quad (14)$$

Computational details

We examined six hydrogen-bonded homodimers of formic acid, acetic acid, trichloroacetic acid, formamide,

formamidine, and urea (Fig. 1). We chose three carboxylic acids from our previous study [22], and added three new compounds in order to check the applicability of our method to nitrogen-containing hydrogen-bonding systems. The geometries of the monomers and dimers were optimized at the Hartree–Fock level of theory with the 6-311G(d, p) basis set in order to maintain continuity with our previous studies [21, 22]. Although we are aware that this level of theory is not sufficient for supramolecular systems, the present method does not depend on the level of calculation. Using the optimized geometries, normal mode analysis was performed at the same level of calculation. These molecular orbital calculations were performed using the Gaussian 09 W [25] suite of programs. The displacement vectors and frequencies of the normal mode vibrations were obtained from the output files and applied in subsequent calculations performed by a coarse-graining program developed in-house in order to determine stiffness constants and the four fidelity indices. The displacement vectors of the molecules were visualized using Jmol ver.14.4.0 [26].

Results and discussion

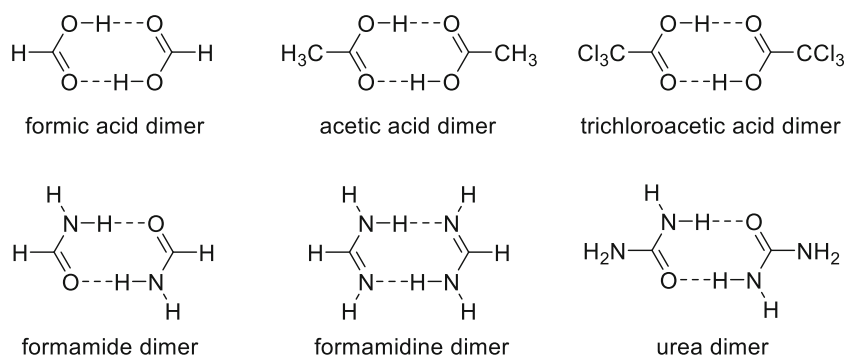
According to the definitions given in the “Theory” section, the indices F_1 to F_4 will be 1 if $\tilde{\mathbf{W}}$ and $\tilde{\mathbf{W}}$ are identical and will be <1 if these matrices are different from each other. F_1 and F_2 are the arithmetic and geometric means of R_{ii} , respectively, so F_1 is always larger than F_2 . Therefore, F_2 is a more severe criterion than F_1 . As n increases, F_2 underestimates the local mismatch between $\tilde{\mathbf{W}}$ and $\tilde{\mathbf{W}}$, while F_3 , the n th power of F_2 , is the severest criterion. In addition to the similarity between $\tilde{\mathbf{w}}_i$ and $\tilde{\mathbf{w}}_i$, F_4 takes into account the orthogonality between $\tilde{\mathbf{w}}_i$ and $\tilde{\mathbf{w}}_j$ ($i \neq j$). As a result, F_4 is considered to be the most severe criterion among the four indices.

Figure 2 compares the four fidelity indices F_1 to F_4 as functions of the coarse-graining dimension n . The component Φ_{T_x, T_x} , corresponding to translation along the x -axis, is also included in each plot as a representative of the set of intermolecular stiffness constants. Except for a few data points, F_1 and

F_2 varied within a narrow range between 0.95 and 1, suggesting that a discrepancy between the columns of $\tilde{\mathbf{W}}$ and $\tilde{\mathbf{W}}$ was averaged over the dimension n . Thus, F_1 and F_2 can be understood as indicators of the average similarity between the original vectors and the restored vectors. The variations in F_3 and F_4 show similar behavior within the range between 0.6 and 1, a much wider range of variation than seen for F_1 and F_2 . This behavior is reasonably explained from the character of F_3 and F_4 that are approximately a product of all the similarity between $\tilde{\mathbf{w}}_i$ and $\tilde{\mathbf{w}}_i$. Thus, F_3 and F_4 can be understood as indicators of the completeness of the restored vectors.

Among the four fidelity indices, we can see that F_3 and F_4 commonly show alternating behavior as a function of n : the values obtained at odd n are lower than those obtained at even n . This phenomenon is caused by the symmetries of the molecular systems. In the case of the formic acid dimer, for example, when going from the coarse-graining dimension $n = 12$ to $n = 13$, we need to add a new MWD vector to construct a $\tilde{\mathbf{W}}$ matrix composed of six zero-frequency motions (translation and rotation) and seven vibrations. The 13th column ($\tilde{\mathbf{w}}_{13}$) corresponds to the seventh vibrational motion, which is characterized as in-phase coupling of the intramolecular bending motions of O–C=O moieties (Fig. 3a). To coarse-grain this motion, the coarse-graining matrix $\tilde{\mathbf{B}}$ needs to contain displacement vectors corresponding to the symmetric combination of the bending motions of O–C=O moieties. In this study, however, each column vector in $\tilde{\mathbf{B}}$ is defined to represent the intramolecular motion of each molecule in the dimeric system (see the “Appendix”). This procedure means that the coarse-graining of $\tilde{\mathbf{w}}_{13}$ requires two vectors corresponding to the bending motions of the respective molecules. Specifically, we need at least 14 columns in $\tilde{\mathbf{B}}$ to adequately compress the dimension of $\tilde{\mathbf{w}}_{13}$. This augmented $\tilde{\mathbf{B}}$ matrix can also adequately compress the dimension of $\tilde{\mathbf{w}}_{14}$, the MWD vector of the eighth vibrational mode, which is characterized as antiphase coupling of the bending motions of O–C=O moieties (Fig. 3b). These considerations explain the rapid increases in the fidelity indices at $n = 14$. Similarly, the vibrational motions corresponding to two sequential columns are in-phase and antiphase couplings of intramolecular vibrations, which

Fig. 1 Structures of the hydrogen-bonded dimers studied



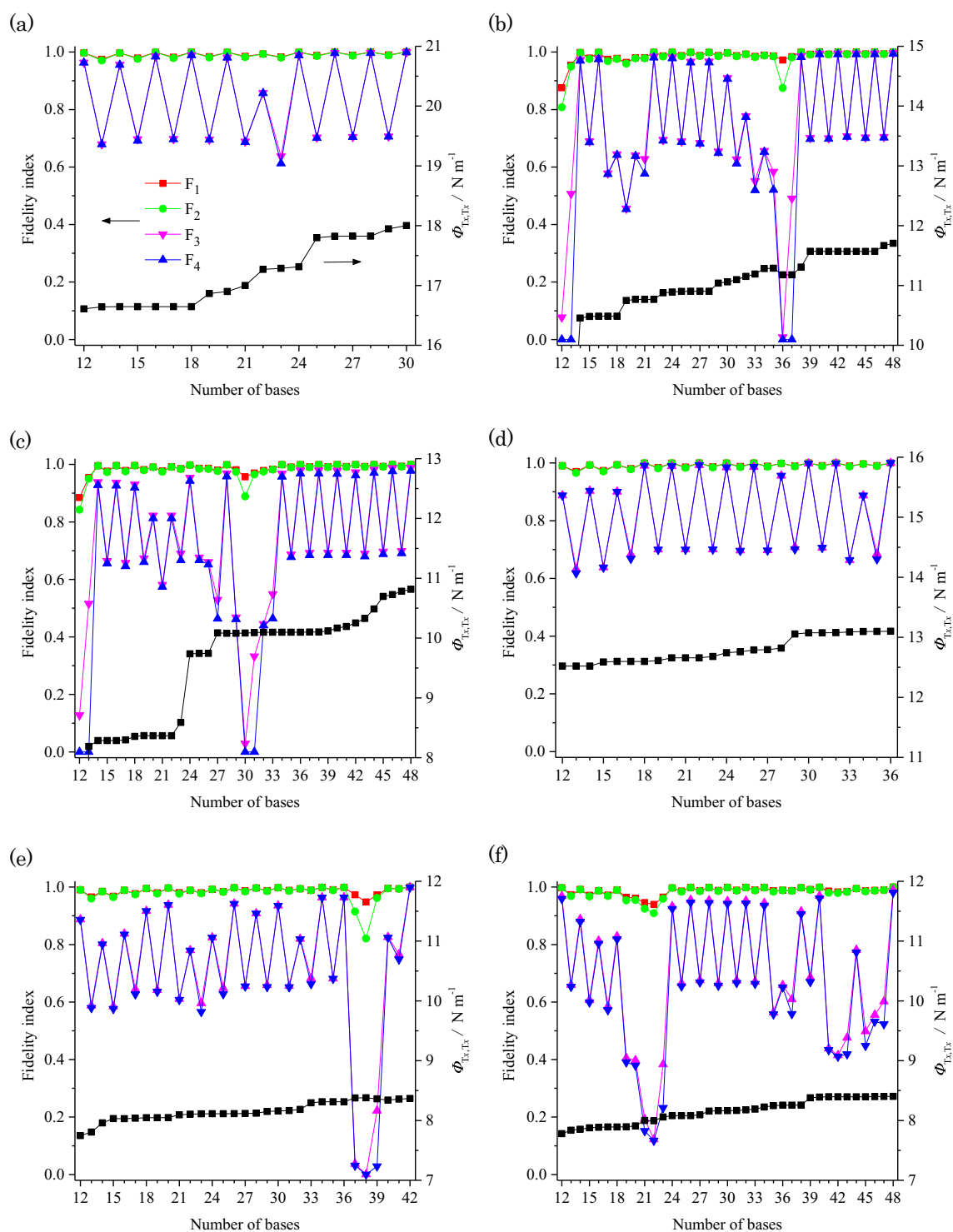


Fig. 2 a–f The four fidelity indices (left-hand y-axes) versus the coarse-graining dimension, as calculated for molecular dimers of **a** formic acid, **b** acetic acid, **c** trichloroacetic acid, **d** formamide, **e** formamidine, and **f**

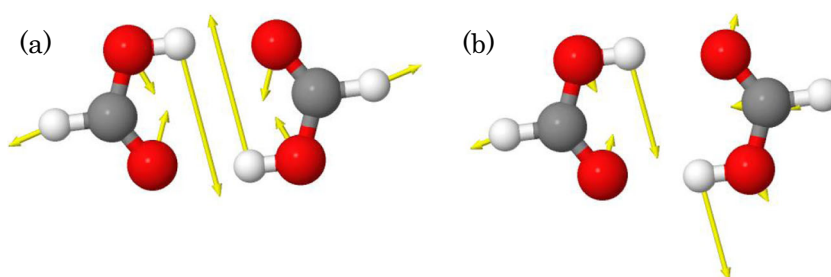
urea. Stiffness constants corresponding to x -axis translation (right-hand y-axes) are superimposed

accounts for the alternating behavior of the fidelity indices in Fig. 2a. The alternating behaviors of F_3 and F_4 for the other molecular dimers can be explained in a similar manner. In view of these results, we can propose a threshold for these fidelity indices: 0.8. Below this value, the intramolecular

vibrational motion(s) in $\tilde{\mathbf{B}}$ would be insufficient to express the vibrations of the molecular system. In contrast, this type of error does not seem to be notable for F_1 and F_2 .

In Fig. 2b, c, f, there are some disturbances of this alternating behavior, implying inconsistency in the vibrational

Fig. 3 a–b Atomic displacement vectors for a the seventh and b the eighth normal-mode vibrations of the formic acid dimer



motions between a monomer in a dimer and an isolated monomer. For the acetic acid dimer (Fig. 2b) and trichloroacetic acid dimer (Fig. 2c), the stiffness constants obtained with $n = 12$ and 13 are unnaturally low (0.0 and 7.9, respectively; data points not shown), implying that coarse-graining of the molecular vibration was conducted inadequately. This error is known to occur due to a lack of displacement vectors corresponding to intramolecular vibration modes (libration of the methyl group) in the $\tilde{\mathbf{B}}$ matrix under these computational conditions [22]. As compared to those seen at higher n , the fidelity indices at these dimension values are rather low: F_1 and F_2 are lower than 0.95 and F_3 and F_4 are only 0.0–0.5. These results suggest that the absence of an intramolecular vibrational motion that contributes to the vibration of the whole molecular

system results in severe decreases in the fidelity indices. In some cases, such severe decreases in fidelity indices occur simultaneously with an anomaly in the trend for the stiffness constant. However, for most coarse-graining dimension values, the stiffness constant shows a stepwise increase with increasing n . As described elsewhere, the stiffness constants calculated under low-dimension coarse-graining conditions are apparent stiffness constants (Φ_{app}), which are normally smaller than the actual ones (Φ_{inter}) due to modulation caused by coupling with the internal elasticity of the constituent molecules [22, 23]. Based on simple four-body models, we formulated a relationship between the height of the stepwise increase and the intramolecular stiffness constant (Φ_{intra}). According to this formulation, an erroneous evaluation of the step height will severely

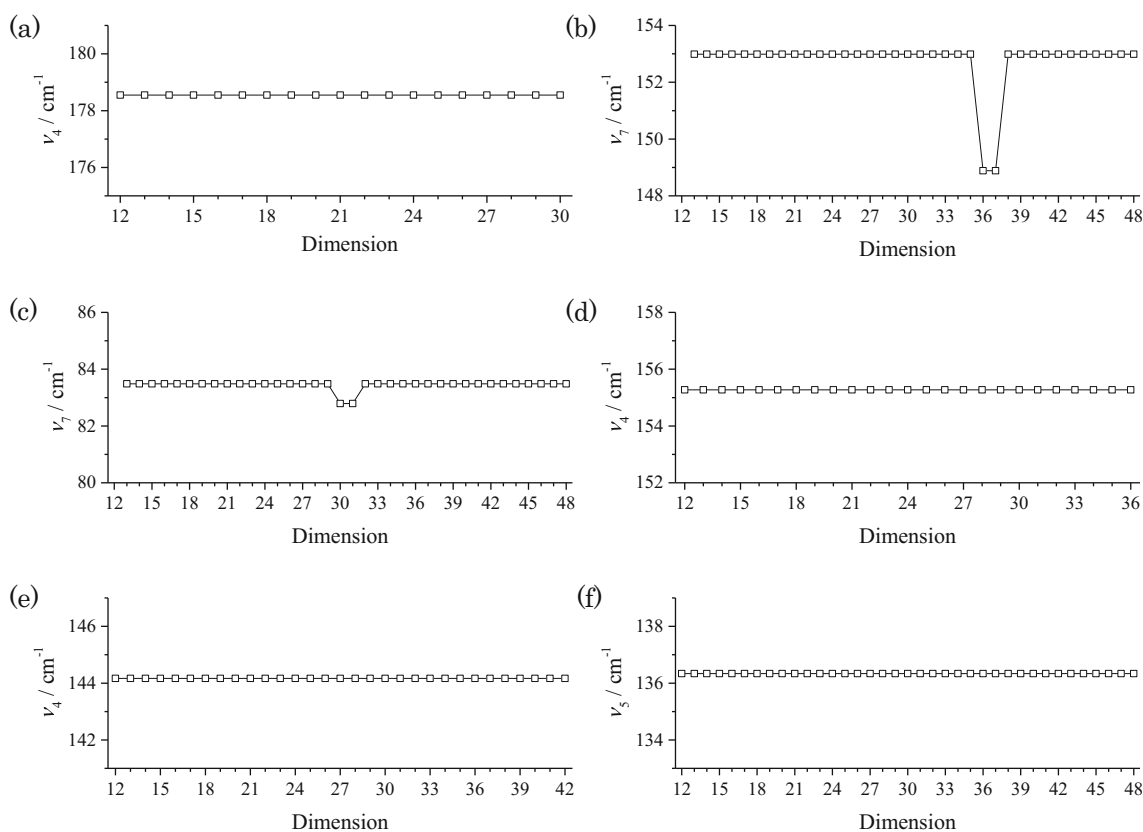


Fig. 4 a–f Restored frequency corresponding to stretching of the intermolecular bond as a function of the coarse-graining dimension for molecular dimers of a formic acid, b acetic acid, c trichloroacetic acid, d formamide, e formamidine, and f urea

compromise the accurate evaluation of the intramolecular stiffness constant. The present study provides a rational criterion for choosing the coarse-graining dimension with which the calculation can yield an adequate value of Φ_{app} . If we adopt a threshold for F_4 of 0.8, we can designate the minimum reliable dimension for the coarse-graining of formic acid, formamide, formamidine, and urea dimers as 12, and that for acetic acid and trichloroacetic acid dimers as 14. These values depend on the degrees of freedom of the LF molecular motion.

We examined the eigenfrequencies reproduced from the stiffness constants and inertial loads with respect to the given coarse-grained internal coordinate system (Eq. 14). Figure 4 shows the wavenumber ν_j (in cm^{-1}) calculated from an element of $\bar{\Omega}$ corresponding to an intermolecular stretch mode (j th) vibration, in which the contribution from antisymmetric translation of the two monomers is predominant. As can be seen from this figure, the frequencies are nearly constant across a wide range of n (dimension) values, although there are some anomalies at positions where the fidelity indices decrease significantly ($n = 36$ and 37 for acetic acid and $n = 30$ and 31 for trichloroacetic acid; see Fig. 2). The reproduced frequencies are nearly identical to those obtained with full-dimensional calculations (i.e., without coarse-graining) at the minimum reliable

dimension or higher. These results clearly show that with respect to the given coordinate system, the ratio of the stiffness constant to inertial load is adequately maintained irrespective of the coarse-graining dimension. We examined the modal masses calculated using Eq. 6. Figure 5 shows the values of λ_j^2 , the j th diagonal elements of Λ^2 calculated in the coarse-graining space, for various dimension values. Except the data for the trichloroacetic acid dimer, the values are almost constant over a wide range of n . A stepwise increase at $n = 24$ for trichloroacetic acid is probably connected to the stepwise increase in $\Phi_{\text{Tx,Tx}}$ at the same dimension. Errors in the apparent stiffness seem to be compensated for by errors in the apparent inertial load when the frequencies are reproduced using Eq. 14. The ideal modal mass matrix is diagonal, but the real Λ^2 is not. In previous studies [21, 22], we approximated Λ^2 to be diagonal when obtaining the normalized \mathbf{U} matrix (Eq. 4). However, the above consideration suggests that an arbitrary modification of Λ^2 may cause unpredictable errors in the subsequent computational procedure. Therefore, for all the calculations in the present study, we used Λ^2 , as it is nondiagonal. As a result, in contrast to our previous prediction [21], the behavior of the modal masses is not a good measure of the completeness of the selected coarse-graining basis set.

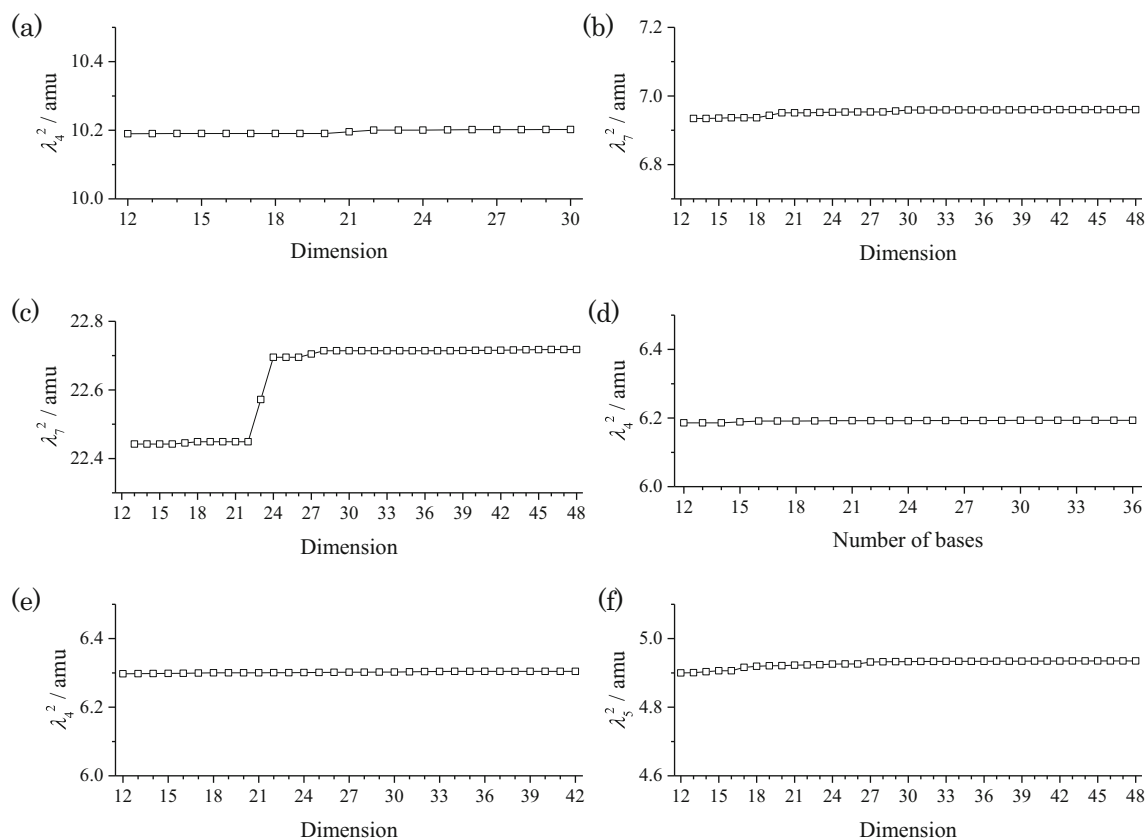


Fig. 5 a–f The modal mass corresponding to stretching of the intermolecular bond as a function of the coarse-graining dimension for dimers of **a** formic acid, **b** acetic acid, **c** trichloroacetic acid, **d** formamide, **e** formamidine, and **f** urea

Conclusions

We verified the reliability of our method of coarse-graining intermolecular vibrations by quantifying the fidelity of the restored MWD vectors. We proposed four fidelity indices F_1 to F_4 , each of which measured the similarity between the original and restored MWD vectors in different ways. Of those four indices, F_4 , the determinant of the product of the two MWD matrices, was the most severe criterion, and we tentatively proposed an index value of 0.8 as the lowest value that would indicate reliable coarse-graining calculations. When this criterion was applied, the minimum reliable dimension for the molecular dimers was found to be 12 or 14, depending on the degrees of freedom of the LF molecular motion. In contrast to an apparent dependence of the stiffness constants on the coarse-graining dimension n , the eigenfrequencies and the modal masses were not very sensitive to n . These findings suggest that coarse-graining analyses using our method are useful for deriving the effective stiffnesses of intermolecular forces, facilitating the efficient calculation of LF-mode vibrations for large systems such as molecular crystals or biomacromolecules. A formulation for calculating LF phonons using coarse-grained force constants is now being prepared and will be published elsewhere.

Acknowledgements This work was supported (in part) by a Grant for Basic Science Research Project in the fiscal year of 2016 received from The Sumitomo Foundation.

Appendix: Definition of the coarse-graining matrix

Based on our previous formulation, we can write matrix $\tilde{\mathbf{B}}$ to transform the atomic displacement vectors of an X-meric molecular assembly from a Cartesian to an internal coordinate system as Eq. A1. Provided that the constituent molecules are nonlinear, matrix $\tilde{\mathbf{B}}$ is composed of \mathbf{B}_m° , representing twelve basic motions, including six translational modes (T_x, T_y, T_z) and six rotational modes (R_x, R_y, R_z). \mathbf{b}_m is a set of vibrational motions selected from the first to the $(3N_m - 6)$ th normal mode, where m designates the constituents I, II, ..., X.

$$\tilde{\mathbf{B}} = \begin{pmatrix} \mathbf{B}_I^\circ & 0 & 0 & 0 & \mathbf{b}_I & 0 & 0 & 0 \\ 0 & \mathbf{B}_{II}^\circ & 0 & 0 & 0 & \mathbf{b}_{II} & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \mathbf{B}_X^\circ & 0 & 0 & 0 & \mathbf{b}_X \end{pmatrix} \quad (\text{A1})$$

To construct an appropriate $\tilde{\mathbf{B}}$ matrix, we need to choose the set of vectors \mathbf{b}_m that best represents the intramolecular vibrational motions, as extracted from the results of normal mode analysis for the constituent in its isolated state. However, there is a slight difference between the optimized

structure of a constituent in its isolated state and its optimized structure in a molecular assembly. To reduce errors resulting from this discrepancy, we adopted a method to adjust the molecular orientation, namely arranging the principal axes of inertia (a, b, c) of a constituent molecule in its isolated state to coincide with those in the molecular assembly. Under this coordination setting, submatrix \mathbf{B}_m° cannot be uniquely determined, but there is an arbitrariness with respect to the definition of the rotating axes. Even though matrix $\tilde{\mathbf{B}}$ is not necessarily orthogonal, the MWD vectors (columns in $\mathbf{M}^{1/2}\tilde{\mathbf{B}}$) will be orthogonal when the rotational axes of a constituent coincide with its principal axes of inertia, and only then will $\Gamma^{-1} = \tilde{\mathbf{B}}^T\mathbf{M}\tilde{\mathbf{B}}$ become diagonal. Even when they do not coincide with each other, the definitions of the axes of translation (x, y, z) and rotation (a, b, c) do not influence the reliability of coarse-graining. We should note that the parameters obtained (stiffness and inertial load) are represented based on the given coordinate system and designated as matrix elements like components (T_x, T_x) or (R_a, R_b). Similar attention should be paid when we construct the \mathbf{C} matrix. The first six columns of \mathbf{C} cannot be uniquely determined, but there is an arbitrariness with respect to the definition of the axes of rotation. The MWD vectors (columns in $\mathbf{W} = \mathbf{M}^{1/2}\mathbf{C}$) are orthogonal to each other only when the rotational axes of the molecular assembly coincide with the principal axes of inertia. Although the definitions of the axes of translation and rotation do not affect the result of the coarse-graining calculation, the orthogonality of \mathbf{W} influences matrix \mathbf{R} and hence the maximum values of the fidelity indices F_1 to F_4 .

For example, the matrix for a dimeric molecular system is given by

$$\tilde{\mathbf{B}} = \begin{pmatrix} \mathbf{B}_I^\circ & 0 & \mathbf{b}_I & 0 \\ 0 & \mathbf{B}_{II}^\circ & 0 & \mathbf{b}_{II} \end{pmatrix} \quad (\text{A2})$$

In our previous studies [21–23], we used $\tilde{\mathbf{B}}_{C_i}$ specified for a molecular dimer with C_i symmetry. However, this treatment does not greatly affect the result because $\tilde{\mathbf{B}}_{C_i}$ and $\tilde{\mathbf{B}}$ are connected by the following relation:

$$\tilde{\mathbf{B}}_{C_i} = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{B}_I^\circ & \mathbf{B}_I^\circ & \mathbf{b}_I & \mathbf{b}_I \\ \mathbf{B}_{II}^\circ & -\mathbf{B}_{II}^\circ & \mathbf{b}_{II} & -\mathbf{b}_{II} \end{pmatrix} = \tilde{\mathbf{B}}\mathbf{P}_{C_i} \quad (\text{A3})$$

using the symmetry-adapted transformation \mathbf{P}_{C_i} composed of the $(6 + \nu) \times (6 + \nu)$ unit matrix \mathbf{E} (ν is the number of intramolecular vibrational modes included in the basis set):

$$\mathbf{P}_{C_i} = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{E} & \mathbf{E} \\ \mathbf{E} & -\mathbf{E} \end{pmatrix} \quad (\text{A4})$$

If we use $\tilde{\mathbf{B}}$ to obtain the various matrices (stiffness, inertial load, displacement, and mass-weighted displacement), the parameters are designated as matrix elements like (T_x, T_x) and

(R_a , R_b) components. Using the \mathbf{P}_{C_i} matrix, we can transform them to the corresponding values for a given coordination system such as the C_i -specific coordination system in which the parameters are designated like ($T_x - T_x$, $T_x - T_x$) or ($R_a + R_a$, $R_b - R_b$) components:

$$\Phi_{C_i} = \mathbf{P}_{C_i}^T \Phi \mathbf{P}_{C_i} \quad (\text{A5})$$

$$\Gamma_{C_i}^{-1} = \mathbf{P}_{C_i}^T \Gamma^{-1} \mathbf{P}_{C_i} \quad (\text{A6})$$

$$\Xi_{C_i} = \mathbf{P}_{C_i}^T \Xi. \quad (\text{A7})$$

Using an appropriate transformation matrix \mathbf{P} , we can also redefine the directions of the x , y , and z axes in order to conveniently clarify the physical meaning of an obtained parameter (e.g., the stiffness of a specific hydrogen bond). In the present study, we defined the xy plane as the least-squares plane of the four atoms participating in the double hydrogen bond and the x -axis as the average of two hydrogen-bonding vectors projected onto the xy plane. Since the obtained parameters of stiffness are represented based on a non-symmetry-adapted coordinate system, those corresponding to translation along the x -axis are designated Φ_{T_x, T_x} .

References

- Plusquellic DF, Siegrist K, Heilweil EJ, Esenturk O (2007) Applications of terahertz spectroscopy in biosystems. *ChemPhysChem* 8:2412–2431. <https://doi.org/10.1002/cphc.200700332>
- Heyden M, Bründermann E, Heugen U et al (2008) Long-range influence of carbohydrates on the solvation dynamics of water—answers from terahertz absorption measurements and molecular modeling simulations. *J Am Chem Soc* 130:5773–5779. <https://doi.org/10.1021/ja0781083>
- Williams MRC, True AB, Izmaylov AF et al (2011) Terahertz spectroscopy of enantiopure and racemic polycrystalline valine. *Phys Chem Chem Phys* 13:11719–11730. <https://doi.org/10.1039/c1cp20594c>
- Taday PF, Bradley IV, Arnone DD, Pepper M (2003) Using terahertz pulse spectroscopy to study the crystalline structure of a drug: a case study of the polymorphs of ranitidine hydrochloride. *J Pharm Sci* 92:831–838. <https://doi.org/10.1002/jps.10358>
- Zhang F, Kambara O, Tominaga K et al (2014) Analysis of vibrational spectra of solid-state adenine and adenosine in the terahertz region. *RSC Adv* 4:269. <https://doi.org/10.1039/c3ra44285c>
- Giraud G, Wynne K (2003) A comparison of the low-frequency vibrational spectra of liquids obtained through infrared and Raman spectroscopies. *J Chem Phys* 119:11753–11764. <https://doi.org/10.1063/1.1623747>
- Della Valle GR, Venuti E et al (2004) Intramolecular and low-frequency intermolecular vibrations of pentacene polymorphs as a function of temperature. *J Phys Chem B* 108:1822–1826. <https://doi.org/10.1021/jp0354550>
- Luckhaus D (2010) Hydrogen exchange in formic acid dimer; tunneling above the barrier. *Phys Chem Chem Phys* 12:8357–8361. <https://doi.org/10.1039/c001253j>
- Xue Z, Suhm MA (2009) Probing the stiffness of the simplest double hydrogen bond: the symmetric hydrogen bond modes of jet-cooled formic acid dimer. *J Phys Chem* 131:054301. <https://doi.org/10.1063/1.3191728>
- Müller A, Talbot F, Leutwyler S (2001) Intermolecular vibrations of the jet-cooled 2-pyridone-2-hydroxypyridine mixed dimer, a model for tautomeric nucleic acid base pairs. *J Chem Phys* 115:5192–5202. <https://doi.org/10.1063/1.1394942>
- Müller A, Talbot F, Leutwyler S (2000) Intermolecular vibrations of jet-cooled (2-pyridone)₂: a model for the uracil dimer. *J Chem Phys* 112:3717–3725
- Müller A, Talbot F, Leutwyler S (2002) S₁/S₂ exciton splitting in the (2-pyridone)₂ dimer. *J Chem Phys* 116:2836–2847. <https://doi.org/10.1063/1.1434987>
- Taddei G, Bonadeo H, Marzocchi MP, Califano S (1973) Calculation of crystal vibrations of benzene. *J Chem Phys* 58:966–978
- Neto N, Righini R, Califano S (1978) Lattice dynamics of molecular crystals using atom-atom and multipole-multipole potentials. *Chem Phys* 29:167–179
- Zhang F, Wang H, Tominaga K, Hayashi M (2016) Mixing of intermolecular and intramolecular vibrations in optical phonon modes: terahertz spectroscopy and solid-state density functional theory. *WIREs Comput Mol Sci* 6:386–409. <https://doi.org/10.1002/wcms.1256>
- Day GM, Zeitler JA, Jones W et al (2006) Understanding the influence of polymorphism on phonon spectra: lattice dynamics calculations and terahertz spectroscopy of carbamazepine. *J Phys Chem B* 110:447–456
- Day GM (2011) Current approaches to predicting molecular organic crystal structures. *Cryst Rev* 17:3–52. <https://doi.org/10.1080/0889311X.2010.517526>
- Zhang F, Hayashi M, Wang HW et al (2014) Terahertz spectroscopy and solid-state density functional theory calculation of anthracene: effect of dispersion force on the vibrational modes. *J Chem Phys* 140:174509. <https://doi.org/10.1063/1.4873421>
- Zhang F, Wang H-W, Tominaga K, Hayashi M (2015) Intramolecular vibrations in low-frequency normal modes of amino acids: L-alanine in the neat solid state. *J Phys Chem A* 119:3008–3022. <https://doi.org/10.1021/jp512164y>
- Williams MRC, Aschaffenburg DJ, Ofori-Okai BK, Schmuttenmaer CA (2013) Intermolecular vibrations in hydrophobic amino acid crystals: experiments and calculations. *J Phys Chem B* 117:10444–10461. <https://doi.org/10.1021/jp406730a>
- Houjou H (2009) Coarse graining of intermolecular vibrations by a Karhunen-Loève transformation of atomic displacement vectors. *J Chem Theor Comput* 5:1814–1821. <https://doi.org/10.1021/cr900169f>
- Houjou H (2011) Evaluation of coupling terms between intra- and intermolecular vibrations in coarse-grained normal-mode analysis: does a stronger acid make a stiffer hydrogen bond? *J Chem Phys* 135:154111. <https://doi.org/10.1063/1.3652102>
- Houjou H (2017) Modelling intra- and intermolecular vibrations under the harmonic oscillator approximation: from symmetry-adapted to coarse-grained coordinate approaches. *J Math Chem* 55:532–551. <https://doi.org/10.1007/s10910-016-0692-x>
- Wilson Jr EB, Decius JC, Cross PC (1980) *Molecular vibrations*. Dover, New York
- Frisch MJ, Trucks GW, Schlegel HB, et al (2013) *Gaussian 09w*, revision D. Gaussian Inc., Wallingford
- Jmol Development Team (2018) Jmol: an open-source Java viewer for chemical structures in 3D. <http://www.jmol.org/>