JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Infrared Spectroscopy of Fluxional Molecules from (ab Initio) Molecular Dynamics: Resolving Large-Amplitude Motion, Multiple Conformations, and Permutational Symmetries

Gerald Mathias,*[,†,‡] Sergei D. Ivanov,[†,§] Alexander Witt,[†,∥] Marcel D. Baer,[†,⊥] and Dominik Marx[†]

[†]Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany

ABSTRACT: The computation of vibrational spectra of complex molecules from time correlation functions generated by ab initio molecular dynamics simulations has made lively progress in recent years. However, the analysis of such spectra, i.e., the assignment of vibrational bands to atomic motions, is by no means straightforward. In a recent article [*J. Chem. Theory Comput.* **2011**, *7*, 2028−2039], Mathias and Baer presented a corresponding analysis method that derives generalized normal coordinates (GNCs) from molecular dynamics trajectories, which furnish band positions, band shapes, and infrared intensities of the separated vibrational modes. This vibrational analysis technique relies on the usual quasi-rigidity assumption; i.e., atomic motions are described by small oscillations around a single reference structure. This assumption, however, breaks down if the molecule undergoes large-amplitude motion and visits different conformations along the trajectory or if the same conformation can be adopted by a different ordering of the atoms, i.e., if permutational symmetries have to be considered. Here, we present an extension of the GNC method that handles such cases by considering multiple reference structures, both for different conformations and for permutational symmetries. By introducing a projection technique and computing probabilities that assign the time frames of the trajectories to these reference structures, the vibrational spectra are split into conformational contributions via a consistent time correlation formalism. For each conformation, the permutational symmetries are resolved, which permits one to determine conformation-local GNCs for the band assignment. The working principle and the virtues of this generalization are demonstrated for the simple case of a methyl group rotation. This is followed by an application to a more intricate case: Upon replacing one proton by a deuteron in protonated methane, CH_5^+, significant changes of its infrared spectrum have been observed since the CH_4D^+ isotopologue features five different isotopomers. Here, a total of 120 conformational and permutational references are required in the projection scheme in order to capture the frequent and versatile structural transitions of this small but utmost floppy molecule and to assign its infrared spectrum. The extended GNC method is general. Thus, it can be applied readily to systems that require more than one reference structure, and it can be transferred to other theoretical spectroscopies that are formulated in terms of time correlation functions.

1. INTRODUCTION

Vibrational spectroscopy is a powerful experimental technique to monitor and explore (bio)chemical processes and molecular structures.[1−6] The computation and especially the decoding of such spectra is a challenge to theory, in particular if the underlying potential energy surfaces are strongly anharmonic, permit large-amplitude motion, or are significantly modified by interactions within (micro- or bulk) solvent environments. In recent years, ab initio molecular dynamics (AIMD) simulations[7] have emerged as a promising approach to theoretical vibrational spectroscopy, especially for systems that are too complex or contain too many coupled degrees of freedom[8−28] to be readily assessed by quasi-exact quantum mechanical treatments of nuclear motion. The appeal of the AIMD approach is that it works almost out of the box after choosing an appropriate electronic structure method. Since energies, forces, and dipoles can be computed on the fly, AIMD does not require the parametrization of a potential energy surface (PES) and dipole surfaces, which is commonly a challenging[29−31] and sometimes unfeasible task, if the number of atoms grossly exceeds the current state-of-the-art capabilities of PES parametrization, e.g., in the condensed phase. Moreover, even fairly small but fluxional ("floppy") systems can be truly challenging in this respect due to the presence of multiple similarly important minima or broad and shallow regions on the PES, which all require special care in the course of parametrization.[32]

Of course, the AIMD approach to theoretical vibrational spectroscopy also faces drawbacks because of its classical approximation of quantum-mechanical motion of the nuclei, while obeying quantum mechanics for the electrons. As a result, the classical Maxwell−Boltzmann distribution may grossly differ from the proper quantum-mechanical Bose−Einstein distribution in the relevant parts of the frequency spectrum, not to mention neglecting symmetry effects on rotations and ro-vibrations due to nuclear spin. On the one hand, due to this discrepancy, the vibrational bands of strongly anharmonic systems may exhibit a pronounced temperature dependence in classical simulations,[33−35] which has to be carefully checked. This cross-checking, on the other hand, may provide deeper insights into the mutual anharmonic couplings of vibrational modes, and thus it may even serve as a sort of analysis tool.[35]

Although the computation of infrared (IR)[36,37] and Raman[36,38] cross-sections from molecular dynamics is both well-defined and straightforward in the frame of the "Heisenberg approach to theoretical spectroscopy", the assignment of the resulting bands to atomic motions is not. In recent years, a few techniques have been suggested to derive approximate normal modes from molecular dynamics simulations and thereby to assign the vibrational

Received: September 21, 2011
Published: November 29, 2011

ACS Publications © 2011 American Chemical Society

bands in the spectra.[16−18,21,39−43] These methods all rely on an equipartition assumption to a lesser or greater extent, which leads to problems in cases when equipartition is hard to achieve.[41] In the preceding paper[44] (paper I), Mathias and Baer introduced an approach that is free from such assumptions. The method constructs generalized normal coordinates (GNCs) for the vibrational analysis by an orthonormal transform of mass-weighted coordinates. The transformation is constructed such that it minimizes the norm of the GNCs mutual Fourier-transformed cross-correlation functions. The prospects of the GNC algorithm had been exemplified by the vibrational analysis of *trans*-isoprene (2-methyl-1,3-butadiene), where the IR bands computed from AIMD had been assigned to the underlying atomic motions. In addition, it has already been successfully applied to assign the IR spectra of floppy molecules and complexes such as protonated methane as well as microsolvated hydronium and Zundel cations.[24,27,35]

Importantly, the GNC algorithm presented in paper I, as well as all other predecessor techniques, is based on the commonly used assumption of quasi-rigidity, namely, that it is sufficient to consider a single reference structure of the molecule and that the dynamical motion can be expressed as (small-amplitude) oscillations around this reference. This assumption is however not justified if large-amplitude motion overrides small-amplitude oscillations such that the trajectories ultimately visit different conformations of a molecule or intramolecular chemical rearrangements lead to new configurations of the atoms. All of these changes correspond to switching between *different minima* on the PES. Examples for conformational changes are *cis*−*trans* isomerizations of polyenes or peptides, whereas different configurations (isomers) may result from intramolecular proton transfer or other chemical reactions. Note that, only for simplicity, we will use the term "conformations" also for different configurations in the aforementioned sense.

A similar problem is rooted in the exploration of the PES by classical dynamics as such. Here, the trajectory can visit chemically *equivalent minima* of the PES, which are related by a permutation of the numbering of equivalent atoms. This problem was already faced in the vibrational analysis of the seemingly simple molecule *trans*-isoprene in paper I . Because of the low rotational barrier of its methyl group (about 11 kJ mol^{-1}),[45] the latter rotates nearly freely under ambient conditions and thus visits three chemically equivalent minima of the PES (degenerate equilibrium structures) separated by methyl group rotations of 120°. The assumption of quasi-rigidity employed in paper I did not account for these different minima, and correspondingly, the methyl group vibrations were not properly disentangled. As a result of this approximation, they still showed correlations among the corresponding stretching and bending modes. Employing internal coordinates for the vibrational analysis in paper I separated the symmetric modes, e.g., the symmetric methyl stretch, but the antisymmetric modes were still coupled.[44]

For the vibrational analysis of the aforementioned cases, one requires not only one but distinct reference structures for each conformation. If multiple reference structures would be considered, the dynamics along the trajectory could again be described as small-amplitude oscillations around the respective reference structure during the residence time within the corresponding "conformation", complemented by the transitions between all of the "conformations" available. Using the same idea, one would be able to handle permutational symmetries, which would be represented by additional (permuted) reference structures for each and every conformation. If the total vibrational spectrum would be dissected into separate contributions of each conformation in this manner, it could be analyzed in terms of the ensemble of thermally populated conformations.

Of course, simply chopping the trajectories into pieces corresponding to a specific instantaneous conformation would be sufficient only if the residence times in each conformation were long compared to the typical periods of the vibrational modes. In this case, the total spectrum could be synthesized a posteriori, including its band assignment, by simply adding the (assigned) conformation-specific spectra using weights according to their free energy differences as obtained from their relative contribution to the total trajectory. However, well-known intricacies in the assignment of IR spectra arise in the opposite limit where switching between conformations is rather frequent on the time scale set by small-amplitude oscillatory motion. In a sense, this limit even defines the aforementioned class of floppy or fluxional molecules that are characterized by effects due to large-amplitude motion. For such cases, a more elaborate and general approach is clearly required.

In this article, we present a generalization of the GNC scheme that can handle systematically multiple reference structures in the vibrational analysis. These structures can either define different conformations of the molecule or represent permutations of atoms of the same species within a given conformation. The extended GNC method is useful for the plethora of cases where a single equilibrium or reference structure is not sufficient to understand vibrational spectra. Furthermore, the very same idea can be readily transferred to other analysis schemes of response properties of molecular systems which are based on time correlation functions as required by the Heisenberg approach to theoretical spectroscopy.

In the following theory section, we first discuss how to dissect the computed IR absorption spectrum into conformational contributions. Then, a similar formalism is developed to compute auto- and cross-correlation functions of the mass-weighted velocities that are specific for each conformation and each permutational pattern. Recombining the permutation-specific correlation functions by employing proper symmetry operations yields correlation functions that are specific only to the conformations. These correlation functions are then used to determine the GNCs of the respective conformations. Further, a projection scheme is introduced that determines the probability for each frame of the trajectory to belong to a certain conformation and permutational pattern, which is used to split the total spectrum quantitatively into conformation-specific contributions.

In the results part, we show how to resolve the permutational symmetry of the methyl group of isoprene to provide a first simple example and test case of the extended GNC method. A much more challenging application in this respect is the analysis of the IR spectra of the small but highly fluxional protonated methane molecule, CH_5^+,[17,46−53] and its partially deuterated isotopologues,[54] $CH_nD_{5-n}^+$, $n = 0−5$, for which this generalization was first successfully employed.[24]

Because of the low intramolecular barriers,[55] CH_5^+ undergoes vivid scrambling dynamics; that is, all protons visit each of the five possible binding sites during the simulation. This leads to 5! = 120 reference structures that are necessary for the vibrational analysis of this small molecule. The merits of the extended GNC approach are exemplified on the partially deuterated CH_4D^+ isotopologue, leading to the fully assigned IR spectrum. To achieve this, the total IR spectrum is split into the contributions

of its isotopomers,[54] for which in turn the vibrational bands are assigned separately.

2. THEORY

We start the derivation of the multiple configuration (extended) GNC scheme by first developing a formalism to split the computed IR spectrum into the underlying conformational contributions. The linear total IR absorption coefficient[36]

$$\alpha^{\mathrm{QM}}(\omega) = \frac{2\pi\omega(1-\mathrm{e}^{-\hbar\beta\omega})}{3V\hbar cn(\omega)} \int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{-i\omega t}\langle \hat{\mathbf{M}}(0)\, \hat{\mathbf{M}}(t)\rangle \tag{1}$$

is given within the Heisenberg approach to theoretical spectroscopy by the Fourier transform of the autocorrelation function of the time-dependent Heisenberg dipole operator $\hat{\mathbf{M}}(t)$, where $\langle \cdot \cdot \rangle$ denotes the quantum-statistical ensemble average at temperature T. The prefactor depends on the sample volume V, the refractive index $n(\omega)$, and the inverse temperature $\beta = 1/k_{\mathrm{B}}T$. To obtain a corresponding $\alpha(\omega)$ from classical dynamics trajectories, one approximates the autocorrelation function of Heisenberg's dipole *operator* $\hat{\mathbf{M}}(t)$ by the autocorrelation function of the dipole *moment* $\boldsymbol{\mu}(t)$. Here, this replacement is done after rewriting eq 1 in terms of the Kubo-transformed quantum correlation function, which readily yields

$$\alpha(\omega) = \frac{2\pi\beta\omega^2}{3Vcn(\omega)} \int \mathrm{d}t\, \mathrm{e}^{-i\omega t}\langle \boldsymbol{\mu}(0)\, \boldsymbol{\mu}(t)\rangle \tag{2}$$

as our basic working equation.[56] Note that many different so-called "quantum corrections factors" have been introduced in the literature mainly to impose the detailed balance condition and that the "harmonic quantum correction" corresponds to the prefactor in eq 2, see ref 56.

Our first goal is to determine individual absorption coefficients $\alpha_\xi(\omega)$ for each reference conformation ξ. Thus, we have to split the dipole autocorrelation function into conformational contributions. Given that we can classify the frames of a trajectory, and thereby assign each frame at time t with a certain probability $p_\xi(t)$ to conformation ξ satisfying

$$\sum_\xi p_\xi(t) = 1, \quad 0 \le p_\xi(t) \le 1 \tag{3}$$

we can split the ensemble average of the dipole autocorrelation function

$$\langle \boldsymbol{\mu}(0)\, \boldsymbol{\mu}(t)\rangle = \sum_\xi \langle p_\xi(0)\, \boldsymbol{\mu}(0)\, \boldsymbol{\mu}(t)\rangle \tag{4}$$

into a sum of conformation-specific correlation functions. These specific correlation functions are restricted to trajectories visiting conformation ξ at $t = 0$. Accordingly, we define the conformation specific IR absorption coefficient as

$$\alpha_\xi(\omega) = \frac{2\pi\beta\omega^2}{3Vcn(\omega)} \int \mathrm{d}t\, \mathrm{e}^{-i\omega t}\langle \boldsymbol{\mu}_\xi(0)\, \boldsymbol{\mu}(t)\rangle \tag{5}$$

where we have introduced the probability-weighted dipole moment $\boldsymbol{\mu}_\xi(t) \equiv p_\xi(t)\, \boldsymbol{\mu}(t)$. Thus, if one can classify the trajectories according to conformations, the computation of conformation-specific spectra is straightforward and mathematically exact. We will discuss possible projection strategies further below. Note that the Fourier transform of the correlation functions introduced here and in the following can be efficiently calculated by employing the cross-correlation theorem as discussed in detail in paper I .

For the subsequent vibrational analysis within the extended GNC scheme, which is the core task for any spectral assignment, the splitting procedure becomes slightly more complicated. We recall from paper I that for GNC we have to compute a tensorial version of the vibrational density of states (VDOS) $\boldsymbol{\Theta}$: $\mathbb{R} \rightarrow \mathbb{R}^{3N} \times \mathbb{R}^{3N}$ for N atoms given by

$$\boldsymbol{\Theta}(\omega) = \frac{\beta}{\pi} \int \mathrm{d}t\, \mathrm{e}^{-i\omega t}\langle \dot{\mathbf{c}}(0) \otimes \dot{\mathbf{c}}(t)\rangle \tag{6}$$

with $\dot{\mathbf{c}}(t) = \mathbf{M}^{1/2}\mathbf{v}(t)$, $M_{ij} = \delta_{ij}m_i$, being the trajectory of the mass-weighted atomic velocities $\mathbf{v}$. Here, the correlations are evaluated as an outer product $\dot{\mathbf{c}}\otimes\dot{\mathbf{c}} = \dot{\mathbf{c}}\dot{\mathbf{c}}^T$ of the mass-weighted velocity components, yielding their autocorrelations on the diagonal of $\boldsymbol{\Theta}$ and their mutual cross-correlations in the off-diagonal elements of $\boldsymbol{\Theta}$. To obtain the generalized normal coordinates, $\boldsymbol{\Theta}$ is brought to diagonal form as close as possible by an orthonormal transform of the $\dot{\mathbf{c}}$, thereby minimizing the cross-correlations.[44]

For a conformation-specific VDOS, i.e., for $\boldsymbol{\Theta}_\xi(\omega)$, we would like to employ the same splitting procedure as we have done to obtain the corresponding IR spectra, i.e., α_ξ. However, we have to additionally consider possible permutational symmetries, i.e., reference structures $\mathbf{x}^{\mathrm{ref}}_{\xi,i}$ belonging to the same conformation ξ but differing from a default reference $\mathbf{x}^{\mathrm{ref}}_{\xi,0}$ by a permutation

$$\mathbf{x}^{\mathrm{ref}}_{\xi,i} = \mathscr{P}_{\xi,i}\mathbf{x}^{\mathrm{ref}}_{\xi,0} \tag{7}$$

where $\mathscr{P}_{\xi,i}$ is a permutation operator interchanging only the positions of atoms of the same species. Therefore, we have to classify the trajectory additionally according to the permutational pattern i with probabilities $p_{\xi,i}$, which satisfy

$$\sum_i p_{\xi,i}(t) = p_\xi(t) \tag{8}$$

We can weight the $\dot{\mathbf{c}}$ with these probabilities, which gives $\dot{\mathbf{c}}_{\xi,i}(t) = p_{\xi,i}(t)\, \dot{\mathbf{c}}(t)$, and introduce conformation- and permutation-specific

$$\boldsymbol{\Theta}_{\xi,i}(\omega) = \frac{\beta}{\pi} \int \mathrm{d}t\, \mathrm{e}^{-i\omega t}\langle \dot{\mathbf{c}}_{\xi,i}(0) \otimes \dot{\mathbf{c}}(t)\rangle \tag{9}$$

which sum up to the global $\boldsymbol{\Theta}$.

Having to evaluate a separate $\boldsymbol{\Theta}_{\xi,i}(\omega)$ for each permutational pattern is however neither desirable nor necessary. Employing the back-transforms $\mathscr{P}^{-1}_{\xi,i}$, we can bring all permutational patterns to the same reference structure and calculate

$$\boldsymbol{\Theta}_\xi(\omega) = \sum_i \mathscr{P}^{-1}_{\xi,i}\, \boldsymbol{\Theta}_{\xi,i}(\omega)\, \mathscr{P}_{\xi,i} \tag{10}$$

as the sum of the transformed $\boldsymbol{\Theta}_{\xi,i}$, or computationally more convenient

$$\boldsymbol{\Theta}_\xi(\omega) = \frac{\beta}{\pi} \int \mathrm{d}t\, e^{-i\omega t} \sum_i \langle \mathscr{P}^{-1}_{\xi,i}\, \dot{\mathbf{c}}_{\xi,i}(0) \otimes \mathscr{P}^{-1}_{\xi,i}\, \dot{\mathbf{c}}(t)\rangle \tag{11}$$

Sometimes it is preferential to analyze the molecular motion in terms of a set of internal coordinates $\mathbf{s}_\xi(\mathbf{x}) = (s_{\xi,1}(\mathbf{x}), \ldots, s_{\xi,3N-6}(\mathbf{x}))$, specifically chosen for conformation ξ (see paper I for details). Here, the permutational symmetries can be resolved in two ways: either one applies the permutation operators directly to the Cartesians

$$\mathbf{s}_{\xi,i}(\mathbf{x}) = \mathbf{s}_\xi(\mathscr{P}^{-1}_{\xi,i}\mathbf{x}) \tag{12}$$

and then evaluates the internal coordinates, which is simple to implement but may involve many redundant evaluations of internal coordinates, or alternatively, $\mathscr{P}_{\xi,i}$ can directly act on the definitions of the $\mathbf{s}_\xi$ to give

$$\mathbf{s}_{\xi,i}(\mathbf{x}) = (\mathscr{P}_{\xi,i}\mathbf{s}_\xi)(\mathbf{x}) \tag{13}$$

That is, one transforms the internal coordinate definitions similar to what has been done with the reference structure according to eq 7. Such internal coordinates are commonly linear combinations of primitive coordinates such as bond lengths, angles, and dihedrals, where a large part of these primitives will be invariant under coordinate permutation or the permuted primitive will be already present in the set of primitives corresponding to the reference structure. Only few primitives will lead to new primitives upon employing the permutation operator. Thus, the set of all relevant primitives has to be determined once for the whole trajectory, and the internal coordinates for a given permutational reference i can be efficiently calculated from these primitives. On the basis of these internal coordinates, the tensorial VDOS evaluates to

$$\mathbf{\Theta}_\xi(\omega) = \frac{\beta}{\pi}\int \mathrm{d}t\, \mathrm{e}^{-i\omega t} \sum_i \langle p_{\xi,i}(0)\, \mathbf{K}^+\dot{\mathbf{s}}_{\xi,i}(0) \otimes \mathbf{K}^+\dot{\mathbf{s}}_{\xi,i}(t)\rangle \tag{14}$$

where $\mathbf{K}^+$ with $\mathbf{G}^{-1} = (\mathbf{K}^+)^T\mathbf{K}^+$ is a root of the inverse Wilson $\mathbf{G}$ matrix that transforms the internal coordinates back to Cartesians.[44,57] Note here that we have employed the common approximation[57] $\mathbf{G} = \mathbf{G}[\mathbf{s}_{\xi,0}] \approx \mathbf{G}[\mathbf{s}_{\xi,0}(\mathbf{x} = \mathbf{x}^{\mathrm{ref}}_{\xi,0})]$; i.e., we evaluate $\mathbf{G}$ solely at the position of the reference structure $\mathbf{x}^{\mathrm{ref}}_{\xi,0}$ and neglect the dependence of this matrix on the internal coordinates.

What remains is to introduce a computationally practical projection scheme to compute the probabilities $p_\xi(t)$ and $p_{\xi,i}(t)$. Here, we choose a general form

$$p_\xi(t) = \frac{\exp\left(-\frac{1}{2\sigma_\mathrm{c}^2}\min_i d_\mathrm{c}(\mathbf{x}(t), \mathbf{x}^{\mathrm{ref}}_{\xi,i})^2\right)}{\sum_\eta \exp\left(-\frac{1}{2\sigma_\mathrm{c}^2}\min_j d_\mathrm{c}(\mathbf{x}(t), \mathbf{x}^{\mathrm{ref}}_{\eta,j})^2\right)} \tag{15}$$

where $d_\mathrm{c}(\mathbf{x},\mathbf{x}')$ is a metric measuring the distance between $\mathbf{x}$ and $\mathbf{x}'$ and the sum over η in the denominator runs over all conformations considered. In order to determine the distance of a given frame to a conformation, we have to keep in mind its permutational symmetries. Therefore, we take the minimum of the distances to all possible permutational references i of a given conformation ξ. Note that $p_\xi(t)$ is properly normalized according to eq 3 by the sum of Gaussians in the denominator.

The width of the Gaussians σ_c determines the width of the switching region between two neighboring conformations. Correspondingly, large values of σ_c lead to slow and smooth transitions, whereas small values of σ lead to fast but possibly noisy transitions if the system oscillates in the switching region between two conformations. For the permutational probabilities, we choose similarly

$$p_{\xi,i}(t) = p_\xi(t)\frac{\exp\left(-\frac{1}{2\sigma_\mathrm{p}^2}d_\mathrm{p}(\mathbf{x}(t), \mathbf{x}^{\mathrm{ref}}_{\xi,i})^2\right)}{\sum_{j=1}^{j_{\max}} \exp\left(-\frac{1}{2\sigma_\mathrm{p}^2}d_\mathrm{p}(\mathbf{x}(t), \mathbf{x}^{\mathrm{ref}}_{\xi,j})^2\right)} \tag{16}$$

Finally, proper metrics $d_\mathrm{c}(\cdot,\cdot)$ and $d_\mathrm{p}(\cdot,\cdot)$ between the coordinate frames and the reference coordinates have to be chosen. These metrics should return similar nearest-neighbor distances

$$r^{nn}_\xi = \min_{\eta,j} d_\mathrm{c}(\mathbf{x}^{\mathrm{ref}}_{\eta,j}, \mathbf{x}^{\mathrm{ref}}_{\xi,0}) \tag{17}$$

for all conformations and

$$r^{\mathrm{nn}}_{\xi,i} = \min_j d_\mathrm{p}(\mathbf{x}^{\mathrm{ref}}_{\xi,j}, \mathbf{x}^{\mathrm{ref}}_{\xi,i}) \tag{18}$$

for all permutations of a given reference, such that transitions between two reference structures occur on similar length scales. These length scales can then be used to choose σ_c and σ_p.

One obvious choice for d_c or d_p is the mass-weighted root-mean-square deviation

$$d_{\mathrm{RMS}}(\mathbf{x},\mathbf{x}') = \min_{A\in\mathscr{O}_3, \mathbf{t}\in\mathbb{R}^3} \sqrt{\frac{\sum_{k=1}^N m_k(\mathbf{A}\mathbf{r}_k + \mathbf{t} - \mathbf{r}'_k)^2}{\sum_{k=1}^N m_k}} \tag{19}$$

where $\mathbf{t}$ is a translation and $\mathbf{A}$ is an orthonormal transform acting on the atomic coordinates $\mathbf{r}_k$ of $\mathbf{x}$, thereby aligning it with $\mathbf{x}'$. Here, all internal degrees of freedom of the molecule contribute to the distance between structures $\mathbf{x}$ and $\mathbf{x}'$. Of course, mass weighting can be omitted or replaced by another weighting scheme if this seems more suitable for the specific problem. Alternatively, one can choose a metric

$$d_{\mathrm{IC}}(\mathbf{x},\mathbf{x}') = \sqrt{\sum_{k=1}^{N_{\mathrm{IC}}} a_k(s_k(\mathbf{x}) - s_k(\mathbf{x}'))^2} \tag{20}$$

based on a set of N_{IC} internal coordinates s_k. The a_k's provide the possibility to combine different kinds of internal coordinates, e.g., bonds and dihedrals, by weighting their contributions. The evaluation of the internal coordinates does not require a structural alignment, and in many cases the choice of the s_k is straightforward. For example, to determine the rotation of a methyl group, one would choose the corresponding torsional angles connecting the protons and the carbon to the rest of the molecule. Similarly simple internal coordinates are also found for many typical conformational changes such as *cis/trans* isomerizations in polyens and peptides or transitions between ring puckers of sugars.[58]

3. COMPUTATIONAL METHODS

As outlined in the Introduction, we consider two problems, which are challenging to our algorithm due to fast transitions between reference structures, either due to switching between permutational patterns or different conformations or a combination of both.

The first illustrative example is the complete vibrational analysis of the methyl group of isoprene. Here, we use the data set described in paper I, comprising 41 trajectories of isoprene in the gas phase, each spanning 25 ps of Born−Oppenheimer ab initio dynamics.[7] Initial conditions have been drawn from the canonical ensemble at 300 K. The electronic structure was treated by the BLYP density functional at a density cutoff of 280 Ry within the CP2k simulation package;[59] details are given in paper I.

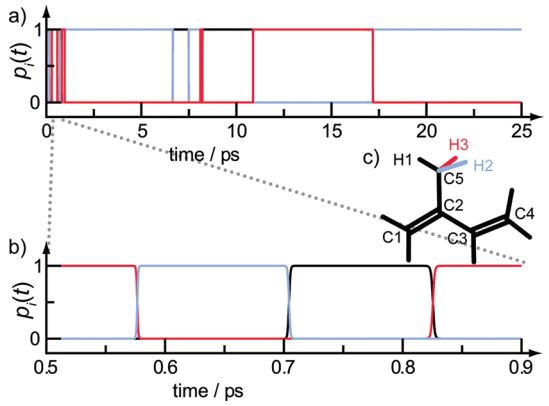

Figure 1. Probabilities $p_i(t)$ for the three permutational references of the methyl group of isoprene for one representative trajectory. In the default reference (black curve), whose structure is shown in panel c, proton H1 lies in the symmetry plane of the molecule. The other two references are obtained by rotating the methyl group by 120°, such that either H2 (blue curve) or H3 (red curve) is in the symmetry plane. The upper panel a shows the full trajectory, whereas b zooms in on a part of the first picosecond, when the methyl group rotates fast.

The second, more challenging example is the vibrational analysis of the CH_4D^+ isotopologue of protonated methane. The isolated CH_4D^+ molecule was simulated by Born−Oppenheimer MD as implemented in the CPMD[60] package. Here, the well-established setup discussed in detail in ref 17 was employed, i.e., the Perdew−Zunger local density approximation supplemented by Becke's exchange gradient correction at a plane−wave cutoff of 35 Ry, as used in previous publications on protonated methane.[15,24,28,46] Starting from canonically distributed initial conditions at a temperature of T = 110 K, 297 trajectories have been integrated for 10.2 ps with a time step of Δt = 0.34 fs.

Both systems have been analyzed with the program normcor, which was extended by the algorithms described in this article. Important parameters that have been chosen for these vibrational analyses will be discussed along with the results.

4. RESULTS AND DISCUSSION

To illustrate the ingredients of the extended GNC algorithm, we will lead through the vibrational analysis of the two examples that address the aspects of permutational symmetries and permutational symmetries combined with conformational changes.

4.1. Methyl Group Rotation: *trans*-Isoprene. We start the discussion with the comparatively simple case of a methyl group rotation, which switches only between chemically equivalent minima, and one does not have to consider different conformations. Already in paper I, it was shown that the vibrational modes of a methyl group cannot be clearly resolved if only one reference structure is considered. Because of the low barrier of the internal rotation of the methyl group, which is about 11 kJ mol^{-1} experimentally[45,61] and about 10 kJ mol^{-1} in our calculation employing the BLYP functional, the methyl group can easily cross this barrier under ambient conditions. Figure 1c introduces the atom labeling scheme and illustrates that due to the C_s symmetry of isoprene the three methyl protons cannot be considered equal. The two protons outside the mirror plane, H2 and H3, are at chemically equivalent sites but differ from the in-plane proton, H1. The rotation of the methyl group by 120° will therefore exchange the character of the three protons, which mathematically represents a cyclic permutation of their numbering. Correspondingly, two additional reference structures have to be considered for the methyl group, which yield the protons H2 and H3, respectively, at the in-plane position occupied by H1 in the default reference.

To discriminate between the reference structures along the trajectory, we used the metric (c.f. eq 20)

$$d_{\mathrm{IC}}(\mathbf{x},\mathbf{x}') = \sqrt{\sum_{j=1,3}\ \sum_{i=1,2,3} \left(\gamma_{\mathrm{Hi,Cj}}(\mathbf{x}) - \gamma_{\mathrm{Hi,Cj}}(\mathbf{x}')\right)^2} \tag{21}$$

constructed from all six dihedral angles $\gamma_{\mathrm{Hi,Cj}}$ across the C5−C2 bond. The permutation probabilities $p_{\xi,i}(t)$ according to eq 16 were computed using a width σ_p = 15° of the Gaussians; note that $p_{\xi}(t)$ = 1 because only one conformation of isoprene is considered.

Figure 1 illustrates the resulting probabilities for a selected trajectory. The whole trajectory in Figure 1a is characterized by long periods where only one reference structure is present. However, during the first picosecond, enlarged in Figure 1b, the methyl group rotates fast, and in turn, protons H3 (red curve) followed by H2 (blue), H1 (black), and finally again H3 (red) are located in the mirror plane. The transitions between the reference structures occur on a time scale of only 10 fs, during which time the probabilities switch continuously and monotonously between zero and one. During this first picosecond of this particular trajectory, the methyl group contains enough energy in its rotational degree of freedom to cross the barrier between two adjacent reference structures. Since the barriers between all reference structures have the same height, the methyl group visits all three reference structures in turn. When the methyl group loses rotational energy due to couplings to other modes, its energy does not suffice any longer for barrier crossings, and thus the molecule gets transiently trapped in one conformation, e.g., during the interval from 2 to 7 ps in Figure 1a, until the methyl group again gains enough energy from other modes to cross the barrier. This behavior, i.e., periods with frequent switching between the reference structures followed by longer residence times within one reference structure, is also found in the other 40 trajectories. Thus, sufficient sampling of initial conditions from a well thermalized trajectory needs to be carried out in order to properly compute canonically averaged infrared spectra.

After computing the tensorial VDOS from eq 11, which combines the contributions of the three reference structures, and minimizing its off-diagonal norm (see paper I), one obtains the bands of the generalized normal modes on the diagonal of , $\Theta(\omega)$. Figure 2 compares the nine vibrational modes associated with the methyl group obtained by (a) using a single reference structure (i.e., plain GNC according to paper I) and (b) considering three reference structures according to the extended GNC method introduced here. For both cases, we have employed a RMSD fitting procedure to remove the global translation and rotation of the molecule; similar results are obtained by using an internal coordinate transform instead (not shown).[44] Already the analysis with a single reference structure in Figure 2a, i.e., using plain GNC, shows the dominant characteristics of the modes. However, their spectra are not singly peaked but show sidebands at the position of other modes, which hint at a remaining coupling that cannot be disentangled by the plain GNC analysis. For example, the symmetric stretch (black curve)

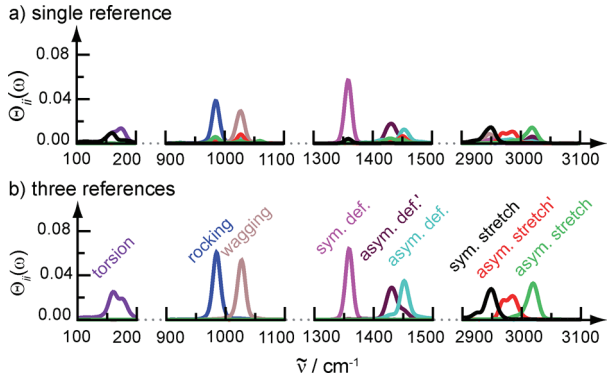

Figure 2. Localized modes $\mathbf{\Theta}_{ii}(\omega)$ obtained for (a) a single reference (as introduced in paper I) and (b) for three references within the extended GNC algorithm introduced here. The color code defines the mode labels and is the same for both graphs.

located at 2950 cm^{-1} has a small sideband at 150 cm^{-1} and seems to couple to the methyl torsion, which is also visible in the associated normal coordinate vectors $\mathbf{q}_i$. To evaluate the quality of the normal coordinates, we have computed their overlaps

$$\vartheta_i = |\mathbf{q}_i \cdot \mathbf{q}_i^{\mathrm{h}}| \tag{22}$$

with the corresponding normal coordinates $\mathbf{q}_i^{\mathrm{h}}$ obtained from the harmonic analysis described in paper I. Since the normal coordinates are normalized, $\vartheta_i = 1.0$ is obtained for $\mathbf{q}_i = \pm\mathbf{q}_i^{\mathrm{h}}$. The average over the nine methyl modes yields only $\langle\vartheta\rangle = 0.25$, implying that the normal coordinates of these modes strongly differ from the harmonic description. Only the symmetric deformation mode shows a significant resemblance at $\vartheta = 0.88$. Interestingly, this mode also shows the best resolved peak in Figure 2a with only very little coupling to other methyl modes.

In stark contrast, employing three reference structures in Figure 2b within the extended GNC scheme completely resolves these couplings, and only a small overlap between the two neighboring asymmetric deformation modes near 1440 cm^{-1} remains. In particular, the normal modes obtained for the deformation and stretching modes reflect the C_s symmetry of the molecule, where we have indicated the modes of odd character by a prime. Note that employing three reference structures hardly changes the peak positions of the vibrational bands, which differ at most by 2 cm^{-1} between Figure 2a and b.

The improved description resulting from using three reference structures also shows up in their overlaps θ_i of the resulting methyl normal coordinates with respect to the harmonic ones. They are larger than 0.98 for eight out of the nine methyl modes, thus indicating a perfect match. Only the methyl torsion around 160 cm^{-1} notably deviates between the extended GNC and the harmonic description with $\vartheta = 0.81$, which is caused by a different coupling to the torsion around the C2−C3 single bond having a similar frequency, 151 cm^{-1}.[44]

Thus, the generalization of the GNC algorithm fully resolves the permutational symmetry of the methyl group and cleanly identifies all modes from the AIMD trajectories. Next, we examine the extended GNC method for a much more involved case where, in addition, multiple conformational states have to be considered.

4.2. Multiple Conformations: CH_4D^+. Protonated methane, CH_5^+, is infamous for its vigorous scrambling dynamics, that is, it actively travels between all minima on its shallow PES.[46,47]

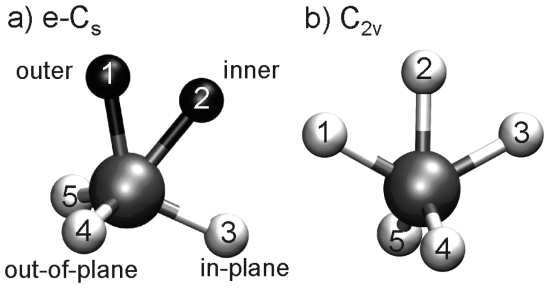

Figure 3. (a) The global minimum structure of the eclipsed C_s (e-C_s) and (b) the C_{2v} transition structure of the CH_5^+ pseudorotation. In the e-C_s structure, the atoms which form the H2 moiety are marked in black to guide the eye.

The global minimum of the PES is of eclipsed C_s (e-C_s) symmetry, see Figure 3a, and may be visualized as a "CH_3 tripod", which includes protons H3, H4, and H5 in Figure 3a, to which the so-called "H_2 moiety", comprising protons H1 and H2 marked in black in Figure 3a, is attached. The H2 moiety is connected to the tripod via a three-center−two-electron bond, which is significantly weaker than the covalent C−H bonds within the tripod.[46] The global minimum can be realized via 5! = 120 permutations of the protons, which all differ from the standpoint of classical dynamics where each atom carries a unique label or number. Most importantly, these 120 minima are connected via extremely low barriers. The transition states between these minima are either a staggered C_s (s-C_s) structure, that is obtained by a rotation of the H_2 moiety by 30° in Figure 3a, or a C_{2v} structure shown in Figure 3b. The respective barrier heights are ∼0.1 kcal mol^{-1} (i.e., about 40 cm^{-1}, 0.004 eV or 50 K) and 0.8 kcal mol^{-1} (∼300 cm^{-1}, 0.03 eV or 400 K) higher in energy than the e-C_s global minimum.[55] The transition over the s-C_s barrier is termed internal (H_2 moiety) rotation, and the transition over the C_{2v} structure is referred to as pseudorotation. From the C_{2v} transition state, a new H_2 moiety can be formed by the atoms at positions H2 and H3. Concomitantly, the atom at position H1 occupies the position H3 in the newly formed tripod. Together with the complementary internal rotation, the pseudorotation permits each atom to travel between all five binding sites, which ultimately leads to what is often called "hydrogen scrambling".

Isotopic substitution complicates the situation even further. Replacing one proton by a deuteron and, thus, forming the CH_4D^+ isotopologue leads to five different isotopomers[54] which have to be considered, since the deuteron can as well reside at any of the five binding sites. The four protons can then be distributed in 4! = 24 ways on the remaining binding sites of each isotopomer. Thus, vibrational analysis of CH_4D^+ requires five conformations with 24 permutations each. In order to set the nomenclature, we classify the tripod atoms with respect to the symmetry plane σ_h; namely, H4 and H5 are referred to as "out-of-plane" atoms (or "sites") whereas H3 is referred to as the "in-plane" atom (site), see Figure 3a. Analogously, H2 in the H_2 moiety is considered to be the "inner" atom, since it is placed between H1 (being the "outer" atom) and the "in-plane" atom H3, all being coplanar with C. As the isotopomers of CH_4D^+ are sufficiently classified by specifying the deuteron position, we employ the same terminology to refer to them, e.g., inner isotopomer is the one with the deuteron in the inner position. Note that since the out-of-plane isotopomers differ only by their handedness, they yield the same spectral contributions.

Table 1. Internal Coordinates s_k and Weights a_k of the Metric d_{IC} as Defined in eq 20 for CH_4D^+[a]

k	s_k	a_k	$s_k(\mathbf{x}^0)$
1	stre(1,2)/Å	1.00	0.98
2	stre(3,4)/Å	1.00	1.79
3	stre(3,5)/Å	1.00	1.79
4	stre(4,5)/Å	1.00	1.89
5	stre(1,5)/Å	1.00	1.72
6	stre(1,4)/Å	1.00	1.72
7	stre(1,3)/Å	1.00	2.05
8	stre(2,3)/Å	1.00	1.44
9	stre(2,4)/Å	1.00	1.95
10	stre(2,5)/Å	1.00	1.95
11	tors(4,C,1,5)/rad	0.33	−2.12 (−121°)
12	tors(3,C,2,1)/rad	0.33	3.14 (180°)
13	tors(3,C,4,1)/rad	0.33	2.30 (132°)
14	tors(3,C,5,1)/rad	0.33	−2.30 (−132°)
15	tors(1,C,2,4)/rad	0.33	1.32 (76°)
16	tors(1,C,2,5)/rad	0.33	−1.32 (−76°)

[a] Reference values $s_k(\mathbf{x}^0)$ are given for the e-C_s structure with the deuteron at the outer position. The coordinates "stre(i,j)" refer to distances between sites i and j. Correspondingly, "tors(i,j,k,l)" refers to the set of dihedral angles. Note that the internal coordinates become dimensionless through the division by the given units (Å, rad).

The first task for our comprehensive vibrational analysis of the CH_4D^+ isotopologue using the extended GNC method is to split its IR spectrum into the contributions of the different isotopomers. For this purpose, we constructed the reference structures of the isotopologues starting from the e-C_s structure with the deuteron at the outer position. The remaining four reference structures were obtained by pair permutations of the deuteron coordinates with the coordinates of each proton. For the assignment of the trajectory frames to the isotopomers, we employed a metric d_{IC} as defined in eq 20 with the internal coordinates s_k and weights a_k given in Table 1.

Here, the structure is determined by all D−H and H−H distances across the molecule, which span a range from 0.98 Å to 2.05 Å in the minimum e-C_s structure. In addition, dihedral angles (k = 11−16) help to determine the topology of the deuteron and protons around the carbon. The a_k for the dihedral angles are chosen smaller than those of stretches since the dihedrals vary over a larger numerical range. With this choice, the contributions of both types of internal coordinates are balanced in the metric.

For a proper assignment of the trajectory frames to the isotopomers, we have to ensure that the reference structures are well separated in the space spanned by the internal coordinates, which define the metric d_{IC}. Figure 4 displays the relative occurrence of the distances between all 120 reference structures, including the 24 permutational references for each isotopologue. Most references are well separated by distances around 4.0, which correspond to multiple pair permutations of the deuteron and protons between the compared reference structures. Only a few reference structures are closer together at distances around 1.0 with a minimum distance of 0.9. These closer distances result from interchanging only the deuteron and one proton. These nearest neighbor distances now permit us to choose a proper value for the width of the Gaussians σ_c, which determines the width of the switching region between conformations in eq 3.

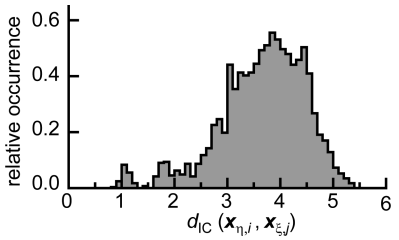

Figure 4. Relative occurrence of the distances between the 120 reference structures. Note that d_{IC} is dimensionless (c.f. Table 1).

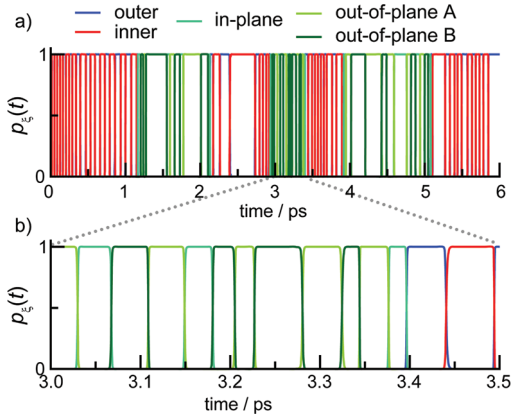

Figure 5. Probabilities $p_\xi(t)$ of the five different isotopomers ξ for one representative trajectory of isotopologue CH_4D^+ akin to Figure 1. The lower panel b zooms in on a selected time interval.

Here, we choose σ_c = 0.1, which separates nearest neighbors by nine standard deviations of the Gaussians. Note that choosing σ_c = 0.2 hardly changes the results, and therefore, the algorithm is robust with respect to this parameter. Furthermore, we employed the very same metric to discriminate between the 24 permutational symmetries of each isotopomer, which we will, therefore, not discuss separately.

The resulting probabilities of the isotopomers along the first 6 ps of a representative trajectory are shown in Figure 5a. Due to the employed metric, the probabilities switch fast and continuously between zero and one. As the close-up on the trajectory in Figure 5b shows, the molecule resides within an isotopomer on a time scale of only a few tens of femtoseconds, which is much shorter than the stability of the permutational states of the methyl group in isoprene, c.f. Figure 1. Furthermore, when we follow the sequence of the isotopomers in Figure 5b, we first see a switching between those three isotopomers where the tripod contains the deuteron. These transitions are caused by the internal rotation of the H_2 moiety with respect to the tripod. At about 3.4 ps, we observe a transition from the in-plane isotopomer to the inner isotopomer, which indicates a pseudorotation event that exchanges the deuteron from the tripod with the proton from the H_2 moiety, thus yielding a HD moiety. Thus, the probabilities obtained by our general projection procedure convincingly reflect the scrambling dynamics of the molecule.

In order to quantify its time scales, the time autocorrelation function of the probabilities $p_\xi(t)$ averaged over all five isotopomers ξ of CH_4D^+ is plotted in Figure 6. The autocorrelation function shows a clear biexponential decay behavior with the fitted time constants of about τ_1 = 58 fs and τ_2 = 630 fs. Comparing these time constants with the isotopomer dynamics

dx.doi.org/10.1021/ct2006665 |J. Chem. Theory Comput. 2012, 8, 224–234

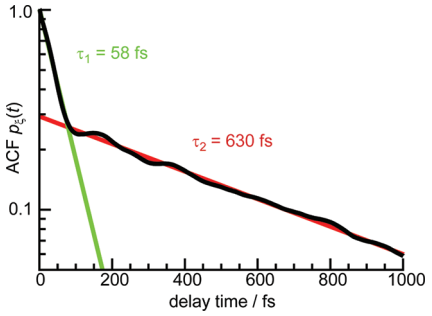

Figure 6. Log-scale time autocorrelation function of the probabilities $p_\xi(t)$ averaged over the isotopomers ξ (black curve). Colored lines display two fitted exponentials of time constants τ_1 = 58 fs and τ_2 = 630 fs.

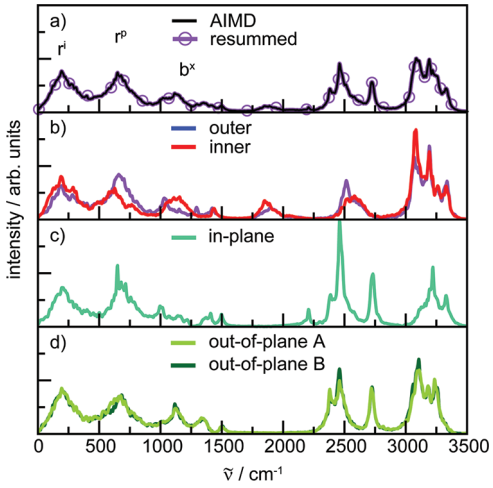

Figure 7. (a) Total IR spectra of CH_4D^+ as generated directly from AIMD simulations at T = 110 K (black line) and total spectrum obtained by resumming the five split spectra depicted in panels b−d weighted by the relative occurrences of the isotopomers (magenta circles). The other panels show the five isotopomer spectra obtained by the splitting procedure: the deuteron is contained in a HD moiety in b, whereas it is in the tripod in panels c and d, according to the color code as indicated.

in Figure 5b, one can deduce that τ_1 is associated with the frequent switching between isotopomers that mostly stems from internal rotations. The slower time constant τ_2 is identified by closer inspection of Figure 5a. Here, longer time intervals are dominated by either the inner and outer isotopomers or the in-plane and the two out-of-plane isotopomers. Thus, τ_2 describes the exchange of the deuteron between the tripod and H_2 moiety, which is the pseudorotation of the deuteron. Obviously, the present analysis yields classical lifetimes that are consistent with the classical approximation to the dynamics and the computation of the IR spectrum according to eq 2. It is noted in passing that quantum IR spectra have been computed and quantum lifetimes have been discussed in ref 28.

Having established the computation of the probabilities, we can now split the computed total IR absorption spectrum depicted in Figure 7a into the distinct spectral contributions of all five isotopomers that contribute, which are displayed in Figure 7b−d. Grossly speaking, three spectral regions can be identified in the total IR spectrum depicted in Figure 7a according to ref 24. The region up to roughly 750 cm^{-1} contains so-called rearrangement modes, i.e., the H_2/HD moiety rotation (r^i) and the pseudorotation motion (r^p), which have been discussed above. The modes between about 750 cm^{-1} and 1500 cm^{-1} can be attributed to bending motions b^x, and the modes above 1500 cm^{-1} are stretching modes of the H_2/HD moiety and the tripod. Particularly the latter are indicative of the different isotopomers because of the different force constants of the weaker bound H_2/HD moiety atoms (due to three-center−two-electron bonding) and the stronger bound tripod atoms, and due to the large frequency shift of the corresponding stretches upon H/D exchange. Thus, already a close inspection of the split spectra should allow for a tentative band assignment. The lowest stretching modes at around 1800 cm^{-1} are found for the outer and inner isotopomers in Figure 7b. Correspondingly, they should be associated with the C−D stretches of the deuteron in the HD moiety. The next modes at 2500 cm^{-1} therein should then represent the C−H stretch of the HD moiety protons, and hence, the modes located above 3000 cm^{-1} belong to the tripod protons. In contrast, the HD moiety C−D stretch at 1800 cm^{-1} is absent in panels c and d, which show the spectra corresponding to the deuterated tripod, but a new band at 2750 cm^{-1} appears. Therefore, one would, at first glance, assign this band to the C−D stretches of the tripod. However, if one considers the reduced mass ratios of a C−H and a C−D bond, a C−H stretch located at around 3200 cm^{-1} would be expected at around 2350 cm^{-1} upon H/D exchange, according to a most simplistic harmonic approximation estimate. Because a blue-shift of 400 cm^{-1} from the estimate seems too large, the band at 2750 cm^{-1} more likely stems from a H_2 moiety stretch. Then, the lower bands below 2500 cm^{-1} should be associated with the remaining C−H stretch of the H_2 moiety and the C−D stretch of the tripod. Going beyond such qualitative discussion by inspection requires one to use the extended GNC analysis scheme, which will provide full and unambiguous assignment of the total IR spectrum in terms of atomic motions as demonstrated below.

Comparing the split spectra of the two out-of-plane isotopomers, which only differ in their handedness and thus should yield strictly identical spectra, one can see that not only is the overall band structure identical, thus supporting the splitting procedure as such, but also that merely a few peak intensities deviate slightly. This implies both that the statistics sampled are sufficient, such that the peak heights of the other isotopomers can also be assumed to be converged, and that the splitting procedure works quantitatively. Note that although the statistics gathered for CH_4D^+ (corresponding to about 3 ns of AIMD trajectory) are 3 times those generated for isoprene, the convergence here is only comparable because the statistics are shared among the five isotopomers.

Finally, Figure 7a also includes the synthetic spectrum obtained by resumming the isotopomer-specific split spectra from panels b to d and weighting them by the relative occurrences of the respective isotopomers in the simulations ranging from 18.6% to 21.4%. These five weights are straightforwardly obtained from expectation values

$$\langle p_\xi \rangle = \frac{1}{N_{tr} t_{max}} \sum_{k=1}^{N_{tr}} \int_0^{t_{max}} \mathrm{d}t\, p_\xi^k(t), \text{ for } \xi = 1, \ldots, 5 \qquad (23)$$

over the N_{tr} = 297 trajectories and satisfy $\Sigma_\xi \langle p_\xi \rangle = 1$ by virtue of eq 3. The so-called resummed spectrum matches the total IR spectrum as directly generated by AIMD one to one, which confirms that the projection and splitting procedure introduced here works quantitatively; i.e., IR intensity is neither gained nor lost by the splitting.

dx.doi.org/10.1021/ct2006665 |J. Chem. Theory Comput. 2012, 8, 224–234

Although inspection of the split spectra seems to permit a tentative assignment of certain modes and yields qualitative insights into the structure of the bands as alluded to above, it is still insufficient for the unambiguous assignment even of the stretching modes. Thus, a quantitative assignment based on the GNC analysis is required to fully resolve the structure of these isotopomer spectra, as will be shown in the following. Here, we will limit the discussion of this analysis to the stretching modes, albeit the algorithm equally assigns the low frequency region; see e.g. ref 24 for the discussion of all modes of the CHD_4^+ isotopologue. Similar to the extended GNC analysis of isoprene, we have employed a RMSD fitting procedure to map the dynamics to the molecular frame of reference, the Eckhart frame, instead of using internal coordinates for this procedure.

Figure 8 displays the IR intensities of the individual stretching modes of the isotopomers as generated by the extended GNC procedure. Their labels have been adopted from the nomenclature introduced in ref 24 and are listed in Table 2 together with the contributions of the leading stretching components for the reader's convenience. In $CH_nD_{5-n}^+$, the stretching modes are localized on either of the two subunits, i.e., the tripod or H_2/HD moiety. Thus, the modes are labeled either "m" for the two H_2/HD moiety stretches or "t" for the three tripod stretches. The subscript H or D indicates whether a proton or a deuteron stretch dominates the mode, whereas the superscript refers to the symmetry of the mode: "a", antisymmetric; "s", symmetric; "i", in-plane; and "u", uncoupled.

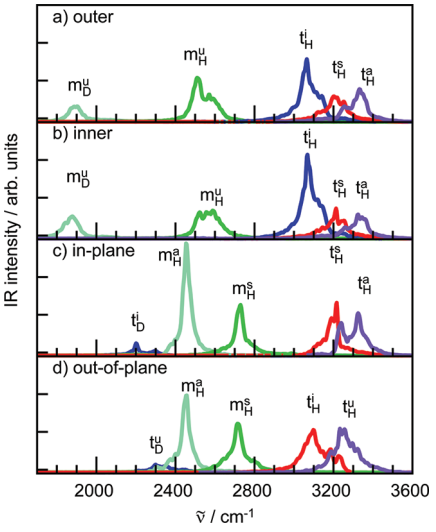

Figure 8. Spectral analysis of the stretching region for the four different isotopomers of the CH_4D^+ isotopologue (recall that the two out-of-plane isotopomers yield identical spectra). The disentangled spectral contributions are depicted in panel a for the outer isotopomer, b for inner, c for in-plane, and d for out-of-plane.

In pure CH_5^+ the tripod modes consist of the antisymmetric and symmetric out-of-plane stretches (t_H^a and t_H^s) and the in-plane stretch (t_H^i), and the H_2 moiety modes are the symmetric and antisymmetric stretches, m_H^s and m_H^a, respectively, see Supporting Information to ref 24 for a pictorial description of the corresponding atomic motions. Upon partial deuteration in the CH_4D^+ case, the symmetries of the CH_5^+ modes may be lowered, and formerly (anti)symmetric modes may now "uncouple" to yield only one dominant stretch. For instance, if the deuteron substitutes a proton in the H_2 moiety, the two former H_2 moiety modes, m_H^a and m_H^s, become m_D^u and m_H^u modes at low and high frequencies, respectively, see Figure 8a,b and Table 2. Notably, both the spectral contributions and the normal modes of the tripod stretches therein remain nearly the same and, in addition, are similar to CH_5^+ (data not shown); t_H^a has the highest frequency followed by t_H^s and t_H^i. Note also that the tripod and HD moiety stretches are not coupled, supporting the adequacy of the concept of two principle structural building blocks of protonated methane species: "H_2 moiety" and "CH_3 tripod".

When the deuteron resides in the tripod, the H_2 moiety modes are similar to those of bare CH_5^+; note that here the symmetric stretch m_H^s is higher in frequency than the antisymmetric stretch m_H^a, see Figure 8c,d and Table 2. However, the coefficients of t_D^i and t_D^u in Table 2 indicate that there is a noticeable coupling between the H_2 moiety stretches H1 and H2 to the tripod deuteron stretches D3 and D4, as expected from the similarity of their frequencies. When the deuteron occupies the in-plane position within the tripod, Figure 8c, the in-plane stretch, t_D^i, simply red-shifts, whereas the other two stretching modes t_H^s and t_H^a of the out-of-plane protons retain their symmetric and antisymmetric character similar to isotopomers characterized by a HD moiety or to bare CH_5^+. Finally, the symmetry is again lowered when the deuteron enters one of the out-of-plane sites, see Figure 8d. In this case the symmetric and antisymmetric tripod modes decouple into t_D^u and t_H^u modes, and t_D^u moves to the region at about 2300 cm^{-1} while t_H^u replaces t_H^a in its frequency range. Note that the stretch coefficients of the out-of-plane modes listed in Table 2 differ by at most 0.02 between the two out-of-plane isotopomers. Furthermore, these coefficients are almost identical if they are determined on a much smaller data set of only 60 trajectories, i.e., the data basis that was used in ref 24, which reflects the good convergence of the GNCs even if the statistics are limited. This discussion makes clear the point that a trustworthy assignment of the IR spectrum of an utmost floppy molecule such as CH_4D^+ not only requires one to split its total spectrum properly into the contributions stemming from the five isotopomers but also requires a full vibrational analysis of the latter by the extended GNC approach.

Thus, our detailed analysis confirms the capability of the extended GNC technique to treat highly nontrivial, truly multiple reference structure cases like the $CH_nD_{5-n}^+$ isotopologues of protonated methane.

Table 2. Dominant Stretch Components HX and DX to the Stretching Modes of the Isotopomers As Indicated

outer		inner		in-plane		out-of-plane	
m_D^u	0.87 D1 − 0.29 H2	m_D^u	0.90 D2 − 0.28 H1	t_D^i	0.93 D3 + 0.36 H1	t_D^u	0.91 D4 + 0.37 H2
m_H^u	0.93 H2 + 0.12 D1	m_H^u	0.91 H1 + 0.18 H3	m_H^a	0.77 H2 − 0.55 H1	m_H^a	0.68 H1 − 0.64 H2
t_H^i	0.97 H3 + 0.16 H5/H4	t_H^i	0.95 H3 + 0.20 H5/H4	m_H^s	0.68 H1 + 0.54 H2	m_H^s	0.62 H1 + 0.60 H2
t_H^s	0.70 H4 + 0.68 H5	t_H^s	0.70 H5 + 0.67 H4	t_H^s	0.71 H5 + 0.69 H4	t_H^i	0.95 H3 − 0.17 H1
t_H^a	0.70 H5 − 0.68 H4	t_H^a	0.70 H4 − 0.67 H5	t_H^a	0.70 H4 − 0.68 H5	t_H^u	0.97 H5 − 0.14 H3

5. CONCLUSIONS AND OUTLOOK

We have presented a significant extension of the generalized normal coordinate (GNC) analysis method introduced recently, which now enables the algorithm to handle multiple reference structures for a comprehensive vibrational analysis of molecular dynamics trajectories. These reference structures correspond to local minima of the PES, which either define chemically different conformations of the molecule or represent chemically equivalent structures that result from a permutation of atoms of the same species.

For each and every time frame of the trajectories, the probabilities of the molecule to occupy either of these reference structures are computed. Thereby, the trajectories are split into the contributions of the references. A time correlation formalism which weights the trajectory frames with the computed probabilities yields split spectra that represent the contributions of the underlying conformations to the total spectrum. Moreover, the scheme is capable to resolve the permutational symmetries, which permits one to compute GNCs for each conformation and, thereby, to assign the peaks in the conformation-specific IR spectra to atomic motion.

We have demonstrated the methodology in detail for two selected spectroscopic problems. The first task was to cleanly assign the vibrational modes of the methyl group of isoprene, which rotates nearly freely at ambient conditions. In paper I, it was shown that these modes are unsatisfactorily resolved when one assumes a quasi-rigid molecule as in the original GNC scheme. However, the extended GNC approach introduced here, employing three reference structures representing the permutational symmetry of the methyl group, fully resolves the vibrational bands, and the GNCs of these modes comply with the C_s symmetry of the molecule.

The second, much more demanding task was the comprehensive vibrational analysis of the CH_4D^+ isotopologue of protonated methane, the latter being widely considered to be among the most prominent and a truly challenging representatives of the class of floppy molecules. Here, 120 reference structures were needed to resolve the spectral contributions of the five isotopomers ("conformations"), each covering 24 permutational symmetries. The analysis showed that, although the dynamics of switching between the isotopomers occurs on a time scale as fast as 60 fs, the algorithm is able to split the computed total IR spectrum into physically reasonable contributions of the conformational states, i.e., into five isotopomer-specific IR spectra to be assigned separately using GNCs.

Having assigned the IR spectrum of a most challenging case such as partially deuterated protonated methane in terms of atomic motions, we expect the extended GNC method to be useful for the plethora of cases where a single equilibrium or reference structure is not sufficient to understand vibrational spectra. Last but not least, the very idea of using a dynamical projection scheme to split trajectories and thus spectra—a posteriori for analysis and not a priori for generation—can be readily transferred to other analysis schemes of response properties of molecular systems which are based on time correlation functions generated by (ab initio) molecular dynamics underlying theoretical spectroscopy in the Heisenberg picture.

AUTHOR INFORMATION

Corresponding Author

*E-mail: gerald.mathias@physik.uni-muenchen.de.

Present Addresses

[‡]Lehrstuhl für BioMolekulare Optik, Ludwig-Maximilians-Universität München, Oettingenstr. 67, 80538 München, Germany

[§]Quantum Dynamics Group, Institut für Physik, Universität Rostock, 18051 Rostock, Germany

[∥]PULSE Institute and Department of Chemistry, Stanford University, Stanford, California 94305, United States

[⊥]Chemical and Materials Science Division, Pacific Northwest National Laboratory, P.O. Box 999, Richland,Washington 99352, United States

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

We are grateful to T. Zelleke for his contributions to the implementation, to DFG (MA 1547/4 to D.M. and SFB 749/C4 to G.M.) and FCI (Chemiefonds—Stipendium to A.W.) for financial support, as well as to HLRB II (München), BOVILAB@RUB (Bochum), and RV-NRW (Dortmund) for computational resources.

REFERENCES

(1) Siebert, F. *Methods Enzymol.* **1995**, *246*, 501–526.
(2) Vogel, R.; Siebert, F. *Curr. Opin. Chem. Biol.* **2000**, *4*, 518–523.
(3) Bieske, E. J.; Dopfer, O. *Chem. Rev.* **2000**, *100*, 3963–3998.
(4) Barth, A.; Zscherp, C. *Q. Rev. Biophys.* **2002**, *35*, 369–430.
(5) Barth, A. *Biochim. Biophys. Acta* **2007**, *1767*, 1073–1101.
(6) Polfer, N. C.; Oomens, J. *Mass Spectrom. Rev.* **2009**, *28*, 468–494.
(7) Marx, D.; Hutter, J. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*; Cambridge University Press: Cambridge, U. K., 2009; pp 11−75.
(8) Silvestrelli, P. L.; Bernasconi, M.; Parrinello, M. *Chem. Phys. Lett.* **1997**, *277*, 478–482.
(9) Nonella, M.; Mathias, G.; Eichinger, M.; Tavan, P. *J. Phys. Chem. B* **2003**, *107*, 316–322.
(10) Vogel, R.; Siebert, F.; Mathias, G.; Tavan, P.; Fan, G.; Sheves, M. *Biochemistry* **2003**, *42*, 9863–9874.
(11) Gaigeot, M.-P.; Sprik, M. *J. Phys. Chem. B* **2003**, *107*, 10344–10358.
(12) Rousseau, R.; Kleinschmidt, V.; Schmitt, U. W.; Marx, D. *Angew. Chem., Int. Ed.* **2004**, *43*, 4804–4807.
(13) Gaigeot, M.-P.; Vuilleumier, R.; Sprik, M.; Borgis, D. *J. Chem. Theory Comput.* **2005**, *1*, 772–789.
(14) Iftimie, R.; Tuckerman, M. E. *J. Chem. Phys.* **2005**, *122*, 214508.
(15) Asvany, O.; Kumar, P. P.; Redlich, B.; Hegemann, I.; Schlemmer, S.; Marx, D. *Science* **2005**, *309*, 1219–1222.
(16) Martinez, M.; Gaigeot, M.-P.; Borgis, D.; Vuilleumier, R. *J. Chem. Phys.* **2006**, *125*, 144106/14.
(17) Kumar, P.; Marx, D. *Phys. Chem. Chem. Phys.* **2006**, *8*, 573–586.
(18) Gaigeot, M.-P.; Martinez, M.; Vuilleumier, R. *Mol. Phys.* **2007**, *105*, 2857–2878.
(19) Mathias, G.; Marx, D. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 6980–6985.
(20) Masia, M.; Forbert, H.; Marx, D. *J. Phys. Chem. A* **2007**, *111*, 12181–12191.
(21) Baer, M.; Mathias, G.; Kuo, I.-F. W.; Tobias, D. J.; Mundy, C. J.; Marx, D. *ChemPhysChem* **2008**, *9*, 2703–2707.
(22) Cimas, A.; Vaden, T. D.; de Boer, T. S. J. A.; Snoek, L. C.; Gaigeot, M. P. *J. Chem. Theory Comput.* **2009**, *5*, 1068–1078.
(23) Thomas, V.; Iftimie, R. *J. Phys. Chem. B* **2009**, *113*, 4152–4160.
(24) Ivanov, S. D.; Asvany, O.; Witt, A.; Hugo, E.; Mathias, G.; Redlich, B.; Marx, D.; Schlemmer, S. *Nat. Chem.* **2010**, *2*, 298–302.
(25) Heyden, M.; Sun, J.; Funkner, S.; Mathias, G.; Forbert, H.; Havenith, M.; Marx, D. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 12068–12073.

(26) Zhang, C.; Donadio, D.; Galli, G. *J. Phys. Chem. Lett.* **2010**, *1*, 1398–1402.
(27) Baer, M.; Marx, D.; Mathias, G. *Angew. Chem., Int. Ed.* **2010**, *49*, 7346–7349.
(28) Witt, A.; Ivanov, S. D.; Mathias, G.; Marx, D. *J. Phys. Chem. Lett.* **2011**, *2*, 1377–1381.
(29) Huang, X. C.; Carter, S.; Bowman, J. *J. Chem. Phys.* **2003**, *118*, 5431–5441.
(30) Huang, X. C.; Braams, B. J.; Bowman, J. M. *J. Chem. Phys.* **2005**, *122*, 044308.
(31) Bowman, J. M.; Carrington, T.; Meyer, H.-D. *Mol. Phys.* **2008**, *106*, 2145–2182.
(32) Jin, Z.; Braams, B.; Bowman, J. *J. Phys. Chem. A* **2006**, *110*, 1569–1574.
(33) Park, M.; Shin, I.; Singh, N. J.; Kim, K. S. *J. Phys. Chem. A* **2007**, *111*, 10692–10702.
(34) Kaledin, M.; Kaledin, A. L.; Bowman, J. M.; Ding, J.; Jordan, K. D. *J. Phys. Chem. A* **2009**, *113*, 7671–7677.
(35) Baer, M.; Marx, D.; Mathias, G. *ChemPhysChem* **2011**, *12*, 1906–1915.
(36) Gordon, R. G. *Adv. Magn. Reson.* **1968**, *3*, 1–42.
(37) Berens, P. H.; Wilson, K. R. *J. Chem. Phys.* **1981**, *74*, 4872–4882.
(38) Putrino, A.; Parrinello, M. *Phys. Rev. Lett.* **2002**, *88*, 176401.
(39) Wheeler, R. A.; Dong, H.; Boesch, S. E. *ChemPhysChem* **2003**, *3*, 382–384.
(40) Wheeler, R. A.; Dong, H. *ChemPhysChem* **2003**, *4*, 1227–1230.
(41) Schmitz, M.; Tavan, P. *J. Chem. Phys.* **2004**, *121*, 12233–12246.
(42) Schmitz, M.; Tavan, P. *J. Chem. Phys.* **2004**, *121*, 12247–12258.
(43) Agostini, F.; Vuilleumier, R.; Ciccotti, G. *J. Chem. Phys.* **2011**, *134*, 084302.
(44) Mathias, G.; Baer, M. *J. Chem. Theory Comput.* **2011**, *7*, 2028–2039.
(45) Towns, T. G.; Carreira, L. A.; Irwind, R. M. *J. Raman Spectrosc.* **1981**, *11*, 487–492.
(46) Marx, D.; Parrinello, M. *Nature* **1995**, *375*, 216–218.
(47) Marx, D.; Parrinello, M. *Science* **1999**, *284*, 59–61.
(48) McCoy, A. B.; Braams, B.; Brown, A.; Huang, X.; Jin, Z.; Bowman, J. *J. Phys. Chem. A* **2004**, *108*, 4991–4994.
(49) Huang, X.; Johnson, L.; Bowman, J.; McCoy, A. *J. Am. Chem. Soc.* **2006**, *128*, 3478–3479.
(50) Huang, X.; McCoy, A. B.; Bowman, J. M.; Johnson, L. M.; Savage, C.; Dong, F.; Nesbitt, D. J. *Science* **2006**, *311*, 60–63.
(51) Johnson, L.; McCoy, A. *J. Phys. Chem. A* **2006**, *110*, 8213–8220.
(52) Wanga, X.-G.; Carrington, T., Jr. *J. Chem. Phys.* **2008**, *129*, 234102.
(53) Hinkle, C. E.; Petit, A. S.; McCoy, A. B. *J. Mol. Spectrosc.* **2011**, *268*, 189–198.
(54) "Isotopologue" (doi: 10.1351/goldbook.I03351) and .isotopomer. (doi: 10.1351/goldbook.I03352) are used according to IUPAC, see http://goldbook.iupac.org (created by Nic, M.; Jirat, J.; Kosata, B.; updates compiled by Jenkins, A.; accessed 11/2011).
(55) Muller, H.; Kutzelnigg, W.; Noga, J.; Klopper, W. *J. Chem. Phys.* **1997**, *106*, 1863–1869.
(56) Ramírez, R.; Lopez-Ciudad, T.; Kumar, P.; Marx, D. *J. Chem. Phys.* **2004**, *121*, 3973–3983.
(57) Wilson, E.; Decius, J. C.; Cross, P. C. *Molecular Vibrations*; McGraw Hill: New York, 1955; pp 54−76.
(58) Cremer, D.; Pople, J. *J. Am. Chem. Soc.* **1975**, *97*, 1354–1358.
(59) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comput. Phys. Commun.* **2005**, *167*, 103–128.
(60) Hutter, J.; Alavi, A.; Deutsch, T.; Bernasconi, M.; Goedecker, S.; Marx, D.; Tuckerman, M.; Parrinello, M. CPMD: Car−Parinello Molecular Dynamics, version 3.10; IBM Corp: Endicott, NY, 1990; MPI für Festkörperforschung Stuttgart: Stuttgart, Germany, 1997. www.cpmd.org (accessed Nov. 2011).
(61) Hsu, S.; Kemp, M.; Pochan, J.; Benson, R.; Flygare, W. *J. Chem. Phys.* **1969**, *50*, 1482–1483.