

## Vibrational modes in partially optimized molecular systems

A. Ghysels, D. Van Neck, V. Van Speybroeck, T. Verstraelen, and M. Waroquier

Citation: *The Journal of Chemical Physics* **126**, 224102 (2007); doi: 10.1063/1.2737444

View online: <http://dx.doi.org/10.1063/1.2737444>

View Table of Contents: <http://aip.scitation.org/toc/jcp/126/22>

Published by the *American Institute of Physics*

---

### Articles you may be interested in

*Cartesian formulation of the mobile block Hessian approach to vibrational analysis in partially optimized systems*  
**127**, 164108164108 (2007); 10.1063/1.2789429

---

COMPLETELY

REDESIGNED!



PHYSICS  
TODAY

*Physics Today* Buyer's Guide  
Search with a purpose.

# Vibrational modes in partially optimized molecular systems

A. Ghysels, D. Van Neck, V. Van Speybroeck, T. Verstraelen, and M. Waroquier<sup>a)</sup>*Center for Molecular Modeling, Laboratory of Theoretical Physics, Ghent University, Proeftuinstraat 86, B-9000 Gent, Belgium*

(Received 27 February 2007; accepted 12 April 2007; published online 12 June 2007)

In this paper the authors develop a method to accurately calculate localized vibrational modes for partially optimized molecular structures or for structures containing link atoms. The method avoids artificially introduced imaginary frequencies and keeps track of the invariance under global translations and rotations. Only a subblock of the Hessian matrix has to be constructed and diagonalized, leading to a serious reduction of the computational time for the frequency analysis. The mobile block Hessian approach (MBH) proposed in this work can be regarded as an extension of the partial Hessian vibrational analysis approach proposed by Head [Int. J. Quantum Chem. **65**, 827 (1997)]. Instead of giving the nonoptimized region of the system an infinite mass, it is allowed to move as a rigid body with respect to the optimized region of the system. The MBH approach is then extended to the case where several parts of the molecule can move as independent multiple rigid blocks in combination with single atoms. The merits of both models are extensively tested on ethanol and di-*n*-octyl-ether. © 2007 American Institute of Physics.

[DOI: [10.1063/1.2737444](https://doi.org/10.1063/1.2737444)]

## I. INTRODUCTION

The applications of molecular modeling nowadays focus more and more on extended systems, in which numerous atoms are involved. Examples are polymer chains, supramolecular assemblies, systems embedded in a solvent, or (macro)molecules adsorbed within porous materials. For the description of the electronic part of the system, one usually resorts to a hybrid model, in which the chemically active part where bonds may be formed or broken is described at a quantum mechanical level, whereas the outer region is described at a lower molecular mechanics level. The previous models are often referred to as quantum mechanical/molecular mechanical methods.<sup>1–4</sup> The basic idea originated from the observation that normally only a portion of the atoms in a reaction is directly involved in bond breaking and forming or in changing of bond order. These methods can be used either in a cluster or a periodic based approach.

In many cases, only part of the system is optimized to restrict the computational cost or to prevent unphysical deformations of the border of periodic systems. Good examples are reactions occurring in porous materials, where the border of the system is kept fixed during the geometry optimization to prevent unphysical deformations due to the neglect of the whole periodic structure. Another example is the simulation of defects embedded in a crystal lattice. In such cases, it is common practice to cut out a cluster around the defect and to keep the cluster border atoms fixed during the geometry optimization.<sup>5</sup> After such a constrained geometry optimization, residual forces remain on the fixed border atoms and the partially optimized structure corresponds to a nonequilibrium state. The usual normal mode analysis (NMA) equations can

be applied to this nonequilibrium configuration using the full Hessian of all atoms in the structure, i.e., the matrix of second derivatives of the potential energy surface with respect to all nuclear coordinates. Such a procedure, however, has some serious defects. The Hessian will have only three eigenvalues equal to zero instead of six, implying that the rotational invariance of the potential surface is not manifest. In addition, spurious imaginary normal frequencies appear, suggesting that the partially optimized system resides in a transition state, even when this is obviously not the case.

Whereas for the determination of the energy the partitioning into chemically active and passive areas is common practice, it is far less applied for the determination of the Hessian of such extended systems. The calculation of the Hessian is one of the most expensive steps in the calculation of free energies, so a partitioning scheme for the Hessian would also seriously reduce the computational cost.

Within this respect, the work of Head and co-workers is especially interesting.<sup>6,7</sup> They introduced a strategy by which only the frequencies of part of a chemical system are computed. In 2002, Li and Jensen introduced the name partial Hessian vibrational analysis (PHVA) and extended the method for the calculation of vibrational enthalpy and entropy changes for chemical reactions.<sup>8</sup> Recently, Besley and Metcalf applied the partial Hessian approximation to calculate the amide I band of polypeptides and proteins.<sup>9</sup> Within this methodology, however, the normal modes are calculated for the system with the fixed atoms frozen at their reference positions as if they were given an infinite mass and only the relaxed atoms can participate in the small amplitude vibrations. Head and co-workers have also developed a more sophisticated partial Hessian method,<sup>10–12</sup> where frequencies are corrected in lowest order perturbation theory for the cou-

<sup>a)</sup>Author to whom correspondence should be addressed. Electronic mail: [michel.waroquier@ugent.be](mailto:michel.waroquier@ugent.be)

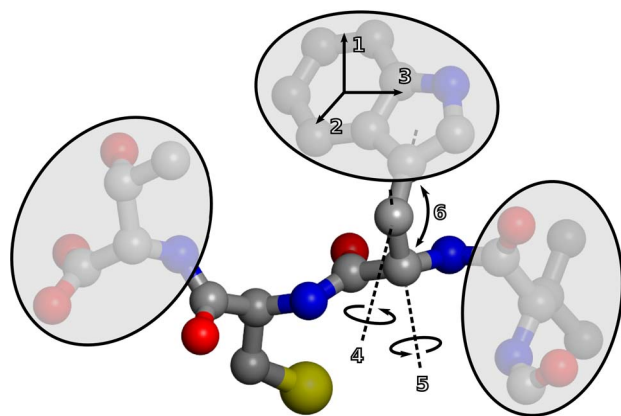


FIG. 1. Schematic representation of the basic idea behind the MBH method. The shaded blocks symbolize the parts of the molecule of which the internal geometry is kept fixed during the partial geometry optimization. In the MBH approach, they are described as rigid bodies with six degrees of freedom (translations and rotations).

pling between PHVA modes and modes in the fixed part. This method has been very useful for the calculation of localized vibrations of adsorbates on surfaces.

In this paper we propose an extended version of the PHVA method, in which the atoms that were kept fixed during the optimization can participate in small amplitude vibrations of the system, with the restriction that only coherent movements as a single block are allowed. The block of the fixed atoms can rotate and translate as a rigid body, while the internal geometry of the block is kept fixed. This model is referred hereafter to as the mobile block Hessian (MBH) approach. A schematic representation of the basic idea behind the MBH method is given in Fig. 1.

It can be expected that both methods would give similar results in those cases where the nonoptimized part of the system is quite rigid. The example of a defect in a crystal lattice is a case in point. For systems in which the molecular environment is more flexible, such as reactions occurring in solvents or polymer chains, the motions of the surroundings cannot *a priori* be neglected. It needs to be investigated to what extent the two methods coincide for such applications and whether such a partial optimization can give an accurate description of the localized modes in the optimized region. Localized modes are characterized by the fact that during the small amplitude vibrations, changes in the geometry only occur in a restricted region of the molecular system. In order to investigate this question, one should be able to construct the normal modes and frequencies of the partially optimized system in a rigorous way and compare them with frequencies of the fully optimized system.

The computational cost of the PHVA and MBH methods is similar, and the NMA equations can be rephrased in terms of a Hessian of reduced dimension in both cases. This leads to a significantly reduced load compared to a full Hessian frequency calculation. Related with a less expensive treatment of the Hessian is the work of Lin *et al.*<sup>13</sup> They have proposed efficient methods for calculating the Hessian in the optimization procedure of the multiconfigurational molecular mechanics method. The partial Hessian vibrational analysis and also the mobile block Hessian approach could be useful

in combination with such methods as they could additionally speed up the computationally expensive task of determining the Hessian for reactions on extended systems.

In Sec. II the problem of determining normal modes in nonequilibrium systems is discussed. The partial Hessian vibrational analysis is revised, and a detailed theoretical derivation is given of the mobile block Hessian approach, which we propose as an extension of the PHVA method for extended systems which are quite flexible. In Sec. III the two methods are compared for the example of the ethanol molecule. This very simple molecule has been chosen as a test case, as it enables one to address various phenomenological issues of the PHVA and MBH methods. The MBH approach can be easily extended to the case where several parts of the molecule can move as independent rigid bodies in combination with single atoms. The framework of this “multiple” mobile block Hessian approach is worked out in Sec. IV and is applied to di-*n*-octyl-ether in Sec. V. Finally, the results are summarized in Sec. VI, and future applications of the MBH model are discussed.

## II. BACKGROUND AND THEORETICAL DEVELOPMENT

### A. Normal modes in nonequilibrium configurations

Consider a molecule with  $N$  masses  $m_A$  ( $A=1, \dots, N$ ), the positions of which are described by Cartesian coordinates  $\mathbf{r}_A \equiv \{r_{A\mu}\}_{\mu=x,y,z}$ , with respect to a space-fixed frame. The energy of the system reads

$$E = \frac{1}{2} \sum_{A\mu} m_A \dot{r}_{A\mu}^2 + V(\{\mathbf{r}_A\}), \quad (1)$$

where  $\dot{r}_{A\mu}$  is a time derivative and  $V$  is the potential energy. Expanding  $V$  around a reference configuration  $\{r_{A\mu}^0\}$  gives

$$V(\{\mathbf{r}_A\}) = V_0 + \sum_{A\mu} \left( \frac{\partial V}{\partial r_{A\mu}} \right)_0 \Delta_{A\mu} + \frac{1}{2} \sum_{A\mu, B\nu} \left( \frac{\partial^2 V}{\partial r_{A\mu} \partial r_{B\nu}} \right)_0 \Delta_{A\mu} \Delta_{B\nu} + \dots \quad (2)$$

in terms of the displacement coordinates  $\Delta_{A\mu} = r_{A\mu} - r_{A\mu}^0$ . By collecting the coordinates in one vector  $x_i = r_{A\mu}$ , with  $i \equiv A\mu = 1, \dots, 3N$ , one can expand the energy in matrix form as

$$E - V_0 = \frac{1}{2} \dot{\Delta}^T M \dot{\Delta} + \Delta^T G + \frac{1}{2} \Delta^T H \Delta \quad (3)$$

up to second order in the displacements  $\Delta_i = x_i - x_i^0$ . In Eq. (3), the mass matrix  $M$  is a diagonal matrix containing the masses,  $G$  is the gradient vector defined by the gradients of the potential energy at the reference point, and  $H$  is the Hessian or second-derivative matrix in the reference point. The NMA equation determining the eigenmodes  $v$  and normal frequencies  $\lambda^{1/2}$  reads

$$Hv = \lambda Mv, \quad (4)$$

representing a generalized eigenvalue problem.

Solution schemes of these equations are implemented in various standard *ab initio* molecular modeling packages, but if the whole molecular system is not in its equilibrium state, the reference configuration is not gradient-free and some of the resulting frequencies are completely unphysical. An additional problem is that a discrimination of the unphysical frequencies from the physical values is not obvious.

We want to make two comments in the situation of non-equilibrium configurations (gradient vector  $G \neq 0$ ). First, one can show that the eigenvalues of Eq. (4) are coordinate-dependent: a second-order expansion of the potential energy expressed in curvilinear coordinates, which are nonlinearly related to the Cartesian coordinates, leads to different normal mode frequencies. This phenomenon is “well known” and is related to the difference between ordinary and covariant derivatives in a nontrivial metric space.<sup>14</sup> Its importance and unpleasant consequences have recently been emphasized by several authors.<sup>15,16</sup>

A second comment deals with the invariance of the potential energy surface under overall rotations. When  $G \neq 0$ , the Cartesian Hessian will only generate three zero eigenvalues (those related to the global translation). The three eigenvalues associated with the global rotation are different from zero,<sup>17</sup> and the absence of these Goldstone modes<sup>18</sup> is due to the use of the second-order expansion in rectilinear Cartesian coordinates, which breaks the rotational symmetry of  $V$ . This defect is simply cured by taking coordinates that respect the symmetries of  $V$ , i.e., internal coordinates.

With any choice of  $3N-6$  internal coordinates  $\{\theta_I\}$ , and a body frame whose origin lies at the center of mass  $\mathbf{r}_{\text{c.m.}}$ , the energy in Eq. (1) can be rewritten in a standard way<sup>19-21</sup> as the sum of the potential energy  $W(\{\theta_I\})$  expressed in internal coordinates and kinetic energy terms arising from the center-of-mass motion, global rotation, Coriolis coupling, and internal motions:

$$E = \frac{1}{2} \mathcal{M} \dot{\mathbf{r}}_{\text{c.m.}}^2 + \frac{1}{2} \bar{\omega} \bar{\mathbf{I}} \omega + \omega \sum_I \mathbf{A}_I \dot{\theta}_I + \frac{1}{2} \sum_{IJ} B_{IJ} \dot{\theta}_I \dot{\theta}_J + W(\{\theta_I\}). \quad (5)$$

Here  $\mathcal{M}$  is the total mass and  $\omega$  the angular velocity vector of the body frame. The inertial tensor  $\bar{\mathbf{I}}$ , the Coriolis coupling  $\mathbf{A}_I$  between the body-frame rotation and internal velocity  $\dot{\theta}_I$ , and the  $B_{IJ}$  matrix are all functions of the internal coordinates.

The potential energy  $W(\{\theta_I\})$ , expanded to second order about a reference geometry  $\{\theta_I^0\}$ , reads

$$W(\{\theta_I\}) = V_0 + \sum_I \left( \frac{\partial W}{\partial \theta_I} \right)_0 \Delta_I + \frac{1}{2} \sum_{IJ} \left( \frac{\partial^2 W}{\partial \theta_I \partial \theta_J} \right)_0 \Delta_I \Delta_J. \quad (6)$$

The set of new coordinates now consists of the center-of-mass coordinates  $\mathbf{r}_{\text{c.m.}}$ , three angles to specify the orientation of the body frame (e.g., Euler angles), and the  $3N-6$  internal coordinates  $\theta_I$ . The NMA equations using internal coordinates show a block structure:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & H^{(ii)} \end{pmatrix} \begin{pmatrix} \mathbf{v}^{(c)} \\ \mathbf{v}^{(r)} \\ \mathbf{v}^{(i)} \end{pmatrix} = \lambda \begin{pmatrix} M^{(cc)} & 0 & 0 \\ 0 & M^{(rr)} & M^{(ri)} \\ 0 & (M^{(ri)})^T & M^{(ii)} \end{pmatrix} \begin{pmatrix} \mathbf{v}^{(c)} \\ \mathbf{v}^{(r)} \\ \mathbf{v}^{(i)} \end{pmatrix}, \quad (7)$$

where  $\mathbf{v}^{(c)}$ ,  $\mathbf{v}^{(r)}$ , and  $\mathbf{v}^{(i)}$  have dimensions 3, 3, and  $3N-6$ , respectively, and the mass matrix entries read  $M^{(cc)} = \mathcal{M} \delta_{\mu\nu}$ ,  $M^{(rr)} = I_{\mu\nu}(\{\theta_I^0\})$ ,  $M^{(ri)} = A_{\mu I}(\{\theta_I^0\})$ , and  $M^{(ii)} = B_{IJ}(\{\theta_I^0\})$ . If the gradient is nonzero, the NMA equations in Eq. (7) will give different eigenvalues from those of Eq. (4). In particular, the eigenvalue equations (7) always generate six zero eigenvalues due to its construction, as the Hessian has vanishing matrix elements in the  $(c)$  and  $(r)$  subspaces, while in Eq. (4) the presence of six zero eigenvalues is not ensured.

In a completely general situation it is, in principle, not possible, due to the coordinate dependence, to define meaningful normal modes in a nonequilibrium point. However, the situations of interest for the present paper are not completely general but arise from physical considerations.

We consider cases where the reference point is obtained by optimizing the energy with respect to a *subset* of coordinates, keeping the remainder fixed during the optimization. The fixed coordinates correspond, e.g., to the geometry of a part of the molecule that is expected to influence the localized mode of interest only slightly. The system corresponding to the subset of coordinates that have been optimized is actually in equilibrium, even though the complete gradient is nonzero.

A detailed theoretical analysis of the coordinate (in)dependence and symmetry properties of the NMA equations in nonequilibrium systems, as well as the link between descriptions using Cartesian and internal coordinates, will be given in a separate publication. In this paper we focus on the comparison of two practical methods for treating such cases.

## B. Partial Hessian vibrational analysis

One assumes a certain reference structure  $\{\mathbf{r}_{A_F}^0\}$  for part of the molecule consisting of  $N_F$  atoms (labeled  $A_F, B_F, \dots$ ). Keeping  $\{\mathbf{r}_{A_F}^0\}$  fixed in space, the positions of the remaining  $N_E = N - N_F$  atoms (labeled  $A_E, B_E, \dots$ ) are optimized, resulting in an “equilibrium” configuration  $\{\mathbf{r}_{A_E}^0\}$ . Obviously, the full gradient is nonzero, since  $(\partial V / \partial r_{A_E \mu})_0 = 0$ , but  $(\partial V / \partial r_{A_F \mu})_0 \neq 0$ . Solutions of the NMA equations taking into account the full Hessian and mass matrix will therefore lead to unphysical results.

In a first model, known as partial Hessian vibrational analysis, one assumes that the fixed atoms  $A_F$  do not participate in the small amplitude vibrations. Hence their displacements  $\Delta_{A_F \mu}$  and velocities  $\dot{\Delta}_{A_F \mu}$  are set to zero. This is consistent with a situation in which infinite masses  $m_{A_F}$  are associated with the fixed atoms. With this assumption, the second-order energy of Eq. (3) reduces to



$$\begin{aligned}
E - V_0 &= \frac{1}{2} \sum_{A_E \mu} m_{A_E} \dot{\Delta}_{A_E \mu}^2 \\
&+ \frac{1}{2} \sum_{A_E \mu, B_E \nu} \left( \frac{\partial^2 V}{\partial \Delta_{A_E \mu} \partial \Delta_{B_E \nu}} \right)_0 \Delta_{A_E \mu} \Delta_{B_E \nu} + \dots \\
&= \frac{1}{2} \dot{\Delta}'^T M' \dot{\Delta}' + \frac{1}{2} \Delta'^T H' \Delta', \quad (8)
\end{aligned}$$

where  $\Delta'$ ,  $\dot{\Delta}'$ ,  $M'$ , and  $H'$  are the displacements, velocities, mass matrix, and Hessian restricted to the reduced  $3N_E$ -dimensional subsystem of the nonfixed atoms with mass  $m_{A_E}$ . Note that the reduced gradient ( $G'=0$ ) vanishes in this subspace, so the  $3N_E$  system is in equilibrium at the reference configuration and provides coordinate-independent normal modes.

In practice, one simply needs to disregard the rows and columns related to the coordinates of the fixed atoms in the full (Cartesian) Hessian and mass matrix. The NMA equation reduces to a  $3N_E \times 3N_E$  generalized eigenvalue problem:

$$\sum_{B_E \nu} \left( \frac{\partial^2 V}{\partial r_{A_E \mu} \partial r_{B_E \nu}} \right)_0 v_{B_E \nu} = \lambda m_{A_E} v_{A_E \mu}. \quad (9)$$

The reduced mass matrix  $M'$  is positive definite, and the presence of negative solutions  $\lambda$  in Eq. (9) gives indication that the reduced system resides in a transition state rather than a minimum. The number of zero eigenvalues depends on the remaining symmetry of the reduced system. Usually no zeros will occur, since the fixed atoms act as an external field breaking translational and rotational invariance for the nonfixed atoms.

### C. The mobile block Hessian approach

In the MBH approach, the fixed atoms are allowed to participate in the small amplitude vibrations by moving as a rigid body (block). This is a physically different situation from the previous approach, since one now takes into account the finite masses of the fixed atoms  $\{m_{A_F}\}$ .

This approach is easily implemented by a suitable choice of the  $3N-6$  internal coordinates  $\{\theta_I\}$ , which is always possible. In the Z-matrix formalism, for instance, one can determine the first  $N_F$  atoms by  $3N_F-6$  Z coordinates (distances, angles, and dihedral angles), and the next  $N_E$  atoms can consecutively be described by  $3N_E$  Z coordinates.

The  $3N_F-6$  internal coordinates, describing the geometry of the fixed atoms  $A_F$ , are labeled  $\{\theta_{I_F}\}$ , and the remaining  $3N_E$  internal coordinates are labeled  $\{\theta_{I_E}\}$ . The imposed reference structure  $\{\mathbf{r}_{A_F}^0\}$  of the fixed atoms then determines the values of the  $3N_F-6$  internal coordinates,  $\{\theta_{I_F}^0\}$ . Optimizing the energy with fixed  $\{\theta_{I_F}^0\}$  and varying  $\{\theta_{I_E}\}$  give the reference structure  $\{\theta_{I_E}^0\}$ . Again, the full gradient is nonzero, since  $(\partial W / \partial \theta_{I_E})_0 = 0$  but  $(\partial W / \partial \theta_{I_F})_0 \neq 0$ . As a result, the NMA equation (7) with the full Hessian will not give correct frequencies (though one does have six zero eigenvalues).

Allowing the fixed block to move as a rigid body, keeping its internal geometry, can be imposed by putting the displacements  $\Delta_{I_F}$  and corresponding velocities  $\dot{\Delta}_{I_F}$  equal to

zero in the second-order expansion of Eq. (5). The part of the molecule described by the  $3N_E$  internal coordinates  $\{\theta_{I_E}\}$  has been relaxed during the geometry optimization and hence is in equilibrium. The gradient term in Eq. (5) thereby vanishes.

The corresponding NMA equations are obtained by omitting the rows and columns related to the  $\{\theta_{I_F}\}$  variables from the Hessian and mass matrix in Eq. (7). The resulting reduced eigenvalue problem of dimension  $3N_E+6$  still has six zero eigenvalues corresponding to overall translation and rotation. These can be decoupled in the usual way<sup>19-21</sup> by a congruent transformation,  $\tilde{\mathbf{v}}^{(c)} = \mathbf{v}^{(c)}$ ,  $\tilde{\mathbf{v}}^{(i)} = \mathbf{v}^{(i)}$ , and  $\tilde{\mathbf{v}}^{(r)} = \mathbf{v}^{(r)} + (M^{(rr)})^{-1} M'^{(ri)} \mathbf{v}^{(i)}$ , yielding the final  $3N_E$ -dimensional NMA equation,

$$H'^{(ii)} \tilde{\mathbf{v}}^{(i)} = \lambda [M'^{(ii)} - (M'^{(ri)})^T (M^{(rr)})^{-1} M'^{(ri)}] \tilde{\mathbf{v}}^{(i)}. \quad (10)$$

The primed matrices do not contain the components related to  $\{\theta_{I_F}\}$ . The transformed mass matrix in the right hand side of Eq. (10) is positive definite and the presence of negative eigenvalues  $\lambda$  will now unambiguously indicate that the reduced system is in a transition state.

### D. Discussion: PHVA versus MBH

Both PHVA and MBH models offer a vibrational analysis in which only a subblock of the Hessian matrix is diagonalized to produce vibrational frequencies for partially optimized systems. In both descriptions, the molecular system is composed of a rigid body with a number  $N_F$  of fixed atoms and  $N_E$  atoms that are free to relax in a partial geometry optimization. Conceptually, the difference between PHVA and MBH lies in the use of near-infinite masses in the former approach for the atoms in the rigid body, while in the MBH model, reduced masses are used.

A more transparent and fundamental comparative study between the PHVA and MBH models can be made when using a specific choice of internal coordinates: for the set of  $3N_E$  internal coordinates  $\{\theta_{I_E}\}$  describing the geometry of the nonfixed atoms, one can take the Cartesian coordinates  $\{\mathbf{r}_{A_E}^{\text{RB}}\}$  in a frame attached to the rigid block (RB) (e.g., the frame constructed by the principal axes of the fixed atoms). The potential energy then becomes

$$W(\{\theta_{I_F}, \theta_{I_E}\}) \equiv V(\{\mathbf{r}_{A_F}^{\text{RB}}, \mathbf{r}_{A_E}^{\text{RB}}\}). \quad (11)$$

Since the  $\{\mathbf{r}_{A_F}^{\text{RB}}\}$  are constant in the fixed geometry of the block, one has

$$\left( \frac{\partial^2 W}{\partial \theta_{I_E} \partial \theta_{I_E}} \right)_0 \equiv \left( \frac{\partial^2 V}{\partial r_{A_E \mu} \partial r_{A_E \mu}} \right)_0. \quad (12)$$

Here it is assumed that at the reference point the frame attached to the rigid body coincides with the space-fixed frame, which can always be done. The reduced Hessian  $H'^{(ii)}$  of MBH now exactly coincides with the Cartesian  $H'$  of the PHVA. This gives evidence that the discrepancy between PHVA and MBH only lies in a different approach of handling the mass matrix.

The difference between both approaches is best illustrated by the example of two masses  $m_1$  and  $m_2$  moving in one dimension and joined by a spring characterized by a

spring constant  $K$ . Following the method of Sec. II B (PHVA), the mass  $m_1$  is considered fixed in space and the normal frequency of the system (vibration of mass  $m_2$ ) is given by  $\omega^2 = K/m_2$ . In the method of Sec. II C (MBH), on the other hand, the mass  $m_1$  is allowed to participate in the vibration and one simply has the normal frequency of the free spring,  $\omega^2 = K/\mu$ , with  $\mu$  the reduced mass. In the limit of  $m_1 \gg m_2$ , the reduced mass tends to  $m_2$  and the eigenvalues of the two approaches converge to the same value.

Generally, the partial Hessian vibrational analysis model can be considered as a limit case of the mobile block Hessian approach. When the mass and the inertial moments of the rigid body increase with respect to the  $N_E$  free masses, the mass matrix of the MBH in the limit becomes equal to the mass matrix of the PHVA. Thus the higher the fixed atom masses, the more the two models converge to each other.

A more stringent comparison is obtained by inspecting the mass term taken up in Eq. (10). Following the various mass definitions given in Sec. II A, one sees that the fixed atom masses  $m_{A_F}$  only occur in the matrix  $M^{(rr)}$ , which coincides with the inertial matrix, while  $M^{(ri)}$  is independent of the  $m_{A_F}$ . The larger the number of atoms in the fixed part of the molecule, the more the total mass of the rigid body increases. As mentioned above, only the elements of the matrix  $M^{(rr)}$  are influenced, and the second term in the right hand side of Eq. (10) will tend to zero. With the specific choice of internal coordinates as stated in Eq. (11), it is clear that the PHVA frequencies form the limits of the MBH frequencies.

Another (non-negligible) aspect concerns the computational cost of both models. Since both models require the numerical evaluation of the reduced Hessian  $H'$  or  $H'^{(ii)}$ , no difference is expected in computing time. The calculation of the corresponding mass matrices is also straightforward. Since most of the required computation time goes to the evaluation of the second derivatives of the potential energy surface, the reduction of the dimension of the full Hessian to  $3N_E$  is essential to get a serious reduction of the computation time.

The applicability of both methods depends on the type of the simulated system under consideration. It can be expected that the MBH method is better suited to describe molecules in the gas phase, because they move quasifreely in space and can be considered as isolated systems. Assigning an infinite mass to the nonoptimized part of the system may be too crude an approximation. The potential energy surface should be invariant under global translation and rotation, and these global modes should effectively decouple from the internal vibrations.

However, when modeling a lattice using the cluster *in vacuo* approach and fixing the border atoms, the situation is different. The cluster cannot move freely and there is no reason why the description of the cluster should be translationally or rotationally invariant. In fact, the border atoms are more or less pinned down at their positions by the presence of the surrounding infinite lattice, which was, however, left out of the simulation cluster. So it is quite a realistic picture to assume that the border atoms are really fixed by external

forces generated by the surrounding lattice. In other words, the rigid body of fixed border atoms may be assumed to have a fixed position and orientation, and thus, the PHVA approach is the appropriate way to calculate eigenvalues.

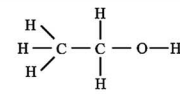
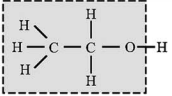
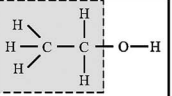
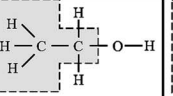
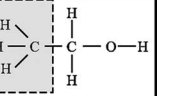
In what follows, both the PHVA and MBH methods are numerically validated and benchmarked against the normal modes extracted from a fully optimized geometry. Independent of the specific application under consideration, some general remarks are useful. Inherent to the PHVA and MBH methods is the introduction of a rigid part of the system that is not optimized, and thus, the number of PHVA and MBH frequencies is always smaller than the number of benchmark frequencies. The benchmark modes generated by atoms belonging to the rigid body are *a priori* not reproduced. Additionally, modes where the displacements are spread out over nonfixed as well as fixed atoms are expected to be very badly described by both the PHVA and MBH. On the other hand, normal modes that are completely localized in the relaxed molecular region with nearly vanishing fixed atom displacements, or normal modes where the fixed atoms move collectively with respect to the optimized region, can be expected to result from the MBH approach.

Note that Head and co-worker investigated the coupling between the PHVA modes and the omitted modes occurring in the fixed atom region.<sup>10-12</sup> It may be possible to apply similar methods to the MBH modes.

### III. APPLICATION TO THE ETHANOL MOLECULE

The PHVA and MBH methods are, in principle, developed to evaluate the frequencies in a large molecular system which cannot be optimized entirely at a high level of theory as the size increases, but in which only part (the active site) of the molecule is optimized at the high level and the remaining part at a substantially lower level of theory. However, in order to understand the advantages and disadvantages of both methods, it is instructive to validate them first on a small molecule, where the exact frequencies are readily available and can be compared with those predicted by PHVA and MBH in partially optimized geometries. To meet this purpose, we have chosen ethanol containing a well localized O–H stretch. The entire molecule has been optimized at a so-called high level [B3LYP/6-31+*g*(*d*)], and this optimized geometry will be used further on as the reference. Frequencies of all present normal modes are obtained by solving the mass-weighted eigenvalue problem [Eq. (4)] of the full  $27 \times 27$  Hessian, and are tabulated in the first column of Table I. They coincide with the values that are obtained from the standard analytical frequency calculation in GAUSSIAN03,<sup>22</sup> and will serve as benchmark values for further comparative studies. An exact treatment should generate six eigenvalues exactly equal to zero, corresponding to the global translation and rotation. In practice, the values differ slightly from zero (varying between  $-1$  and  $8 \text{ cm}^{-1}$ ). Translational frequencies are sensitive to numerical errors in the construction of the Hessian. Rotational frequencies are, in addition, affected by the small residual forces due to the finite convergence criteria. The effect of the almost zero fre-

TABLE I. Normal mode frequencies (in  $\text{cm}^{-1}$ ) of ethanol derived from the benchmark geometry, which corresponds to the geometry optimization obtained at B3LYP/6-31+ $g(d)$ . The rigid body is composed of the atoms in the shaded region. In the left column, translational and rotational frequencies from the full Hessian calculation are plotted before and after projection. Vibrational frequencies are not affected by this projection. The PHVA and MBH frequencies were ordered according to the maximum overlap with the benchmark modes.

									
Full	Projected	PHVA	MBH	PHVA	MBH	PHVA	MBH	PHVA	MBH
-1	0	-	-	-	-	72	-	59	-
-1	0	-	-	-	-	242	-	94	-
0	0	-	-	-	-	678	-	211	-
1	0	-	-	-	-	-	-	355	-
2	0	-	-	-	-	-	-	-	-
8	0	-	-	-	-	-	-	-	-
internal rotation:									
244		-	-	-	-	-	249	-	244
295		275	289	299	289	279	297	292	295
C-C-O stretch:									
417		-	-	-	491	-	458	-	419
mixed modes:									
825		-	-	272	807	823	992	1024	831
902		-	-	-	-	-	1011	-	906
1037		-	-	-	-	-	-	748	1040
1104		-	-	-	1031	-	-	1058	1111
1185		-	-	-	-	-	-	-	1195
1270		1164	1208	1224	1245	1210	1218	1226	1277
1305		-	-	715	-	1254	1269	1262	1308
1419		-	-	-	-	-	-	-	-
1461		-	-	-	-	1317	1378	1445	1452
1505		-	-	-	-	-	-	-	-
1521		-	-	-	-	-	-	-	-
C-H stretch:									
1544		-	-	-	-	1501	1534	1541	1541
2995		-	-	245	-	2928	2977	2995	2995
3021		-	-	-	-	2877	3015	3023	3023
3049		-	-	-	-	-	-	-	-
3116		-	-	-	-	-	-	-	-
3124		-	-	-	-	-	-	-	-
O-H stretch:									
3756		3645	3708	3755	3756	3755	3756	3756	3756

quencies on the other 21 frequencies is negligible here: projecting out the overall translation and rotation, which is implemented in most of the program packages such as GAUSSIAN03, gives six eigenmodes exactly equal to zero and does not affect the vibrational frequencies. Some of the

modes have a clear interpretation: the two lowest frequencies (244 and 295  $\text{cm}^{-1}$ ) correspond to internal rotations of the methyl top and the hydroxyl group, while the highest (3756  $\text{cm}^{-1}$ ) is associated with the highly localized O-H stretching mode.

We applied the PHVA and MBH methods to two different cases: first to the fully optimized structure and then to several partially optimized structures. The PHVA equations (9) are constructed with the submatrix  $H'$  of the Hessian that contains the second derivatives of the potential energy with respect to the Cartesian coordinates of the free atoms. The MBH frequencies corresponding to Eq. (10) are calculated, following the discussion in Sec. II D, with the same submatrix  $H'$  [see Eq. (12)] but with a different mass matrix.

The ordering of the calculated frequencies in the tables is determined by the maximum of the square of the overlap  $|\langle \nu_{\text{bench}} | \nu \rangle|^2$  between the benchmark mode with frequency  $\nu_{\text{bench}}$  and the calculated mode with frequency  $\nu$ . For modes without a pronounced maximum overlap, this is of course rather arbitrary.

### A. PHVA and MBH applied to the equilibrium structure

Frequencies are calculated for the fully optimized geometry, while the part of the Hessian chosen to be included in the vibrational analysis is varied. In this case, there are no remaining residual forces and the difference between the two models (PHVA or MBH) can easily be studied. The results are given in Table I for various rigid body sizes. The shaded box indicates the part of the molecule that is not included in the calculation of the Hessian. For instance, the second column of the table reports the three frequencies corresponding to the modes generated by a rigid body and one single atom that can vibrate. The only nonfixed atom in the vibrational analysis is the hydrogen atom of the hydroxyl group. In the third column, the whole hydroxyl group can vibrate, while in the last column we display the situation with only the methyl group fixed at its reference geometry.

Only three PHVA and MBH modes are calculated in the case where only the H atom of the hydroxyl group is taken into account (second column of Table I). Visualization of the three modes revealed that they all qualitatively correspond to a normal mode of the benchmark frequency spectrum, but that their values are underestimated (benchmark values are 295, 1270, and 3756  $\text{cm}^{-1}$ ). In the PHVA method, the underestimation is even more pronounced due to the use of infinite masses. This is a general conclusion: the PHVA fails in accurately reproducing the benchmark values, especially in the small and medium frequency regions, while the quantitative agreement of the MBH predictions of localized modes—taking place in the nonfixed region—with the benchmark values is manifestly present. The agreement is even very close in the last case, where the rigid body consists of a fixed methyl group. The O–H stretch mode is always present in the frequency spectrum, only in the case of one relaxed atom (the hydrogen) is the obtained frequency slightly underestimated.

Looking at lower frequencies, the MBH approach gives consistently better results than the PHVA, because the reduced mass effect is taken into account. As the total mass of the fixed block decreases, it is obvious that the PHVA induces spurious low frequency modes of the order of 60–100  $\text{cm}^{-1}$ , corresponding to translations/rotations of the nonfixed atoms in the field of their environment (the fixed

atoms). Summarizing, when working with molecules in the gas phase, the use of the MBH model is highly recommended as soon as the total mass of the fixed block becomes of the same order as the total mass of the relaxed atoms.

### B. PHVA and MBH applied to partially optimized structures

We introduce partially optimized geometries and verify whether the relevant calculated frequencies in the active site can be reproduced by both the PHVA and MBH models in an accurate way with respect to the benchmark frequencies. Initially, the ethanol molecule has been optimized at the lower HFO/STO-3g level. Then the system was partially optimized at the higher B3LYP/6-31+g(d) level while keeping the atoms of the rigid block fixed at their initial HF/STO-3g positions. Results are collected in Table II. The atoms belonging to the shaded box were not optimized at the high level but were kept fixed at their low level geometries. For each case, the frequencies resulting from a full Hessian [calculated at B3LYP/6-31+g(d)] diagonalization, i.e., from a standard normal mode analysis in Gaussian, as well as the normal modes resulting from the PHVA and MBH methods are tabulated. The standard frequency analysis always gives a number of spurious imaginary frequencies, as could be expected since residual forces on the nonoptimized atoms disturb the evaluation of the frequencies. In addition, the positive frequencies deviate substantially from the benchmark values given in Table I. This means that the normal frequencies generated by standard procedures in program packages such as GAUSSIAN03,<sup>22</sup> in molecules whose atomic positions have not been optimized at the same level of theory, are far from being accurate.

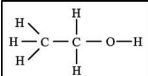
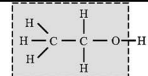
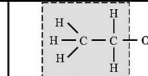
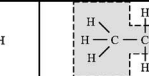

By applying the PHVA or the MBH model, the unphysical imaginary frequencies disappear, but the resulting frequencies, however, differ significantly between the two methods. The MBH results converge rapidly to the benchmark values, highlighting the efficiency of the proposed MBH model. A striking resemblance is even observed in the last column where the fixed body is restricted to the ending methyl group. The low frequency spectrum of the PHVA method, however, deviates largely from the benchmark values. This is entirely due to the reduced mass effect, inducing spurious unphysical modes.

While the full Hessian frequencies are very sensitive to the exact molecular geometry, it is remarkable that PHVA and MBH frequencies from the partially optimized geometry are very close to those obtained with the benchmark geometry with the same block size. This indicates that the PHVA and MBH models are less sensitive to the exact internal geometry of the fixed block. This could be important for future applications in large systems.

It is also interesting to compare the thermodynamical quantities associated with the vibrational part of the partition function of the system. In Fig. 2 the vibrational contribution to the entropy  $S$  and the free enthalpy  $G$  are given for the different partially optimized ethanol configurations at  $T = 298.15$  K. The values calculated with the benchmark frequencies are also indicated. It is clear that reducing the num-



TABLE II. Normal mode frequencies (in  $\text{cm}^{-1}$ ) of ethanol derived on the basis of partially optimized geometries at the B3LYP/6-31+ $g(d)$  level of theory. The rigid body is composed of atoms in the shaded region and its geometry is originated from a geometry optimization of the whole molecule at the low HF/STO-3g level. Benchmark frequencies are given in the left column for comparison. The PHVA and MBH frequencies were ordered according to the maximum overlap with the benchmark modes.

												
bench	Full PHVA MBH			Full PHVA MBH			Full PHVA MBH			Full PHVA MBH		
-1	-258	-	-	-264	-	-	-254	71	-	-251	60	-
-1	-130	-	-	-132	-	-	-80	243	-	-90	95	-
0	-62	-	-	-75	-	-	-61	678	-	-66	212	-
1	-1	-	-	-2	-	-	-1	-	-	-1	355	-
2	-1	-	-	0	-	-	1	-	-	1	-	-
8	1	-	-	1	-	-	1	-	-	1	-	-
244	91	-	-	42	245	-	124	-	248	124	-	246
295	294	281	295	284	296	285	289	280	298	289	294	297
417	403	-	-	396	-	490	397	-	458	397	-	419
825	760	-	-	758	272	813	793	824	992	793	746	839
902	873	-	-	877	-	-	880	-	-	883	-	909
1037	1021	-	-	1025	-	1032	1025	-	1013	1029	1025	1041
1104	1070	-	-	1096	-	-	1097	-	-	1098	1058	1116
1185	1147	-	-	1156	-	-	1177	-	-	1177	-	1201
1270	1257	1161	1205	1261	1224	1244	1266	1210	1218	1267	1225	1280
1305	1267	-	-	1277	715	-	1298	1252	1268	1300	1262	1313
1419	1384	-	-	1385	-	-	1388	-	-	1390	-	-
1461	1422	-	-	1429	-	-	1452	1315	1376	1454	-	-
1505	1489	-	-	1488	-	-	1488	-	-	1488	1444	1452
1521	1505	-	-	1506	-	-	1507	-	-	1508	-	-
1544	1527	-	-	1527	-	-	1544	1502	1535	1544	1542	1542
2995	3140	-	-	3137	-	-	2995	2877	2978	2994	2995	2995
3021	3162	-	-	3162	-	-	3023	2929	3016	3022	3023	3023
3049	3177	-	-	3173	-	-	3162	-	-	3162	-	-
3116	3242	-	-	3243	-	-	3243	-	-	3242	-	-
3124	3253	-	-	3252	-	-	3251	-	-	3251	-	-
3756	3751	3640	3703	3755	3755	3755	3756	3756	3756	3756	3756	3756

ber of modes by introducing the rigid block affects both the vibrational entropy and free enthalpy. However, the MBH values tend consistently to the benchmark value when the block size decreases, whereas the PHVA values do not. This is to be expected, since PHVA assumes infinite mass for the fixed block and for the present application the MBH is physically more relevant.

#### IV. EXTENSION: MULTIPLE MOBILE BLOCKS

The application field of MBH can be easily extended to multiple mobile blocks. Suppose the molecule is decomposed into  $K$  rigid blocks and  $N_E$  freely relaxed atoms. The whole molecule is optimized at a low level of theory determining the position of each atom in each of the rigid blocks. In a second step, a higher level of theory is used and one optimizes the positions of the  $N_E$  free atoms, as well as the positions and orientations of the rigid blocks, keeping their original internal geometry. As a consequence, residual forces

remain between the fixed atoms within a block, thus the full gradient is nonzero, and a normal frequency analysis along the lines of Eq. (7) with the full Hessian will therefore produce unphysical frequencies.

In the extended method, one assumes that each block is allowed to participate in the small amplitude vibrations, moving as a rigid body. To implement this approach, a suitable choice of internal coordinates is necessary, e.g., the Z-matrix formalism. It is, in fact, sufficient to choose the numbering of the atoms in the Z-matrix construction such that the atoms in each fixed block are numbered consecutively.

As in Sec. II C, the geometry of each block  $b$  is described by  $3N_{F,b}-6$  internal coordinates  $\{\theta_{I_{F,b}}\}$ . The imposed reference structure  $\{\mathbf{r}_{A_{F,b}}^0\}$  of the fixed atoms of block  $b$  then determines the values of the  $3N_{F,b}-6$  internal coordinates  $\{\theta_{I_{F,b}}^0\}$ . The remaining  $3N_E+6(K-1)$  internal coordinates will

be labeled  $\{\theta_{I_E}\}$ . Optimizing the energy with fixed  $\{\theta_{I_{F,b}}^0\}$  and varying  $\{\theta_{I_E}\}$  yield the reference structure  $\{\theta_{I_E}^0\}$ .

To impose the fixed internal geometry of each block during the vibrational analysis, the displacements  $\Delta_{I_{F,b}}$  and corresponding velocities  $\dot{\Delta}_{I_{F,b}}$  are set equal to zero in the second-order expansion of Eq. (6). The gradient term in Eq. (6) thereby vanishes, so the reduced system of  $3N_E+6(K-1)$  internal coordinates  $\{\theta_{I_E}\}$  is in equilibrium.

The corresponding NMA equations are obtained by omitting the rows and columns related to the  $\{\theta_{I_{F,b}}^0\}$  variables from the Hessian and mass matrix in Eq. (7). The resulting reduced eigenvalue problem of dimension  $3N_E+6K$  has six zero eigenvalues corresponding to overall translation and rotation. These can be decoupled as in the mobile block Hessian approach with a congruent transformation similar to Eq. (10).

The discussion of this multiple MBH approach is similar to the single MBH with only one rigid body. First, the disturbing negative eigenvalues due to residual forces are eliminated, and imaginary frequencies will only occur if the reduced system of relaxed atoms and  $K$  blocks is actually in a transition state. The presence of six zeros shows that the translational and rotational invariance are respected in the second-order expansion of the potential energy of the reduced system. Concerning the mass effect, the finite mass of each block is taken into account. Moreover, the method provides a reduction in computer time, because for each block  $b$  with  $N_{F,b}$  atoms, only derivatives with respect to six coordinates instead of  $3N_{F,b}$  have to be calculated.

Of course, the main question is whether the multiple MBH approach is capable of reproducing the normal mode frequencies of the fully optimized benchmark structure. Generally, two types of benchmark modes are well reproduced by the multiple MBH in analogy with the single MBH ver-

sion: (i) modes localized in the relaxed part of the molecule (ii) and modes where the fixed atoms move as a whole. Some other modes present in the benchmark, on the other hand, are obviously not reproduced by the multiple MBH. These are modes in which the internal geometry of the rigid blocks is changed. Such modes are not localized in the chemically active part of the molecule and are therefore not relevant for the present study. Modes where the displacements are spread out over the relaxed as well as fixed atoms are mostly not reproduced in the multiple MBH model, except those where the atoms of the rigid bodies move coherently [modes of type (ii)]. Some examples of such modes are shown in Fig. 1. For instance, the rigid bodies can rotate about the single bonding axis connecting them with the active part of the molecule or they can contribute to a stretching, a bending motion, etc.

For practical purposes, it is important to select the freely relaxed region of the molecule attentively. One of the main advantages of this extension to multiple MBH is the computational profit as a consequence of the reduced number of Hessian matrix elements to be evaluated, but unfortunately this option is not (yet) present in most of the standard program packages.

## V. APPLICATION TO DI-N-OCTYL-ETHER

A suitable example to validate the multiple MBH approach is di-*n*-octyl-ether ( $C_8H_{17}-O-C_8H_{17}$ ). Ethers are known to have a C–O stretching band that falls in the fingerprint region at  $1050-1260\text{ cm}^{-1}$ . This vibration is a localized mode and very characteristic for all ethers. The ability of the various models in reproducing this stretching band is a strong validation for the MBH method.

The geometry of the molecule is first fully optimized at the B3LYP/6-31+ $g(d)$  level of theory, and a normal mode

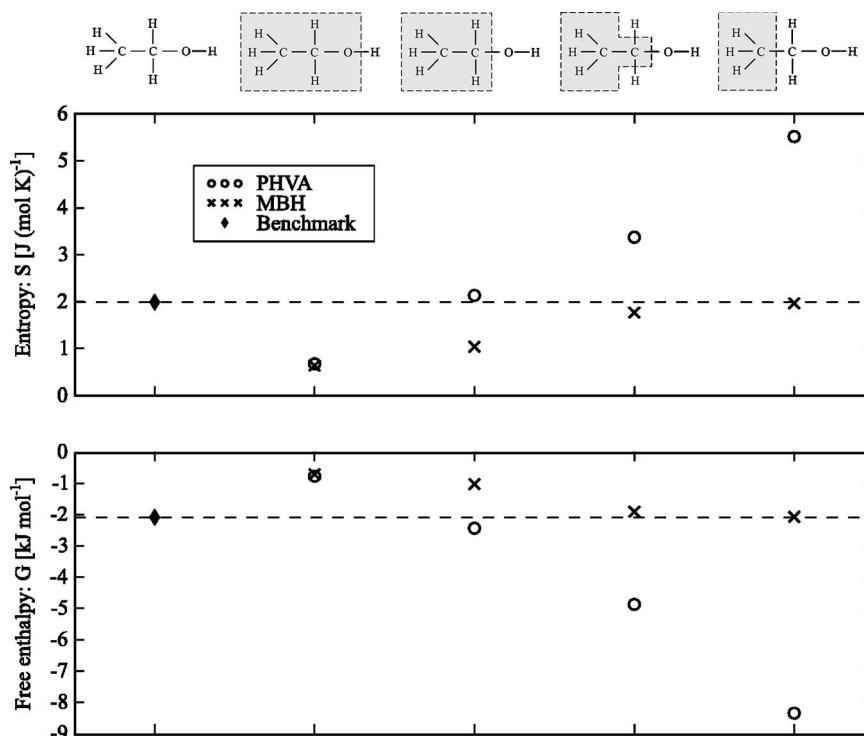


FIG. 2. Vibrational contribution to the entropy  $S$  and the free enthalpy  $G$  calculated with PHVA ( $\circ$ ) and MBH ( $\times$ ) frequencies are given for the different partially optimized ethanol configurations at  $T=298.15\text{ K}$ . Benchmark values are indicated by the dashed lines. The fixed block in the MBH calculation consists of the atoms in the shaded box.

TABLE III. Frequencies  $\lambda^{1/2}$  (in  $\text{cm}^{-1}$ ) of di-*n*-octyl-ether of various partially optimized configurations defined in Fig. 3 are compared with the benchmark frequencies of the fully optimized geometry (left column). Three approaches are used: the full Hessian calculation (Full), the PHVA method, and the MBH approach. The size of the rigid bodies is defined by the configuration label. The Full/PHVA/MBH frequencies are ordered according to the maximum overlaps  $|\langle \nu_{\text{bench}} | \nu \rangle|^2$  with benchmark eigenmodes, which are given by the values between parentheses (in %).

Bench	Configuration 8						Configuration 5						Configuration 2					
	Full	PHVA		MBH		Full	PHVA		MBH		Full	PHVA		MBH				
0	0	(100)	-	(-)	-1	(100)	0	(100)	29	(33)	-1	(100)	0	(100)	12	(56)	-1	(100)
0	0	(100)	-	(-)	0	(100)	0	(100)	-	(-)	-1	(100)	0	(100)	134	(73)	0	(100)
0	0	(100)	-	(-)	0	(100)	0	(100)	-	(-)	0	(100)	0	(100)	22	(58)	0	(100)
1	-8	(95)	-	(-)	1	(100)	7	(92)	-	(-)	1	(100)	9	(88)	228	(17)	0	(100)
1	-12	(96)	-	(-)	0	(100)	-12	(96)	-	(-)	1	(100)	-9	(98)	-	(-)	-1	(100)
4	-156	(61)	-	(-)	4	(97)	-124	(29)	-	(-)	4	(99)	-	(-)	34	(67)	3	(100)
11	9	(65)	-	(-)	16	(95)	27	(54)	-	(-)	12	(99)	34	(53)	228	(17)	11	(100)
19	-15	(98)	-	(-)	24	(98)	-15	(99)	50	(40)	19	(100)	-9	(99)	-	(-)	19	(100)
27	-154	(40)	-	(-)	42	(67)	-124	(32)	-	(-)	20	(99)	-36	(29)	63	(49)	27	(100)
33	53	(37)	-	(-)	64	(54)	-52	(43)	-	(-)	35	(98)	64	(34)	34	(59)	33	(100)
51	36	(96)	-	(-)	114	(78)	37	(95)	132	(44)	55	(99)	40	(97)	53	(60)	51	(100)
53	-104	(34)	215	(30)	283	(52)	-53	(41)	-	(-)	60	(94)	37	(53)	-	(-)	53	(100)
65	-	(-)	-	(-)	-	(-)	57	(50)	57	(39)	79	(86)	-	(-)	63	(42)	65	(100)
84	-96	(25)	-	(-)	-	(-)	61	(32)	79	(30)	101	(71)	75	(66)	-	(-)	84	(100)
92	-	(-)	-	(-)	-	(-)	90	(48)	101	(33)	117	(63)	-	(-)	95	(78)	92	(100)
96	84	(98)	261	(22)	-	(-)	86	(97)	-	(-)	114	(93)	87	(98)	102	(72)	96	(100)
113	77	(32)	-	(-)	-	(-)	90	(34)	-	(-)	-	(-)	109	(51)	95	(51)	113	(100)
129	90	(57)	-	(-)	-	(-)	118	(76)	-	(-)	362	(54)	153	(34)	133	(87)	130	(100)
133	110	(41)	244	(50)	300	(47)	116	(52)	144	(48)	-	(-)	132	(68)	122	(62)	133	(100)
138	137	(100)	-	(-)	180	(88)	138	(100)	-	(-)	145	(99)	138	(100)	276	(56)	138	(100)
150	-	(-)	-	(-)	-	(-)	139	(83)	228	(44)	362	(51)	169	(53)	149	(62)	150	(100)
154	143	(98)	-	(-)	-	(-)	144	(99)	-	(-)	198	(79)	145	(99)	167	(71)	154	(100)
160	127	(44)	-	(-)	-	(-)	152	(80)	-	(-)	149	(54)	158	(98)	157	(68)	160	(100)
162	128	(46)	-	(-)	-	(-)	150	(80)	130	(36)	143	(55)	162	(73)	154	(76)	162	(100)
178	160	(83)	125	(29)	143	(41)	175	(95)	230	(57)	169	(78)	182	(82)	178	(95)	178	(100)
220	209	(99)	-	(-)	336	(27)	210	(99)	243	(55)	209	(56)	212	(99)	233	(83)	220	(100)
...																		
1149	1146	(99)	1146	(94)	1147	(95)	1148	(100)	1148	(100)	1148	(100)	1149	(100)	1149	(100)	1149	(98)
1461	1457	(99)	1442	(83)	1442	(83)	1461	(100)	1461	(100)	1461	(100)	1461	(100)	1461	(100)	1461	(100)
1549	1547	(98)	1545	(88)	1545	(88)	1549	(100)	1548	(100)	1548	(100)	1549	(100)	1549	(100)	1549	(100)
2954	2953	(100)	2953	(99)	2953	(99)	2954	(100)	2954	(100)	2954	(100)	2954	(100)	2954	(100)	2954	(100)
2965	2965	(99)	2965	(99)	2965	(99)	2965	(100)	2965	(100)	2965	(100)	2965	(100)	2965	(100)	2965	(100)
2983	2983	(99)	2984	(98)	2984	(98)	2983	(100)	2983	(100)	2983	(100)	2983	(100)	2983	(100)	2983	(100)
2985	2985	(99)	2986	(98)	2986	(98)	2985	(100)	2985	(100)	2985	(100)	2985	(100)	2985	(100)	2985	(100)

analysis at the same level, constructed with analytical second derivatives, provides the whole spectrum of frequencies, which serve as benchmark values. Some eigenfrequencies are listed in the left column of Table III. The full Hessian frequency analysis on this completely optimized geometry generates six (almost) zero eigenvalues, as it should be. In the *low* frequency spectrum, typical modes are found that involve large parts of the molecule. In this class, bending and torsionlike modes are found where large massive blocks of the molecule are involved. Also internal rotations and collective accordionlike motions are present. For instance, the three lowest benchmark frequencies (11, 19, and 27  $\text{cm}^{-1}$ ) are identified as the relative rotation/torsion of the two octyl chains around the three axes of inertia of the molecule. In view of later application in which the frequencies are used as input for the construction of the partition function in the harmonic oscillator approximation, this low frequency spectrum is extremely important as they give the largest contri-

bution to the vibrational partition function. In the *high* frequency range ( $\sim 3000 \text{ cm}^{-1}$ ), localized C–H stretches are found. Some specific modes are interesting as they are localized near the central O atom: 2985, 2983, 2965, and 2954  $\text{cm}^{-1}$  correspond to C–H stretching modes—symmetrically or antisymmetrically—of the  $\text{CH}_2$  moieties attached to the middle oxygen, 1461  $\text{cm}^{-1}$  is mainly a bending of these two  $\text{CH}_2$  units with respect to each other and 1149  $\text{cm}^{-1}$  is the typical C–O stretch for ethers.

In what follows, the multiple MBH model is applied to a set of partially optimized structures at the B3LYP/6-31 +  $g(d)$  level of theory, while part of the molecule is kept fixed at the geometry optimized at the lower HF/STO-3g level. The central part of the molecule is always allowed to relax. Two nonoptimized blocks are systematically taken into account, located at both sides of the central part. They differ in the length of the chain that is considered in the nonoptimized

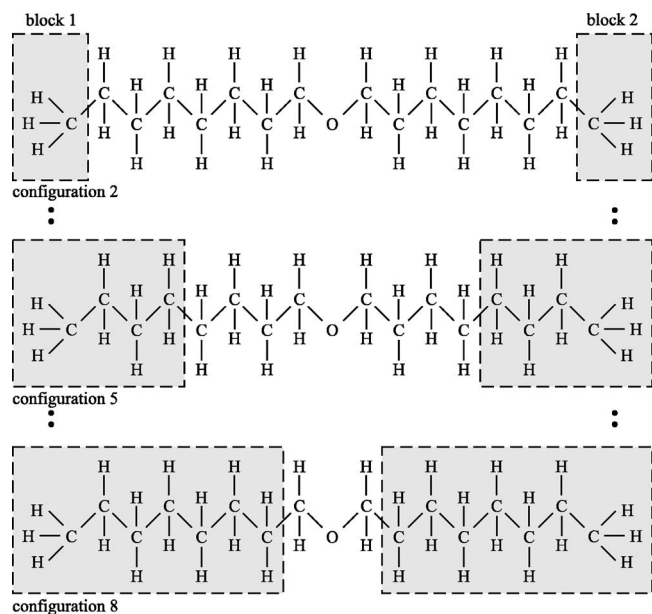


FIG. 3. Specification of the various configurations of di-*n*-octyl-ether with rigid bodies indicated as shaded regions. Atoms in shaded boxes are fixed at HF/STO-3 $g$  positions during the partial geometry optimization at the B3LYP/6-31+ $g(d)$  level.

block, as schematically depicted in Fig. 3. The influence of varying chain length on the reproduction of the frequency spectrum will be investigated. The fully optimized case is referred to as configuration 1 (the benchmark geometry), whereas in configuration 8, two heptyl chains are constrained during the optimization.

The low energy part of the spectra obtained by diagonalizing the full Cartesian Hessian (as would be done in the common normal mode analysis) is represented in Fig. 4. The same deficiencies related to nonequilibrium geometries as in the ethanol example are noticed: the appearance of non-negligible negative (imaginary) frequencies of the order of  $-225\text{ cm}^{-1}$  and the lack of three eigenvalues equal to zero belonging to the global rotation. Moreover, the number of spurious negative frequencies can vary depending on the number of atoms belonging to the fixed block and is thus

unpredictable. The repercussion of the negative frequencies is mainly found in the low frequency modes, where large parts of the molecule, including some nonoptimized atoms, are involved. Values for a selected number of modes are given in Table III. The reproduction of the three lowest eigenfrequencies is very poor and even negative values are found. It is, nevertheless, remarkable that the previously mentioned localized modes are very well reproduced even within the presence of some negative frequencies. For instance, the localized C–O band, situated at  $1149\text{ cm}^{-1}$ , is well reproduced, probably owing to the fact that only optimized atoms are involved in this internal motion.

This numerical example illustrates the necessity to strictly reduce the dimension of the Hessian, following the procedure explained in the (multiple) MBH approach, in order to get rid of the disturbing negative frequencies. The low frequency part of the resulting spectra is plotted in Fig. 5 for the different configurations. Thereby all negative eigenvalues are eliminated, six eigenvalues are equal to zero as could be expected, and the other frequencies are all physically significant. Figure 5 allows qualitative comparison between the MBH normal mode frequencies and the benchmark values. Obviously, the MBH results belonging to the largest fixed blocks (configuration 8) differ substantially from the benchmark values. As the number of relaxed atoms increases, the agreement improves (configuration 2). Quantitative values are given for a selected number of modes in Table III, where the MBH modes were sorted according to their resemblance to the benchmark modes (see caption of table). The PHVA method was applied as well, combining the two individual fixed blocks into one fixed block, as shown in Fig. 6. The PHVA method is capable of reproducing localized modes, but frequencies in the lower spectrum are very poorly reproduced or are absent. Apparently, the MBH model is able to reproduce accurately not only the previously mentioned localized modes, but more importantly, the low frequency modes have much more realistic values compared to the full Hessian values or the PHVA values.

A more pronounced study can be made by evaluating the overlap between the calculated multiple MBH modes and

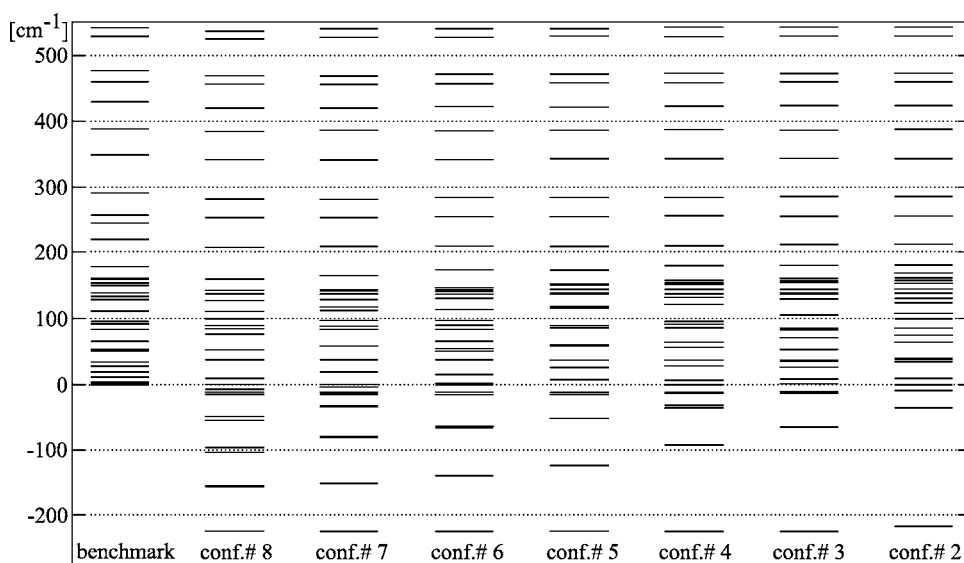


FIG. 4. Lowest frequencies  $\lambda^{1/2}$  (in  $\text{cm}^{-1}$ ) of di-*n*-octyl-ether based on the full Cartesian Hessian belonging to the various partially optimized configurations defined in Fig. 3. Partial optimization at the B3LYP/6-31+ $g(d)$  level. Plot on the left displays the exact normal mode frequencies (full geometry optimization) that serve as benchmark.



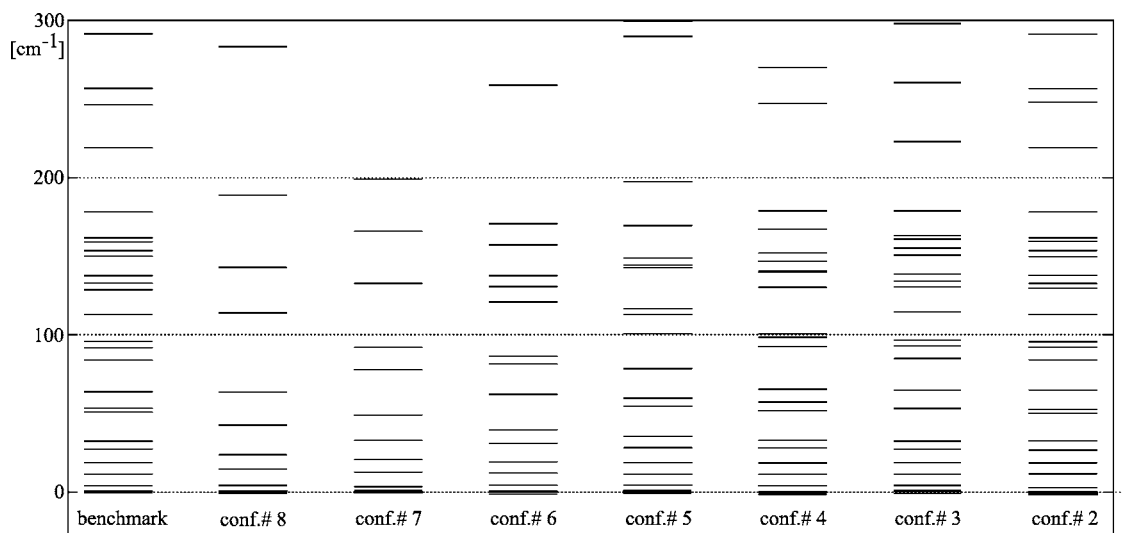


FIG. 5. Lowest frequencies  $\lambda^{1/2}$  (in  $\text{cm}^{-1}$ ) of di-*n*-octyl-ether based on the multiple MBH model belonging to the various partially optimized configurations defined in Fig. 3. Partial geometry optimization at the B3LYP/6-31+*g*(*d*) level. Plot on the left displays the exact normal mode frequencies (full geometry optimization) that serve as benchmark for the other plots where two rigid bodies (defined by the configuration label) are taken into account in the frequency analysis.

benchmark modes. The square of the overlap  $|\langle \nu_{\text{bench}} | \nu_{\text{MBH}} \rangle|^2$  gives a measure of how strong the benchmark mode with frequency  $\nu_{\text{bench}}$  is involved in the specific MBH mode with frequency  $\nu_{\text{MBH}}$ . Due to the completeness of the basis of benchmark modes, the sum of these numbers over all benchmark frequencies is equal to 1. The results for configurations 8, 5, and 2 are given in a scatter plot in Fig. 7. All values above a certain limit (we take 20% throughout this work) are indicated by a circle, and the darker the fill intensity of the circle, the larger the magnitude of the overlap. The most ideal case is a black filled circle on the diagonal, as it means that the multiple MBH model has reproduced the benchmark mode in an excellent way. If the strength of a specific MBH mode is spread out over various benchmark modes, a larger scattering of circles is noticed off the diagonal, which is indeed the case for configuration 8.

By enlarging the optimized region, the discrepancy with the benchmark almost disappears. In fact, already from configuration 5 results have essentially converged (see Fig. 7). By retaining only the ending methyl groups in the fixed boxes (configuration 2), all relevant frequencies are excellently reproduced. In the low frequency region (below  $1500 \text{ cm}^{-1}$ ), they are identical within a margin of  $1 \text{ cm}^{-1}$ . Only a few nonrelevant modes that are localized in the methyl tops are absent.

In Table III the maximum overlap  $|\langle \nu_{\text{bench}} | \nu \rangle|^2$  between benchmark modes  $|\nu_{\text{bench}}\rangle$  and the selected full Hessian,

PHVA, and MBH modes  $|\nu\rangle$  is added between brackets. The overlap for the localized modes is consistently excellent. The overlap for the full Hessian or PHVA modes is quite poor for the low frequencies, whereas the MBH modes show very reasonable overlaps even in this region.

## VI. SUMMARY AND CONCLUSIONS

In this paper a new method, referred to as the MBH approach, was introduced to calculate the normal modes of partially optimized molecular systems. The MBH approach is an extension of the previously introduced PHVA method in the sense that the nonoptimized regions of the system are allowed to move as rigid blocks with respect to the optimized part of the system. Both the MBH and PHVA methods eliminate the imaginary frequencies that result from the common full Hessian normal mode analysis on a partially optimized structure. In cases where the surrounding nonoptimized part of the system effectively are kept immobile by external forces at its reference position, such as in a lattice, the PHVA and MBH methods perform equally well, but for a flexible surrounding medium the MBH method is more appropriate. The various methods were outlined in two examples, i.e., ethanol and di-*n*-octyl-ether. It was found that the localized modes in the optimized part of the system are always well reproduced, irrespective of the applied method. Even a full Hessian normal mode analysis gives quite accurate values for the frequencies of localized modes. However, for normal modes that involve a larger part of the molecular system, the MBH method performs better, since the nonoptimized blocks can move coherently with respect to the chemically active part of the system.

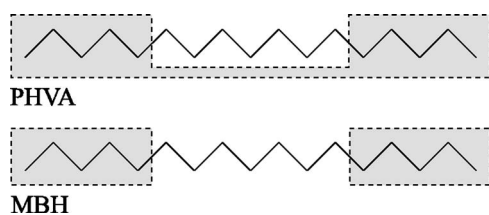


FIG. 6. The PHVA method implies the introduction of one block. For the multiple MBH method, two rigid blocks were used.

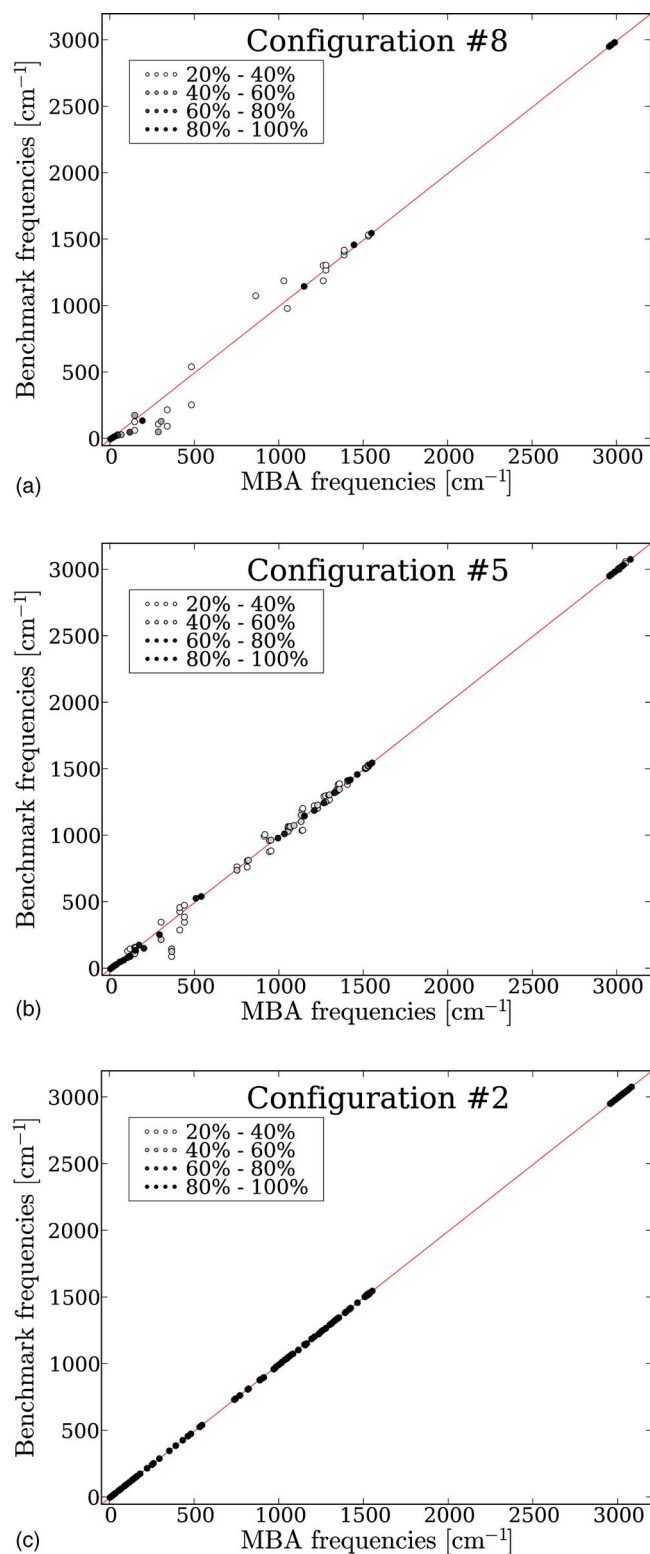


FIG. 7. Square of the overlap between the MBH normal modes and the benchmark normal mode frequencies of di-*n*-octyl-ether. The sum of the strengths is always normalized to 1 for each MBH frequency.

The MBH method does not only eliminate spurious negative frequencies but implicates also a serious reduction of the computational cost for large molecular systems, as the calculation of the Hessian is the most expensive part after a geometry optimization. Molecular modeling focuses more and more on these extended molecular systems, and hence, such efficient techniques are indispensable. The multiple MBH model looks highly suited for use in systems in which the molecular environment is rather flexible, such as reaction in solvents. Most of the solvent molecules can be regarded as rigid bodies moving in all directions with respect to the optimized central part of the system. Their participation to the normal modes can be simulated by the MBH model. The application of the MBH approach for the calculation of partition functions and derived quantities will be further investigated in the future.

## ACKNOWLEDGMENTS

This work is supported by the Fund for Scientific Research-Flanders and the Research Board of Ghent University. This work was partly performed within the framework of the SBO-BIPOM program of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen).

- <sup>1</sup> A. Warshel and M. Levitt, *J. Mol. Biol.* **103**, 227 (1976).
- <sup>2</sup> X. Assfeld and J. L. Rivail, *Chem. Phys. Lett.* **263**, 100 (1996).
- <sup>3</sup> J. L. Gao, P. Amara, C. Alhambra, and M. J. Field, *J. Phys. Chem. A* **102**, 4714 (1998).
- <sup>4</sup> Y. K. Zhang, T. S. Lee, and W. T. Yang, *J. Chem. Phys.* **110**, 46 (1999).
- <sup>5</sup> F. Stevens, H. Vrielinck, V. Van Speybroeck, E. Pauwels, F. Callens, and M. Waroquier, *J. Phys. Chem. B* **110**, 8204 (2006).
- <sup>6</sup> S. Q. Jin and J. D. Head, *Surf. Sci.* **318**, 204 (1994).
- <sup>7</sup> M. D. Calvin, J. D. Head, and S. Q. Jin, *Surf. Sci.* **345**, 161 (1996).
- <sup>8</sup> H. Li and J. Jensen, *Theor. Chem. Acc.* **107**, 211 (2002).
- <sup>9</sup> N. A. Besley and K. A. Metcalf, *J. Chem. Phys.* **126**, 7 (2007).
- <sup>10</sup> J. D. Head, *Int. J. Quantum Chem.* **65**, 827 (1997).
- <sup>11</sup> J. D. Head and Y. Shi, *Int. J. Quantum Chem.* **75**, 815 (1999).
- <sup>12</sup> J. D. Head, *Int. J. Quantum Chem.* **77**, 350 (2000).
- <sup>13</sup> H. Lin, J. Z. Pu, T. V. Albu, and D. G. Truhlar, *J. Phys. Chem. A* **108**, 4112 (2004).
- <sup>14</sup> A. Tachibana and K. Fukui, *Theor. Chim. Acta* **49**, 321 (1978).
- <sup>15</sup> D. J. Wales, *J. Chem. Phys.* **113**, 3926 (2000).
- <sup>16</sup> R. Murry, J. T. Fourkas, L. Wu-Xiong, and T. Keyes, *J. Chem. Phys.* **110**, 10410 (1999).
- <sup>17</sup> I. Kolossvary and C. McMartin, *J. Math. Chem.* **9**, 359 (1992).
- <sup>18</sup> J. Goldstone, A. Salam, and S. Weinberg, *Phys. Rev.* **127**, 965 (1962).
- <sup>19</sup> V. Van Speybroeck, D. Van Neck, and M. Waroquier, *J. Phys. Chem. A* **104**, 10939 (2000).
- <sup>20</sup> V. Van Speybroeck, P. Vansteenkiste, D. Van Neck, and M. Waroquier, *Chem. Phys. Lett.* **402**, 479 (2005).
- <sup>21</sup> P. Vansteenkiste, D. Van Neck, V. Van Speybroeck, and M. Waroquier, *J. Chem. Phys.* **124**, 044314 (2006).
- <sup>22</sup> M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 03, Revision C.02, Gaussian, Inc., Wallingford, CT, 2004.