Theor Chem Acc (2009) 123:413–429
DOI 10.1007/s00214-009-0532-1

REGULAR ARTICLE

An adaptive density-guided approach for the generation of potential energy surfaces of polyatomic molecules

Manuel Sparta · Daniele Toffoli · Ove Christiansen

Received: 16 December 2008 / Accepted: 10 February 2009 / Published online: 1 March 2009

Abstract We present an adaptive density-guided approach for the construction of Born–Oppenheimer potential energy surfaces (PES) in rectilinear normal coordinates for use in vibrational structure calculations. The procedure uses one-mode densities from vibrational structure calculations for a dynamic sampling of PESs. The implementation of the procedure is described and the accuracy and versatility of the method is tested for a selection of model potentials, water, difluoromethane and pyrimidine. The test calculations illustrate the advantage of local basis sets over harmonic oscillator basis sets in some important aspects of our procedure.

Keywords Potential energy surfaces · Grid construction · Adaptive density-guided approach · Vibration · Vibrational self consistent field

1 Introduction

Within the Born–Oppenheimer (BO) approximation, the atomic nuclei move in a potential generated by the electrons. For all but the simplest molecules, the large dimensionality of the energy hyper-surface ($3N - 6$ or $3N - 5$ for linear molecules, where N is the number of atoms in the molecule) prevents the calculation of the fully coupled potential. The standard route is, therefore, to retain in the expansion of the potential only a selection of the mode couplings (MCs) or entire classes of MCs [1–7]. The approximate potential energy surface (PES) could, for example, include all one- two- and three-mode coupling terms while the higher mode couplings are simply neglected.

Even if a restricted MC representation of the fully coupled potential effectively addresses the issue of the large dimensionality of the quantum dynamical problem, an efficient implementation of the procedure has to provide an accurate representation of the potential surface with as few electronic structure calculations as possible. In order to achieve such efficiency the PES should be sampled in the relevant configurational space and this depends in a more or less complex way on the number and character of the quantum states and/or dynamics of interest. The latter point is apparently an aspect which is often taken into account manually in the construction of PESs for vibrational calculations.

In a recent work [8], we outlined a strategy for an efficient sampling of the PES in a set of grid points input by the user. In this article, we describe an adaptive density-guided approach (ADGA) that allows for a dynamical generation of the set of grids onto which each of the potential energy terms included in the restricted mode-coupling expansion of the fully coupled PES (vide infra) are evaluated. In the ADGA scheme, the optimal grid extensions and the mesh of grid points are determined iteratively by using approximate densities for the target states in the calculation; the PES generation step and the

M. Sparta (✉) · D. Toffoli · O. Christiansen
Department of Chemistry, The Lundbeck Foundation Center for Theoretical Chemistry, University of Aarhus, Langelandsgade 140, 8000 Aarhus C, Denmark
e-mail: msparta@chem.au.dk

O. Christiansen
e-mail: ove@chem.au.dk

Present Address:
D. Toffoli
Department of Chemistry, Middle East Technical University, 06531 Ankara, Turkey
e-mail: dtoffoli@metu.edu.tr

vibrational structure calculation are combined in such a way that outputs of the vibrational calculations are used to estimate where the next set of evaluation points (i.e. single-point electronic structure calculations) have to be added in the grid domains, until specific convergency criteria are met.

Our procedure, in some sense, generates the potential and the wave function hand in hand. Thus, the configurational space explored in the potential sampling is adjusted relative to the vibrational quantum states calculated, thus assuring that in the calculation of the states of interest only the relevant part of the potential comes in play. Clearly, there must be some assumptions and approximations behind this and they will be described and discussed throughout the article. However, as we seek to develop quantum dynamical methods for larger molecular systems, we must create methods that automatically put the emphasis where it is needed. In this way, savings in man-time and computer-time can be achieved while, at the same time, the accuracy of the PES increases since it is ensured that enough points are distributed in critical regions.

The idea of calculating only points of the potential according to need is not new. The direct dynamics approach to classical molecular dynamics is a simple example, where each potential value and derivative (the force) are calculated on the fly when needed for the numerical integration of the classical equations of motion. Also in vibrational structure theory direct algorithms have been developed where the points needed for calculating the vibrational wave function are obtained by calling an electronic structure program [9, 10]. The idea followed in the ADGA is somewhat different in that (i) we actually build up an intermediate potential representation in contrast to the classical direct dynamics approach above and (ii) we do not define before hand the parameters of the grids of points that have to be calculated as the aforementioned direct methods and as was done in our previous work [8]. In some sense, our approach is in the spirit of the growing a surface strategy of Collins et al. [11] where a surface is generated based on classical mechanical simulations of a reactive event, will be later used in a more advanced calculation. In our context, we use quantum wave functions for bound vibrational motion obtained at some level of theory as a guide in identifying the important contributions to the potential. This potential may subsequently be used in conjunction with more extensive wave function calculations. Other recent works, such as the neural network approaches of Manzhos and Carrington [12, 13], and the approach of Dawes et al. [14, 15] based on interpolating moving least squares are also worth to cite. We also refer to the literature for recent discussions on other issues relating to effective and accurate PES construction, such as fitting methods with advanced selection of evaluation points [16], permutational invariance [17], and more general coordinates [18, 19]. We also point out that Rauhut[5] discussed iterative approaches as means of error control and, in conclusion, with other clever tricks, reported huge computational savings.

The idea underlying the ADGA procedure to be discussed, namely the use of the density of the wave function for nuclear motion in the decision steps of the iterative algorithm, is worth to emphasize. In the expectation value of the potential, the potential is multiplied with the density in the integral, $\langle V \rangle = \int \rho(q) V(q)\, \mathrm{d}q$. The ADGA takes advantages of this in an iterative build up of the analytic form of the potential, emphasizing regions of space with large contributions to such average values, while less attention is given to regions with small values for the density times potential. Although the concept is explored here within our particular setup and directly aimed at an efficient calculation of potential energy and molecular property surfaces for vibrational structure calculations, the basic idea can conceivably be extended to other setups (including some of the before mentioned) and other final goals. The basic algorithm can be combined with other important techniques to further reduce the number of calculated points, such as the explicit use of the molecular symmetry and the complete neglect of some mode-combinations [5–7, 20]. Furthermore, the combination of different levels of electronic structure theory [5, 6, 7, 8, 15], as well as the use of various interpolation schemes, and derivatives [7, 8, 11, 15, 21] can boost the efficiency of the proposed method. The main features of the ADGA can be appreciated without these important tools. However, even if they are not considered explicitly here, they are currently under investigation in our group.

Since the vibrational wave functions control the dynamic generation of the PES through the density, so will the choice of the one-mode primitive basis used in the vibrational structure calculations. This aspect has been explored as well in the article. We have implemented a semi-local basis set constituted of distributed gaussian functions, augmenting the existing harmonic oscillator basis available in the MidasCpp [22] suite of vibrational structure programs. The use of distributed Gaussians has a long history [23, 24] and has been used many times in combination with vibrational wave function calculations. The basic aspects of the gaussian basis sets will be briefly described to fully define our calculations.

The article is organized as follows: in Sect. 2, we describe the algorithm used for the generation of the grids of evaluation points, while in Sect. 3 the basic features and implementations of the gaussian basis set for vibrational calculations are reported. The computational details are presented in Sect. 4. In Sect. 5, the procedure is applied to the calculation of anharmonic vibrational energies of

model potentials such as the Morse oscillator, a double-well potential model, as well as water, difluoromethane and pyrimidine. For water we use the accurate analytic representations of the BO potential from Partridge et al. [25] while the potential energy values for difluoromethane and pyrimidine are calculated at the density functional theory (DFT) level, by using the DALTON quantum-chemistry code [26] and GAMESS suite of codes [27], respectively. Results are compared to calculations based on a standard non-iterative grid approach. Our conclusions and perspectives are then summarized in the final Sect. 6.

2 Description of the method

In the following, the implementation of the ADGA approach is outlined. We start by defining the Hamiltonian form we use and the expansion of the fully coupled BO potential in terms of restricted mode-coupling contributions. The procedure adopted for obtaining an analytical representation of the potential, in a form suitable for the subsequent vibrational structure calculations, will be briefly summarized, followed by a short account of the parametrization used for the molecular vibrational wave function. The main features of the new implementation are fully documented in the final section.

2.1 The vibrational Hamiltonian and the potential energy operator

For a non-rotating molecular system, assuming the BO approximation, the Hamiltonian in terms of normal coordinates reads [8, 28]:

$$H = -\frac{1}{2}\sum_k \frac{\partial^2}{\partial q_k^2} + \frac{1}{2}\sum_{\alpha\beta} \pi_\alpha \mu_{\alpha\beta} \pi_\beta - \frac{1}{8}\sum_\alpha \mu_{\alpha\alpha} + V(\mathbf{q}). \quad (1)$$

In Eq. 1, π_α and $\mu_{\alpha\beta}$ are Cartesian components of the vibrational angular momentum and the inverse of the effective moment of inertia, respectively (for further details and a discussion of the implementation of the Watsonian in MidasCPP, we refer to [8]). $V(\mathbf{q})$ in Eq. 1 is the BO potential energy operator. The kth normal coordinate is denoted q_k and $\mathbf{q}$ denotes the vector of all normal coordinates. Often the terms involving the inverse effective moment of inertia are neglected, which leaves us with the following Hamiltonian:

$$H = -\frac{1}{2}\sum_k \frac{\partial^2}{\partial q_k^2} + V(\mathbf{q}). \quad (2)$$

Due to its large dimensionality, (the number of vibrational degrees of freedom, M), the computation of the potential energy term is by far the most time-consuming part in the construction of the Hamiltonian, and can generally only be accomplished in an approximate fashion. A widely used approach in studies of vibrational dynamics is to approximate the full potential, $V(\mathbf{q})$, in a hierarchical manner including potential energy terms of lower dimensionality up to a certain level [1–5]:

$$V^{(1)}, V^{(2)}, V^{(3)}, \ldots, V^{(M)}. \quad (3)$$

To this aim, a set of potential energy functions (PEFs) are defined which include the coupling among a subset n of the M coordinates:

$$\begin{aligned} V^{m_1} &= V(0,\ldots,0,q_{m_1},0,\ldots,0) \\ V^{m_1,m_2} &= V(0,\ldots,0,q_{m_1},0,\ldots.0,q_{m_2},0,\ldots,0) \end{aligned} \quad (4)$$

and so forth up to $V^{m_1,m_2,\ldots,m_M}$, the fully coupled potential, $V(\mathbf{q})$. For the sake of simplicity the set of modes (referred to as a mode combination, MC, hereafter) defining the particular PEF are collected in an n-dimensional vector $\mathbf{m}_n$, so that a n-dimensional PEF is denoted as $V^{\mathbf{m}_n}$. Note that $V^{\mathbf{m}_n}$ is symmetric with respect to the permutation of mode indices in $\mathbf{m}_n$. One typical way to calculate the PEFs, commonly known as a grid approach, is to compute the potential energy values on a set of grid-points and interpolate or fit functions accordingly.

The PEFs can be used to define the sequence of potential energy terms (see Eq. 3) that converge to the fully coupled potential. However by summing over all MCs over-countings are introduced since each PEF includes all the lower-dimensional PEFs corresponding to the set of $\mathbf{m}_{n'} \subset \mathbf{m}_n$. Following the general formalism of [2] we introduce potentials $\overline{V}^{\mathbf{m}_n}$ defined such that:

$$\overline{V}^{\mathbf{m_n}} = S^{\mathbf{m_n}} \sum_{l=1}^{n} (-1)^{n-l} \binom{n}{l} V^{\mathbf{m_l}} \quad (5)$$

where $S^{\mathbf{m_n}}$ is an operator that symmetrizes with respect to n and m indices. Note that by construction, $\overline{V}^{\mathbf{m}_n}$ vanishes if any of the n internal coordinates is zero [2]:

$$\bar{V}^{\mathbf{m}_n}(\ldots, q_i = 0, \ldots) = 0. \quad (6)$$

Thus the following expression defines one hierarchy of potentials converging to the exact limit, for $n = M$, the total number of vibrational degrees of freedom:

$$V \approx V^{(n)} = \sum_{k=1}^{n} \sum_{\mathbf{m_k}} \overline{V}^{\mathbf{m_k}}. \quad (7)$$

Generally, one may include in the expansion only the MCs that are relevant for the representation of the potential:

$$V = \sum_{\mathbf{m} \in \mathbf{MCR}\{\mathbf{V}\}} \overline{V}^{\mathbf{m}}, \quad (8)$$

where MCR is a mode-combination range - the set of MCs we want to include in the potential.

2.2 The analytical representation of the potential and vibrational wave functions

In this work, we adopt a grid-based approach for potential sampling, fully documented in [8]. Very briefly, at a given mode combination level n, the $V^{\mathbf{m}_n}$ PEFs, are sampled in a set of grid points by means of ab-initio electronic structure calculations. From the $V^{\mathbf{m}_n}$ PEFs the $\overline{V}^{\mathbf{m}_n}$ terms of Eq. 5 are generated and an analytic form is obtained via a polynomial fitting procedure. The multidimensional linear least-squares fitting is realized in a direct product polynomial basis using frequency-scaled mass-weighted normal coordinates, $y_i = \sqrt{\omega_i} Q_i$, where ω_i is the classical harmonic frequency for the ith normal motion [8]. Through the fitting procedure, the vibrational Hamiltonian is cast to a sum over T products of single particle operators:

$$\hat{H} = \sum_{t=1}^{T} c_t \prod_{m=1}^{M} \hat{h}^{m,t}. \tag{9}$$

where c_t are expansion coefficients and each term in the summation is a product of M one-mode operators, $h^{m,t}$, of the form q_m^n or a derivative $(\mathrm{d}/\mathrm{d}q_m)^n$ (with $n \geq 0$ and $q_m^0 = (d/dq_m)^0 = 1$). The computational advantages of Hamiltonians of the form in Eq. 9 are important and well-established in time-dependent quantum dynamics [29] and in vibrational structure theory [30]. The fitting step is therefore a convenient step for the representation of the vibrational Hamiltonian in a sum over-product form which can be handled by the MidasCpp suite of vibrational programs. We use simple polynomial q_m^n but the procedure can easily be extended to other functions.

In MidasCpp, the calculation of vibrational energies can be done with a variety of approximate vibrational structure methods, such as VSCF [31–33], arbitrary order vibrational Møller-Plesset, (VMP) [4, 9, 34, 35, 36], and vibrational configuration interaction (VCI) [32, 33, 37, 38, 39, 40] parameterizations as well as newer (second quantization based [41]) approaches such as vibrational coupled-cluster (VCC) [40] and vibrational response theory [42, 43]. Within the ADGA approach, the calculation of molecular vibrational wave functions is required in two crucial steps. In the first step as a source for providing an effective density of the vibrational wave functions during the iterative construction of the PES. Once a converged PES has been obtained, a final calculation (VCI, VMP, VCC) gives the vibrational levels of interest. The first step is repeated many times during the construction of the potential, and high efficiency is therefore important and for simplicity we use VSCF. The VSCF wave function and densities rigorously separate into a product form, and furthermore VSCF is exact for uncoupled modes. We note that it has been illustrated very recently [44] that VMP, VCI, and VCC wave functions with two-mode coupling can be very efficiently implemented, so it is certainly a realistic future perspective to use also correlated wave functions in the initial step. However VSCF is anticipated to be good enough for the systems investigated in the present article, and for the purpose of illustrating the basic features of our procedure.

2.3 The ADGA for PES sampling

The previous implementation of the potential sampling driver [8] has been extended to allow for a dynamic generation of the grid of points on which each $V^{\mathbf{m}_n}$ term is evaluated. The new implementation should be therefore contrasted with the former where static grids were generated at the outset by specifying grid boundaries and the number of evaluation points. One shortcoming of the latter approach relates to the difficulties associated with a careful tailoring of the grids for the molecular system at hand. In the new procedure, at each iteration step, output from the vibrational structure program is used as a guide for the placement of the new set of evaluation points, and the accuracy of the sampling procedure is always monitored. The procedure uses one-mode density informations from VSCF runs for the determination of the grid boundaries on which each one-mode coupling potential term, $V^{\mathbf{m}_1}$, is to be sampled. When the adaptive procedure is converged for the one-mode coupling terms of the potential, the one-mode grid boundaries obtained define the extension of the grids where the higher mode-couplings ($n > 1$) terms of the potential will be sampled. In a multidimensional sampling, a direct product grid is employed. At each mode coupling level the adaptive procedure stops when specific convergence criteria are met, as outlined below. Then the iterative sampling is continued up to the maximum mode-coupling level used in the expansion of the potential in Eq. 3. Since one-mode grids are generated slightly differently compared to the multidimensional ones, we will describe the adaptive procedure separately for the monodimensional and multidimensional cases.

2.3.1 Adaptive generation of the monodimensional grids

The implementation is based on the partition of the configuration space into sub-sectors defined as the intervals between two adjacent sampling points. For each vibrational coordinate the procedure starts with the evaluation of the potential at the boundaries of the input grid (see Fig. 1a) as in our static grid setup [8]. The initial boundaries for the monodimensional grids can be specified by the user through input of a non-negative integer v in such a manner that the cut-off in the ith normal coordinate axis

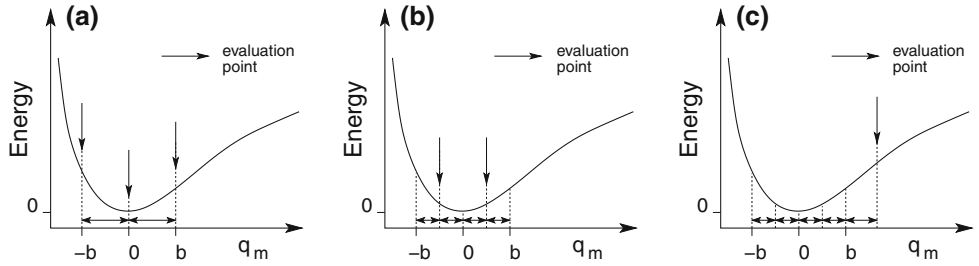

Fig. 1 Graphical representation of the adaptive procedure for the determination of the monodimensional PESs. **a** Iteration 0: two evaluation points (placed at the position of the grid boundaries specified in the input) are considered, together with the equilibrium geometry ($q_m = 0$), defining thus two sub-intervals. **b** Iteration 1: evaluation points are placed in the middle points of the sectors defined in **a**. **c** Extension of the grid outside the initial boundaries

corresponds to the classical turning point, $x_{i,\mathrm{TP}}$, for the harmonic motion in that direction

$$x_{i,\mathrm{TP}} = \pm\sqrt{\frac{2\hbar}{\omega}\left(v+\frac{1}{2}\right)}. \tag{10}$$

Alternatively, the initial grid boundaries can be explicitly given in the input by the user.

An analytical representation of the $V^{\mathbf{m}_1}$ terms is then provided via a polynomial fitting. The potential is then used in a VSCF calculation for the ground vibrational state. The VSCF modals are then read back by the program and a VSCF mean vibrational density for mode q_m is constructed:

$$\rho_{av}(q_m) = \frac{\sum_{i^m}^{N}\left|\phi_{i^m}^{m}(q_m)\right|^2}{N} \tag{11}$$

In Eq. 11, N is the maximum number of modals included in the evaluation of the average density and is input by the user; $\phi_{i^m}^{m}(q_m)$ are the set of VSCF modals for mode m. By construction, the average density integrates to one over the entire domain, a feature used later in the optimization of the grid boundaries. One can alternatively construct a maximum density for the given number N of modals: $\rho_{\max}(q_m) = \max_{[1,2,\ldots,N^m]}\left|\phi_{i^m}^{m}(q_m)\right|^2$. Nevertheless, for sake of simplicity, we will exclusively use the mean density in the remainder of this article.

Before the first iteration is started, testing points are placed in the middle of each sub-interval, and given the density information from the vibrational structure calculation and the analytical representation of the potential, the program evaluates the following integral

$$\int \rho(q_m) V^{\mathbf{m}_1}(q_m)\,\mathrm{d}q_m \tag{12}$$

for each of the four sub-intervals. Integration is performed with a standard Gauss–Legendre (GL) quadrature scheme and the integral values are stored.

The first iteration starts with the explicit evaluation of the potential for each of the newly added testing points by means of electronic structure calculations (see Fig. 1b). An updated analytical representation of the $V^{\mathbf{m}_1}$ is then obtained via a polynomial fitting, and a new mean vibrational density is computed in a ground-state VSCF calculation. New values of the integral $\rho V^{\mathbf{m}_1}$ of Eq. 12 are then calculated for each of the four sub-intervals defined in the preceding iteration, and compared with the stored values.

A further testing point is added at the middle point of the intervals for which the condition

$$\frac{\int (\rho_{\mathrm{new}} V_{\mathrm{new}}^{\mathbf{m}_1})\,\mathrm{d}q_m - \int (\rho_{\mathrm{old}} V_{\mathrm{old}}^{\mathbf{m}_1})\,\mathrm{d}q_m}{\int (\rho_{\mathrm{new}} V_{\mathrm{new}}^{\mathbf{m}_1})\,\mathrm{d}q_m} < \epsilon_{\mathrm{rel}}. \tag{13}$$

is not fulfilled. In other words, an interval where the relative variation of the integral value of $\rho V^{\mathbf{m}_1}$ between two iterations is found larger than a specified threshold (usually of the order of 1.0%) is further subdivided with the addition of a new testing point.

Note that the quantity defined in Eq. 12 can be regarded as a measure of the energy contribution from the given interval for mode m. For this reason it was decided that intervals with a small contribution to the energy should not be further subdivided, regardless of the relative error computed by means of Eq. 13. In particular, a sector is not further divided if the condition

$$\int (\rho_{\mathrm{new}} V_{\mathrm{new}}^{\mathbf{m}_1})\,\mathrm{d}q_m - \int (\rho_{\mathrm{old}} V_{\mathrm{old}}^{\mathbf{m}_1})\,\mathrm{d}q_m < \epsilon_{\mathrm{abs}} \wedge \int (\rho_{\mathrm{new}} V_{\mathrm{new}}^{\mathbf{m}_1})\,\mathrm{d}q_m < \epsilon_{\mathrm{abs}} \tag{14}$$

is fulfilled, with ϵ_{abs} being usually of the order of 10^{-6}.

The algorithm is therefore designed to require the calculation of additional evaluation points (i.e. single point energy calculations) in regions of the grid domain where sensible variations of the integral in Eq. 12 are observed between two successive iterations.

The density information is also used as a guide for the determination of the extension of the monodimensional grids. The number of modals N included in the calculation of the mean density of Eq. 11 for each vibrational mode relates to the number of vibrational states that needs to be

accurately calculated. Therefore, the number of vibrational states of interest controls the size of the monodimensional grid and determines the extension of the configurational space of the PES explored in the whole iterative procedure. This feature is highly desirable for an automatic procedure that constructs a semi-global representation of a potential energy surface since it ensures that the vibrational wave functions are localized in the configurational space explored with the evaluation points, i.e. a variational wave function does not fall into artificial "holes" of the PES, see also later. Moreover, since the potential is sampled only in the regions needed for accuracy of the requested states, evaluation points are not placed in regions where the solution of electronic problem is irrelevant and perhaps more difficult.

The procedure for the automatic extension of the grid boundaries relies on the fact that the integral, over the full configurational space, of the mean density, Eq. 11, is equal to one. The program checks that the integral value of the mean density summed over the sub-sectors approaches 1 according to:

$$\sum_i \int_i \rho_{av}(q_m)\,\mathrm{d}q_m > 1 - \epsilon_\rho \tag{15}$$

where i defines the ith sub-sector and ϵ_ρ is the fraction of density allowed to be outside the grid domain (in our tests ϵ_ρ ranges between 10^{-2} and 10^{-4}). If the condition of Eq. 15 is not fulfilled, the program extends the grid (see Fig. 1c).[1]

The adaptive procedure stops when all the convergency criteria of Eqs. 13–15 are met for each of the sub-intervals generated in the current iteration.

2.3.2 *Adaptive generation of the multidimensional grids*

The adaptive construction of the multidimensional grids is a straightforward generalization of the procedure outlined above. However, some distinctions are needed. First, the multidimensional grids are not allowed to extend beyond the boundaries defined by the corresponding monodimensional ones. Moreover, the multimode densities needed in the convergency checks (the convergency criteria are multidimensional generalizations of Eqs. 13 and 14) are taken for simplicity to be a direct product of the converged one-mode densities as:

$$\rho(q_{m_1}, q_{m_2}, \ldots, q_{m_n}) = \rho(q_{m_1}) \times \rho(q_{m_2}) \times \cdots \times \rho(q_{m_n}) \tag{16}$$

Such a product form holds for one state described by a VSCF wave function, but does not hold exactly for densities averaged over several states and for correlated wave functions. Clearly, the use of correlated densities could be considered.

In the current implementation, the procedure converges hierarchically up to a user specified maximum mode-combination level, i.e. monodimensional contributions are converged before the bi-dimensional and so on. This allows to use the converged analytic representation of the $V^{\mathbf{m}_{n'}}$ during the construction of the $\bar{V}^{\mathbf{m}_n}$ when the mode-combination $\mathbf{m}_{n'}$ is a subset of $\mathbf{m}_n$. This feature has been implemented as an option in the program.

It is worth to note that a specific grid is adaptively constructed for each of the multidimensional MCs, i.e. the multidimensional grids are not simply obtained as direct product of the corresponding monodimensional grids. A graphical representation of the procedure for the adaptive construction of bi-dimensional grids is reported in Fig. 2 with a detailed explanation in the caption.

Since the bi-dimensional and higher dimensional surfaces are known to be zero along axis/planes where any coordinate is zero (see Eq. 6) and, at least, can be assumed to be small in some domains, one may conceivably choose a zero function as the zero-th iteration estimate of the real surface. If the function values are indeed essentially zero for a particular mode-combination in the first iteration, further iterations are avoided for that specific cut of the PES. This illustrates clearly what is achieved and not

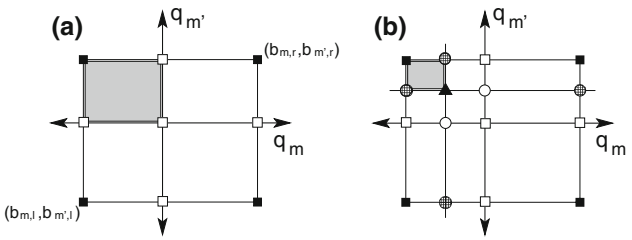

Fig. 2 Graphical representation of the adaptive procedure for the determination of the bi-dimensional PESs. **a** Four evaluation points at the boundaries of the corresponding monodimensional grids are considered in the first iteration (*black filled squares*). Together with the points already calculated for the monodimensional surfaces (*white filled square*), the available points define four sub-sectors (*one highlighted*). **b** Example of subdivision of a sector in the bi-dimensional procedure: the new evaluation point (*black filled triangle*) is placed in the middle of the sector. Due to the direct product representation, extra evaluation points have to be computed (*circles*) and the bi-dimensional surface is eventually partitioned into nine sub-sectors (*one highlighted*).

[1] In relation with the check of Eq. 15, the extension is set to 1/8 of the initial spanned space. The program evaluates the amount of density recoverable in a specific direction by assuming a linear decay of the density outside the grid boundaries; if the estimated residual density is larger than half of ϵ_ρ, the extension is trigged on. If the condition of Eq. 15 is not fulfilled, but none of the contributions in either direction is found large enough to trigger on the extension, i.e. due to an erroneous estimate of the density recoverable, the grid is extended in both directions with 1/16 of the initial grid domain.

achieved by the adaptive approach: some points need to be calculated for each MC, but only few points are needed if the particular mode combination turns out to be unimportant, while more points will be calculated in case of a strong and complicated coupling. The sampling of a bi-dimensional surface will thus require a minimum of four points. Getting rid of the MCs without ever calculating this minimal number of points is considered to be another, very interesting, activity of research, but is outside the scope of this article.

As a final remark, the ADGA can be used to converge not only the analytic representation of the PES but also the analytic representation of any molecular property surfaces since in all technical aspect molecular properties are handled in the same way as the energy. For tests and performance of the ADGA with respect to the construction of molecular property surfaces we refer to a later work.

2.3.3 Technical details

The calculation of the grid points in the construction of potential energy surfaces is an ideal task to be tackled by parallel computing. Once the individual grid points are defined, their calculation may be addressed independently: this goes under the name of "embarrassingly parallel problem". At each iteration, the ADGA defines a list of grid points that are to be computed. In our computational setup, these points are considered in parallel thanks to the combined action of shell scripts and the queue system. In this way, a very efficient parallelization of the grid calculations (by far the most time-demanding part of the procedure) was achieved. This is, in a certain sense, a proof of principle for a genuine master–slave parallelization.

Symmetry considerations can give a substantial speed up in the construction of PES for small and medium-sized molecules. In the current implementation of the ADGA the use of the symmetry properties of the PES are not exploited to reduce the number of single point calculations, though symmetry is used for the individual points. The implementation of a strategy to exploit symmetry seems straightforward into the current algorithm, in fact, when a specific grid point is requested, the algorithm can check if symmetry related points have been already requested. If the grid point is recognized as symmetrically related to an other one, it will be not explicitly calculated but its property values will be duplicated from its symmetry related point. The use of symmetry to reduce the number of explicit calculation is currently under development in our group and this aspect will be considered in future applications. The neglect of symmetry does not affect the forthcoming discussion on the computational efficiency and accuracy of the ADGA approach compared to the static-grid implementation [8] since computational savings due to symmetry considerations represent a common factor in speed-up for both the static and the ADGA approaches compared in this work. Note also that our aim in this article is to investigate the usefulness of density information in guiding molecular PESs construction, rather than providing the most efficient implementation of the new concept introduced.

3 Localized Gaussian basis

Although harmonic oscillator wave functions provide a good zero order description for the lowest vibrational states of rigid molecules, their global nature requires the sampling of large portions of the molecular PES when functions of high vibrational quantum number are used for expanding the modals. An incorrect sampling of the outer portions of the potential surface can cause instabilities in the vibrational calculation due to contamination of low-lying vibrational states from higher excited ones. We therefore implemented and tested the performance of a localized basis set of distributed Gaussians functions for expanding the modals, within the procedure presented in the article.

Note that in the step determining the optimal extension of the sampling grid, the adaptive procedures requires that the vibrational wave function probes the space immediately outside the boundaries of the grid. Therefore, the support for the localized basis is chosen to go somewhat beyond the boundaries.[2]

We consider a set of distributed Gaussians, of the independent variable x:

$$G_i(x) = \left(\frac{2\zeta_i}{\pi}\right)^{\frac{1}{4}} \exp(-\zeta_i(x - Z_i)^2) \tag{17}$$

Each individual Gaussian indexed by i is centered at $Z_i \in [a; b]$ and has exponents ζ_i. Thus, the set of Gaussians are specified by the set of Z_i and ζ_i. A particular simple Gaussian basis set is provided by the case of an equidistant set of points with identical exponents, so the set of Gaussians can be specified by four numbers: the upper and lower bounds of the interval where the Gaussians are distributed, namely a and b, the total number of Gaussians N, and a common exponent ζ. We can define a basis set density as:

$$\rho = \frac{N}{b - a}. \tag{18}$$

[2] It was found adequate to enforce the support for the localized basis set to be 50% larger (with a maximum allow 20 a.u.). than the space spanned by the evaluation points.

The individual Gaussian's centers are given by:

$$Z_i = a + \frac{b-a}{N-1} \times i \tag{19}$$

where $i = 0, 1, \ldots, N-1$. In practice there has to be a balance between the density of the basis set, Eq. 18, and the exponents ζ, in order to avoid linear dependencies in the basis. In our implementation the exponents are generated as suggested by Hamilton et al. [23] $\zeta_i = 4c^2/(Z_{i+1} - Z_{i-1})^2$ for the Gaussians centered in the interior points and $\zeta_1 = c^2/(Z_2 - Z_1)^2$ and $\zeta_N = c^2/(Z_N - Z_{N-1})^2$ for the Gaussians centered on the boundaries of the interval $[a,b]$. The c parameter is the generator of the Gaussian's exponents, and sensible values [23] range between 0.5 and 1.1. For a polyatomic molecule with M vibrational modes, each mode has its own basis set. The program can generate a set of localized Gaussians in several ways. One way is to define explicitly the interval boundaries, the c parameter and the number of basis functions or the density of the Gaussians, Eq. 18, for each of the vibrational coordinates. Alternatively the program accepts in input a vector of M integers specifying the number of basis functions or M real numbers for the basis set densities. The interval boundaries for each of the M vibrational coordinates can be read from file to correspond to those defining the grid extensions in the PES sampling procedure or can be generated through input of a non-negative integer v to correspond to the classical turning point for the harmonic motion in that direction. Gaussians are then evenly distributed between the boundaries.

The calculation of integrals for polynomial expanded potentials can for the Gaussians functions be done analytically, and are detailed in the Appendix for completeness.

4 Computational details

One-mode densities used in the adaptive procedure are obtained after VSCF calculations on the ground vibrational state. Since a VSCF wave function is exact for a system of a single mode of vibration, VSCF calculations have been performed in the monodimensional test cases. For water we use Full-VCI (FVCI) parameterization of the wave function, whereas a VCI[gs,4] wave function is used for difluoromethane to compare with previous work. In a VCI[gs,4] calculation [40] all configurations, which have up to four excited modes relative to the VSCF ground-state are included. In the case of water, the lowest 15 VSCF modals per mode were retained in the FVCI calculations, whereas the 6 lowest modals per mode were used in the VCI[gs,4] on difluoromethane. For pyrimidine, the second order vibrational Møller–Plesset (VMP2) Perturbation Theory was adopted, with the inclusion in the correlated calculation of the lowest eight VSCF modals per mode.

In addition to HO wave functions, the use of localized Gaussian basis sets has been explored for all the cases presented. The HO basis is specified by a maximum vibrational quantum number and we report some tests of the sensitivity on this number for each of the molecules investigated in the article. The exponents for the HO basis are obtained from the quadratic term of the potential. The Gaussian basis sets were generated given a density value, Eq. 18, for the functions equal to 0.8. The exponents of the Gaussians were generated according to Hamilton et al. [23] (see Sect. 3 above) with $c = 1$. These settings may easily generate around 100 basis functions per mode, but this is not a problem in the very fast VSCF calculations.

While a spline interpolation technique *prior* to the fitting procedure was found useful to avoid spurious oscillations in the fitted surfaces within the static grid framework, it is not used in the ADGA as the new algorithm would automatically detect and correct artifacts in the fitting. In fact, the ADGA proved to be more stable and economic without spline interpolation. The maximum polynomial degree used for the fitting of the monodimensional and multidimensional surfaces is 12, except in the double-well problem where it was set equal to 14. We checked that the analytic representation of the PES is converged with respect to the maximum order of the fitting polynomials (variations in the maximum order of the polynomial gave variations of less than 0.01 cm^{-1} in the vibrational energies). In the early cycles of the iterative procedure few evaluation points are available and the maximum degrees of the fitting polynomials are reduced to $n - 1$ with n being the number of evaluation points.

5 Sample applications

5.1 The Morse potential

A standard model for approximating the potential of a diatomic molecule is given by the Morse function:

$$V(q) = D_e\left(1 - e^{-a(q-q_e)}\right)^2 \tag{20}$$

where $V(q)$ is the potential energy, D_e is the dissociation energy and q_e the equilibrium distance, while a is a parameter controlling the width of the potential well. The parameters chosen in the simulation are $D_e = 0.1026$, $r_e = 1.9972$ and $a = 0.732$ in atomic units, and the curve approximates the potential of the ground state of the H_2^+ molecule. The advantage of the use of the Morse potential for benchmark is due to the fact that the exact solutions for the eigenstates are available analytically and the corresponding results are reported in Table 1.

Table 1 Vibrational transition energies in the Morse potential as a function of the number of states and the density threshold, ε_ρ, used in the adaptive construction of the PES

Exact	Adaptive approach[a]								
		Two states[b]			Four states[c]			Six states[d]		
	ε_ρ	10^{-1}	10^{-2}	10^{-3}	10^{-1}	10^{-2}	10^{-3}	10^{-1}	10^{-2}	10^{-3}
	$\%\rho$	93.17	99.63	99.94	91.05	99.77	99.98	90.66	99.64	99.97
	# calc.	18	20	23	15	19	21	19	22	24
0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2,273.75		0.01	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01
4,419.41		*0.15*	*−0.01*	*−0.01*	−0.02	−0.01	−0.01	−0.01	−0.01	−0.01
6,436.97		*1.30*	*−0.06*	*−0.04*	−0.12	−0.01	−0.01	−0.01	−0.01	−0.01
8,326.44		*7.18*	*−0.45*	*−0.29*	*−0.88*	*−0.01*	*−0.01*	0.01	−0.01	−0.01
10,087.82		*29.84*	*−2.60*	*−1.81*	*−5.09*	*0.19*	*0.13*	0.33	0.11	0.11

The differences between the analytic and VSCF values are reported (cm^{-1}). A basis set of 30 HO functions is used in the VSCF calculations. Numbers in italic refer to energy levels of the states not included in the evaluation of the mean density

[a] The thresholds for the convergence of the sub-sectors are $\varepsilon_{rel} = 10^{-2}$ and $\varepsilon_{abs} = 10^{-6}$

[b] The states corresponding to $v = 0, 1$ are used in the computation of the mean density, and the classical harmonic turning point of the state $v = 0$ sets the initial boundaries

[c] The states corresponding to $v = 0{-}3$ are used in the computation of the mean density, and the classical harmonic turning point of the state $v = 2$ sets the initial boundaries

[d] The states corresponding to $v = 0{-}5$ are used in the computation of the mean density, and the classical harmonic turning point of the state $v = 4$ sets the initial boundaries

In Table 1, the ability of the adaptive procedure to construct an accurate representation of the potential as a function of the number of the states addressed and the amount of mean density allowed outside the grid, ϵ_ρ, is tested. The fraction of the mean density inside the grid boundaries is consistent with the corresponding thresholds, as can be seen from the results in the first row of Table 1. Moreover, the analysis of the data reveals that the agreement with the analytic results is quite good even when the less tight threshold ($\epsilon_\rho = 10^{-1}$) is used (with this set up, the ADGA uses 4 iterations to converge). The agreement is improved when ϵ_ρ is reduced by an order of magnitude, requiring only a few more evaluation points. A further decrease of ϵ_ρ does not bring any improvement (but it requires a few more evaluation point as well as 1–2 extra iterations), confirming that the results are converged for the particular calculation settings. The energy values pertaining to the states not directly addressed in the calculations are emphasized in italic in Table 1. It can be observed that fair agreement is obtained for the lower states when the density threshold is low. The analysis of the tabulated results suggests that a good compromise between accuracy and computational efficiency is achieved with a threshold value of $\epsilon_\rho = 10^{-3}$.

Concerning the choice of the initial grid boundaries, the use of the classical harmonic turning points corresponding to the nth state, n being the highest state of interest (or a state 1–2 units lower than the highest state of interest) is found to be a good compromise.

We report an analysis of the convergence of the vibrational energies with respect to the absolute and relative thresholds (ϵ_{abs} and ϵ_{rel}) controlling the subdivision of the sub-sectors in Table 2. Moreover, the influence of the primitive basis set used in the VSCF calculation is explored. In all the cases, the AGDA converged within 8 iterations and the number of evaluation point ranges between 23 and 30 for all the tests but the final one (vide infra). Clearly, the energy levels of the Morse potential are well reproduced up to the state $v = 5$ by using 30 HO functions and the less tight convergence criteria ($\epsilon_{abs} = 10^{-5}, \epsilon_{rel} = 10^{-1}$). On the other hand, the vibrational energies for the 0→6 and 0→7 transitions are poorly reproduced. These vibrational energies do not improve after the use of tighter thresholds suggesting that these errors are not dependent on the adaptive procedure itself. This is confirmed by the second set of calculations: in this case, by simply increasing the number of primitive HO basis functions up to 50, an excellent accuracy is achieved for all of the requested transition energies.

Given that the procedure here described is meant to be a "black box" tool for the generation of highly accurate PES, the behavior observed in the first two sets of ADGA calculations should be avoided. We therefore explored the use of localized Gaussian functions as primitive basis for the vibrational calculations. The centers of the Gaussian functions are equidistantly placed and the density of the basis is kept constant in the adaptive procedure. This means that the number of basis functions increases as the

Table 2 Vibrational transition energies of the Morse potential as a function of the ε_{abs} and ε_{rel} thresholds controlling the subdivision of the sub-sectors in the adaptive construction of the PES

Exact	Adaptive approach[a]						
	Basis	30 HO		50 HO		Gaussians	
	ε_{abs}	10^{-5}	10^{-6}	10^{-5}	10^{-6}	10^{-5}	10^{-6}
	ε_{rel}	10^{-1}	10^{-2}	10^{-1}	10^{-2}	10^{-1}	10^{-2}
	# calc.	23	25	23	30	25	50
0.00		0.00	0.00	0.00	0.00	0.00	0.00
2,273.75		−0.01	−0.01	−0.01	−0.01	0.00	0.00
4,419.41		−0.01	−0.01	−0.01	−0.01	0.00	0.00
6,436.97		−0.01	−0.01	−0.01	−0.01	0.00	0.00
8,326.44		−0.01	−0.01	−0.02	−0.02	0.00	0.00
10,087.82		0.11	0.11	−0.02	−0.02	0.00	0.00
11,721.10		3.52	3.52	−0.02	−0.02	0.00	0.00
13,226.29		42.65	42.65	0.02	0.02	0.02	0.01

The differences between the analytic and VSCF values are reported (cm^{-1}). The ε_ρ parameter controlling the extension of the grid boundaries was set to 10^{-3}

[a] The states corresponding to $v = 0, \ldots, 7$ are used in the computation of the mean density, and the classical harmonic turning point of the state $v = 6$ sets the initial boundaries

configurational space spanned in the adaptive procedure increases. The results obtained with the Gaussians basis set are found extremely accurate even with the less tight set of thresholds. This confirms the flexibility of the localized basis sets in following the gradual expansion of the sampling grid, while saturating the area of interest. A drawback of the use of the localized basis set is the greater number of evaluation points required with the tighter thresholds. In particular, the number of calculations doubles, because the localized basis sets confine the vibrational wave function in space and therefore the density. Whenever the grid expands, the wave function relaxes in space, and this provides a "global" change in the ρV quantity that may trigger the subdivision of several of the inner sub-sectors. Nevertheless, it should be considered that, under normal circumstances, the number of calculations required in the construction of the monodimensional surfaces is never a bottleneck, and on the other side, the optimization of the boundaries achieved by using localized basis helps to minimize the number of evaluation points on the multidimensional surfaces which is much more critical. In order to overcome this drawback of the localized basis in the construction of the monodimensional surfaces it would be possible to perform a pre-screening of the potential adopting an inexpensive method of calculation. The boundaries obtained could be used in the subsequent adaptive procedure as starting guess, this is expected to improve the convergence of the PES with respect to the number of evaluation points and iteration cycles.

In the case of monodimensional surfaces a standard static grid approach can certainly be optimized by human influence to use fewer points than the adaptive procedure while achieving the same accuracy. Nevertheless, this kind of optimization requires an *a priori* knowledge of the surface and implies significant trial-and-error. The ADGA is fully automatic and does not require any prior knowledge of the potential surface.

5.2 The symmetric double-well potential

As a second monodimensional model, we consider a symmetric double-well potential of the form of a harmonic function perturbed with a Gaussian-type barrier [45]:

$$V(q) = \frac{1}{2}\omega^2 q^2 + Ae^{-\alpha q^2}. \tag{21}$$

In our test, we use the parametrization given by Lin et al. [46] for modeling the inversion mode of ammonia and their calculated vibrational transition energies were taken as a reference for benchmark.

From the point of view of expanding a potential around a stationary point, the double-well problem can be approached in two different ways: either starting from the local maximum (transition state of the inversion barrier) or from one of the two minima. While, the two minima are treated equivalently in the first case this is not true in the second case.

Table 3 reports the results obtained with the adaptive procedure when the maximum of the barrier is chosen as the reference point. For all the tests considered, the adaptive procedure converged within five iterations, and the vibrational excitation energies obtained are found in excellent agreement with the reference data of Lin et al. [46] confirming the robustness of the ADGA. The agreement with the reference data is found to be independent on the basis set adopted in the vibrational calculation, and, as observed in the previous example, the use of the Gaussians basis may imply an increase in the number of evaluation points required.

Table 4 lists the results obtained when sampling the symmetric double-well potential starting from one of the two equivalent minima. With respect to the previous case, this situation is intrinsically more complicated because the procedure has to cross the barrier searching for the second minimum, as a result, the adaptive procedure require more iteration to converge (10–14).

As shown in Fig. 3, the procedure is capable of reconstructing the shape of the PES step by step. The results in Table 4 highlight that an overall good accuracy is achieved even when employing less tight convergency thresholds. The use of a set of tighter thresholds does not improve the quality of the results, but it confirms that the adaptive

Table 3 Vibrational transition energies (in cm^{-1}) for the symmetric double-well potential

Reference	Adaptive approach[a]					
	Basis	30 HO		50 HO		Gaussians	
	ε_{abs}	10^{-5}	10^{-7}	10^{-5}	10^{-7}	10^{-5}	10^{-7}
	ε_{rel}	10^{-1}	10^{-3}	10^{-1}	10^{-3}	10^{-1}	10^{-3}
	# calc.	19	29	43	45	19	27
0.00		0.00	0.00	0.00	0.00	0.00	0.00
0.83		0.00	0.00	0.00	0.00	0.00	0.00
932.40		0.06	0.05	0.05	0.05	0.06	0.05
967.80		−0.04	−0.05	−0.05	−0.05	−0.04	−0.05
1,603.00		−0.10	−0.11	−0.11	−0.11	−0.10	−0.11
1,882.80		0.00	0.01	−0.01	−0.01	0.01	0.00
2,385.00		-0.06	−0.06	−0.07	−0.07	−0.05	−0.06
2,891.70		0.02	0.01	0.01	0.01	0.03	0.02
3,453.20		0.00	−0.01	−0.01	−0.01	0.01	0.00
4,047.40		0.02	0.02	0.02	0.02	0.03	0.03

The VSCF results as a function of the thresholds and basis sets in the adaptive construction of the PES are given relative to the published data taken from [48]. The potential is expanded around the local maximum

[a] The lowest ten states are used in the computation of the mean density. The ε_ρ parameter controlling the extension of the grid boundaries was set to 10^{-3}

Table 4 Vibrational transition energies (in cm^{-1}) for the symmetric double-well potential

Reference	Adaptive Approach[a]					
	Basis	50 HO		80 HO		Gaussians	
	ε_{abs}	10^{-5}	10^{-7}	10^{-5}	10^{-7}	10^{-5}	10^{-7}
	ε_{rel}	10^{-1}	10^{-3}	10^{-1}	10^{-3}	10^{-1}	10^{-3}
	# calc.	50	320	148	1,251	56	320
0.00		0.00	0.00	0.00	0.00	0.00	0.00
0.83		0.00	0.00	0.00	0.00	0.00	0.00
932.40		0.06	0.06	0.03	0.01	0.05	0.04
967.80		−0.03	−0.03	−0.07	−0.09	−0.05	−0.06
1,603.00		−0.05	−0.05	−0.17	−0.16	−0.13	−0.13
1,882.80		0.23	0.22	−0.05	−0.06	−0.02	−0.02
2,385.00		0.61	0.61	−0.13	−0.12	−0.08	−0.07
2,891.70		1.96	1.95	−0.04	−0.04	0.00	0.01
3,453.20		5.07	5.05	−0.07	−0.05	−0.02	0.00
4,047.40		12.24	12.20	−0.04	−0.02	0.01	0.02

The VSCF results as a function of the thresholds and basis sets in the adaptive construction of the PES are given relative to the published data taken from [48]. The potential is expanded around one of the equivalent minima

[a] The lowest ten states are used in the computation of the mean density. The ε_ρ parameter controlling the extension of the grid boundaries was set to 10^{-3}

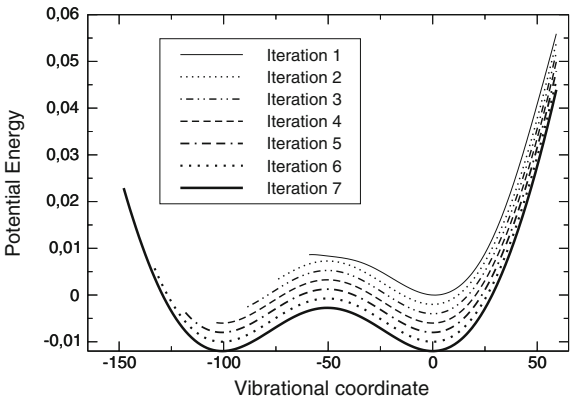

Fig. 3 The adaptive procedure applied to the double-well potential: the potential is gradually expanded from the *rightmost* minimum to the *leftmost* minumum. Offsets in the energy scale are used for the sake of clarity.

procedure provides a PES of excellent quality and the discrepancy with the references are due to shortcomings of the basis set in the vibrational calculation, (see the "50 HO" columns in Table 4).

In the case of the double well potential, the side effect of the increase in the number of evaluation points when low thresholds are used is particularly pronounced. However, such effect was expected: at each iteration the global representation of the wave function changes significantly due to the gradual "discovery" of the second minimum and this, in conjunction with tighter thresholds, triggers the subdivision of a larger number of sectors near the starting minimum. A clever choice for the initial boundaries could reduce the number of evaluation points significantly, but our aim here is rather to demonstrate that the algorithm can, by brute force, cope with this difficult situation. While this can not generally be expected to carry over to the multidimensional case without additional problems, it does indicate a significant robustness of the procedure.

The double-well problem provides a clear example on the shortcomings of the use of a global primitive basis in the vibrational calculation: the symmetric HO wave functions are not suited for this highly asymmetric (if seen from one of the minimum) potential and a large vibrational quantum number has to be chosen to ensure that the basis functions cover the region far from the starting minimum in the direction of the second minimum. A lack of basis function and saturation was found to be the problem for the "50 HO" basis set, 30 more functions had to be included in the basis set to properly reproduce the reference values. The use of HO functions with high vibrational quantum number implies that the space outside the "controlled" region is probed and therefore the wave function risks to collapse into an artificial hole. The algorithm deals with these wild oscillations by enlarging progressively the configurational domain at the expense of an unnecessarily

Table 5 FVCI transition energies for the fundamental and first overtones vibrations of water for both a static and an adaptive sampling of the full PES

(v_1, v_2, v_3)[a]	Reference		Standard approach[b]		Adaptive approach[c]			
		Basis	20 HO		20 HO		Gaussians	
		Thresh	Grid A[d]	Grid B[e]	Loose[f]	Tight[g]	Loose[f]	Tight[g]
		# calc.	1,377	3,601	456	884	462	890
(1, 0, 0)	3,657.04		−0.43	0.05	−0.14	0.00	−0.13	0.01
(0, 1, 0)	1,594.76		−0.04	0.28	−0.44	−0.28	−0.43	−0.27
(0, 0, 1)	3,755.96		−0.74	−0.23	−0.77	−0.28	−0.76	−0.27
(2, 0, 0)	7,201.55		−0.73	0.06	−0.25	−0.10	−0.23	−0.08
(0, 2, 0)	3,151.63		−0.92	−1.16	−1.26	−1.09	−1.25	−1.08
(0, 0, 2)	7,445.12		−1.39	−0.42	−1.80	−0.57	−1.78	−0.55
	MAD		0.72	0.37	0.78	0.39	0.77	0.38
	MaxAD		1.39	1.16	1.80	1.09	1.78	1.08

Vibrational energies (in cm^{-1}) are given as the difference with respect to the results of [25], mean absolute deviations (MAD) and the maximum absolute deviations (MaxAD) are reported for each calculation

[a] v_1: symmetric stretch, A_1; v_2: bending, A_1; v_3: anti-symmetric stretch, B_1

[b] Static grid approach [8], the classical harmonic turning point of the state $v = 4$ sets the boundaries for each mode

[c] The lowest three VSCF modals for each vibrational mode are used in the construction of the average density. The ε_ρ parameter was set to 10^{-3}

[d] The static grids were constructed according to the scheme $(32_1)(16_1)(8_1)$, with a Fine Grid 6 times more dense (see [8] for details)

[e] The static grids were constructed according to the scheme $(48_1)(24_1)(12_1)$, with a Fine Grid 6 times more dense (see [8] for details)

[f] The 1D surfaces were converged with $\varepsilon_{rel} = 5 \times 10^{-3}$ and $\varepsilon_{abs} = 5 \times 10^{-7}$, the 2D surfaces were converged with $\varepsilon_{rel} = 5 \times 10^{-2}$ and $\varepsilon_{abs} = 5 \times 10^{-6}$ and the 3D surface was converged with $\varepsilon_{rel} = 3 \times 10^{-1}$ and $\varepsilon_{abs} = 3 \times 10^{-5}$

[g] The 1D surfaces were converged with $\varepsilon_{rel} = 5 \times 10^{-3}$ and $\varepsilon_{abs} = 5 \times 10^{-7}$, the 2D surfaces were converged with $\varepsilon_{rel} = 2 \times 10^{-2}$ and $\varepsilon_{abs} = 2 \times 10^{-6}$ and the 3D surface was converged with $\varepsilon_{rel} = 1 \times 10^{-1}$ and $\varepsilon_{abs} = 1 \times 10^{-5}$

large number of evaluation points. As a matter of fact, the converged grid boundaries do not reflect the spatial extension of the wave functions but of the more diffuse primitive basis functions. On the other hand, by construction, the localized basis set is confined in the region where the analytical PES is well conditioned. The saturation in the space of interest is automatically achieved in accordance with the "black box" philosophy motivating this work.

5.3 The water molecule

The water molecule has long been used as a benchmark system to test the accuracy and the performances of procedures to calculate vibrational energies of polyatomic molecules [3, 8, 47]. In this work, we use the spectroscopically accurate potential of Partridge and Schwenke [25]. In all the test calculations the full representation of the potential has been used in the Watsonian (Eq. 1). The vibrational excitation energies to the fundamentals and to the first overtones were addressed.

Table 5 presents a comparison between the reference vibrational excitation energies [25] for the selected states of a non-rotating water molecule and those obtained with both the ADGA presented here and the static grid approach of [8]. Mean absolute deviations (MAD) and the maximum absolute deviations (MaxAD) are reported for each calculation. Two different grid settings were used in the non iterative approach while the adaptive procedure uses two different sets of thresholds (see footnotes to Table 5 for details).

In the standard grid approach, the quality of the results increases with the number of the evaluation points used in the potential sampling: the MADs are found to be 0.72 and 0.37 cm^{-1} for Grid A (1,377 evaluation points) and Grid B (3,601 evaluation points), respectively. It should be noticed that, even if the accuracy obtained with the standard grid approach is clearly satisfactory, the number of evaluation points needed is rather large. The results listed in Table 5 clearly demonstrate that the ADGA is computationally more efficient and the MAD of the calculations with the "loose" set of thresholds is comparable with the results obtained with Grid A, in spite of requiring only 1/3 of the evaluation points. With the "loose" set of thresholds, the adaptive procedure converges the 1D, 2D and 3D surfaces requiring about 60 (6 iterations), 340 (4 iterations) and 64 (2 iterations) energy evaluations, respectively. Moreover, if the "tight" set of thresholds is used, the number of

Table 6 Fundamental vibrational frequencies of difluoromethane (cm^{-1}) computed with ADGA, relative to the values computed with a standard grid

Mode	Standard[a]	ADGA[b]			
		Thresh	Loose[c]	Tight[d]	Tight[d,e]
	10,776	# calc.	1,786	4,023	6,519
ν_1 (A_1) CH_2 symm. stretch	2,945.2		−1.8	−0.6	−0.5
ν_2 (A_1) CH_2 scissor	1,498.0		−0.5	−0.1	0.2
ν_3 (A_1) CF_2 symm. stretch	1,111.9		−0.2	0.0	0.5
ν_4 (A_1) CF_2 bend	525.8		0.1	0.0	0.6
ν_5 (A_2) CH_2 twist	1,246.2		−7.5	−0.8	−1.0
ν_6 (B_1) CH_2 anti-symm. stretch	3,026.6		7.1	−0.7	−11.5
ν_7 (B_1) CH_2 rock	1,169.0		−4.2	0.9	1.1
ν_8 (B_2) CH_2 wag	1,428.1		−7.8	−0.9	−1.1
ν_9 (B_2) CF_2 anti-symm. stretch	1,076.5		0.1	−0.1	0.6
MAD			3.2	0.5	1.9
MaxAD			7.8	0.9	11.5

Mean absolute deviations (MAD) and the maximum absolute deviations (MaxAD) are reported. The contruction of the 3D PES is based on CamB3LYP/aug-cc-pVTZ single point calculations. Vibrational calculation: VCI[gs,4]

[a] Reference values are taken from [8]. Static grid approach: the classical harmonic turning point of the state $v = 10$ sets the boundaries for each mode. The grids were constructed according to the scheme $(48_1)(24_1)(12_1)$, with a Fine Grid 8 times more dense (see [8] for details)

[b] The lowest two VSCF modals for each vibrational mode are used in the construction of the average density. ε_ρ was set to 10^{-3}

[c] The 1D surfaces were converged with $\varepsilon_{rel} = 5 \times 10^{-3}$ and $\varepsilon_{abs} = 5 \times 10^{-7}$, the 2D surfaces were converged with $\varepsilon_{rel} = 1 \times 10^{-1}$ and $\varepsilon_{abs} = 1 \times 10^{-5}$ and the 3D surfaces were converged with $\varepsilon_{rel} = 3 \times 10^{-1}$ and $\varepsilon_{abs} = 3 \times 10^{-5}$

[d] The 1D surfaces were converged with $\varepsilon_{rel} = 5 \times 10^{-3}$ and $\varepsilon_{abs} = 5 \times 10^{-7}$, the 2D surfaces were converged with $\varepsilon_{rel} = 5 \times 10^{-2}$ and $\varepsilon_{abs} = 5 \times 10^{-6}$ and the 3D surfaces were converged with $\varepsilon_{rel} = 1 \times 10^{-1}$ and $\varepsilon_{abs} = 1 \times 10^{-5}$

[e] The PES includes 4D mode couplings, the 4D surfaces were converged with $\varepsilon_{rel} = 3 \times 10^{-1}$ and $\varepsilon_{abs} = 3 \times 10^{-5}$

evaluation points increases up to about 900 (still only 70% of the points required for Grid A) while the associated MAD equals the one for the Grid B calculation. When the "tight" set of thresholds is adopted, the 1D, 2D and 3D surfaces require about 60 (6 iterations), 496 (4 iterations) and 336 (2 iterations) potential energy evaluations respectively. The CPU time used in the ADGA (excluding the time required for the single point calculations) was about 100 s,[3] the VSCF calculations required about 50 s and while the rest was used in the steps of grid update and convergency checking.

Once again, we notice that water is such a well known and simple system that optimal grids could be designed manually. Nonetheless, the results suggests that the ADGA method is capable to provide a high efficient, general and black-box strategy.

5.4 Difluoromethane

We consider the fundamental vibrational frequencies of difluoromethane. Potential energy values have been computed by means of DFT calculations using the DALTON quantum-chemistry program [26] and the CAM-B3LYP exchange-correlation potential [48]. The computational details are the same as those in Ref. 8 and the results there presented were addressed.

In Table 6, the VCI[gs,4] values obtained with the PES constructed with standard grid are used as reference. In the adaptive construction of the PES we used different sets of thresholds ("loose" and "tight" in Table 6). The agreement between the different sets of computed vibrational frequencies is fair, but the highly automatized ADGA uses 1/6 of the evaluation points requested with the static grid approach. Whereas 10,776 single-point calculations are required by the grid-sampling algorithm of [8], the adaptive procedure is converged after only 1,786 single-point calculations (1D surfaces: 117 single points, 5 iterations; 2D surfaces: 845 single points, 3 iterations and 3D surfaces: 824 single points, 2 iterations). In spite of the great reduction in the number of single point calculations required, the MAD and MaxAD with respect to the reference values are found equal to 3.2 and 7.8 cm^{-1}. The use of a tighter set of thresholds improves the agreement and the differences with the static grid results are smaller than 1.0 cm^{-1}. Such accuracy is obtained with a consistently lower number of single-point

[3] On a Opteron 2.2 GHz processor, 2 GB memory RAM.

calculations (1D surfaces: 117 single points, 5 iterations; 2D surfaces: 1,418 single points, 4 iterations and 3D surfaces: 2,488 single points, 3 iterations); clearly the computational efficiency of the adaptive approach does not compromise the accuracy of the results. Concerning the time spent in the ADGA itself (excluding the time for the single point calculations) 15 CPU min were needed, see footnote3 among them the VSCF calculations required 3 min. The cost of the ADGA itself is negligible compared to the potential saving of the many electronic structure single points.

Due to the ability of the adaptive procedure to handle unimportant mode combinations with the explicit computation of a relative small amount of points, the inclusion of the four-mode couplings terms in the approximated PES is now feasible without an overwhelming number of evaluation points. In particular only 2 of the 126 four-mode couplings were not converged after the first iteration. The construction of the PES up to the four-mode coupling terms was carried out as an extension of the previously described "tight" three-mode PES, and the four-mode terms required two iterations and 2,496 additional evaluation points leading to a total number of 6,519 single point calculations. It is interesting to note that the number of additional points for the 4 mode terms is smaller than the total number of points required for one-, two-, and three-mode couplings.

The vibrational energies obtained with a VCI[gs,4] calculation using the four-mode potential are found in good agreement with the data obtained with the three-modes PES, except for the CH_2 anti-symmetric stretch. Clearly, a direct comparison between the values obtained with a 3D- and a 4D-PES is unfair, but it shows that at least for one of the fundamental frequencies the mode combination range was not converged. This appears consistent with the observation that the CH_2 anti-symmetric stretch mode was present in both the four-mode combination terms not converged within the first iteration of the procedure. In order to check that the excitation level of the VCI[gs,4] was adequately balanced to the potential containing also four-mode terms, a VCI[gs,5] calculation was carried out and the results confirm the analysis based on the VCI[gs,4] data.

5.5 Pyrimidine

As final test case, we consider the 24 fundamental vibrational frequencies of pyrimidine (1,3-diazine). The single point calculations needed for the construction of the PES (including the two mode couplings) were computed by means of DFT adopting the hybrid meta GGA functional M06 [49] and the Dunning/Hay double zeta basis sets improved with polarization functions [50] as available in GAMESS suite of codes [27].

The reference values for this test were computed at the VMP2 level of theory as described in Sect. 4 and adopting a PES representation obtained with the static grid approach. Such PES was constructed according to the scheme $(24_1)(12_1)$ [8], with a Fine Grid, 8 times more dense and placing the boundaries for each mode at the classical harmonic turning points of the state $v = 10$, and it required a total amount of 40,321 single point calculations. In order to investigate the quality of the potentials constructed with the ADGA, the fundamental frequencies obtained in the vibrational calculations where compared to the reference values and the mean absolute deviation and the maximum absolute deviation were computed for each of the PES. Table 7 shows the fundamental frequencies computed with each of the PES here described.

Three PES representations, named "A", "B" and "C", were constructed with the adaptive approach adopting tighter and tighter thresholds[4] and Gaussian basis sets in the VSCF calculations. The three PESs required 5,181, 8,507 and 15,839 single point calculations, respectively. Moreover, regardless to the set up, the convergence was achieved within 6 iterations for the monodimensional grid and within 4 iterations for bidimensional ones.

As expected, the statistical analysis in terms of MAD and MaxAD revealed that the fundamental frequencies, computed with the PES obtained iteratively, approach the reference values as the thresholds get tighter and the number of calculations increases. With the "A" potential, the MAD and MaxAD were found equal to 4.78 and 19.47 cm^{-1}, respectively, while if the "B" PES was used, these quantities decreased to 2.29 and 6.41 cm^{-1}, respectively. Finally, with the potential constructed with the tightest set of thresholds, "C", a MAD of 0.79 cm^{-1} and a MaxAD of 4.29 cm^{-1} were computed with respect to the reference set. Concerning the time spent in the ADGA (excluding the time for the single point calculations) for the "B" and "C" PES 45 and 100 CPU min were needed respectively see footnote3 of which the VSCF calculations required about 15 min in both cases.

[4] A: the 1D surfaces were converged with $\epsilon_{rel} = 1 \times 10^{-2}$ and $\epsilon_{abs} = 1 \times 10^{-6}$, whereas the 2D surfaces were converged with $\epsilon_{rel} = 2.5 \times 10^{-1}$ and $\epsilon_{abs} = 2.5 \times 10^{-5}$. Gaussian basis sets were used in the VSCF calculations. B: the 1D surfaces were converged with $\epsilon_{rel} = 5 \times 10^{-3}$ and $\epsilon_{abs} = 5 \times 10^{-7}$, whereas the 2D surfaces were converged with $\epsilon_{rel} = 1.2 \times 10^{-1}$ and $\epsilon_{abs} = 1.2 \times 10^{-5}$ Gaussian basis sets were used in the VSCF calculations. C: the 1D surfaces were converged with $\epsilon_{rel} = 3 \times 10^{-3}$ and $\epsilon_{abs} = 3 \times 10^{-7}$, whereas the 2D surfaces were converged with $\epsilon_{rel} = 5 \times 10^{-2}$ and $\epsilon_{abs} = 5 \times 10^{-6}$ Gaussian basis sets were used in the VSCF calculations. D: the 1D surfaces were converged with $\epsilon_{rel} = 5 \times 10^{-3}$ and $\epsilon_{abs} = 5 \times 10^{-7}$, whereas the 2D surfaces were converged with $\epsilon_{rel} = 1.2 \times 10^{-1}$ and $\epsilon_{abs} = 1.2 \times 10^{-5}$ HO basis sets were used in the VSCF calculations.

Table 7 Fundamental vibrational frequencies of pyrimidine (cm^{-1}) computed with ADGA, relative to the values computed with a standard grid

Mode		Reference	ADGA[a]			
			A[a]	B[a]	C[a]	D[a]
	# calcs.	40,321	5,181	8,507	15,839	9,351
1 A_1		2,976.11	−1.60	0.00	0.31	0.00
2 A_1		2,974.17	−5.00	−1.99	−0.20	−1.99
3 A_1		2,964.55	−1.20	3.29	4.29	3.29
4 A_1		1,606.37	−2.22	−1.46	−0.26	−1.46
5 A_1		1,464.98	−5.54	−3.29	−0.58	−3.28
6 A_1		1,144.85	−4.37	−1.37	−0.15	−1.37
7 A_1		1,058.10	−1.33	−0.58	−0.21	−0.58
8 A_1		983.53	−11.97	−6.41	−2.34	−6.40
9 A_1		678.03	−3.48	−1.49	−0.53	−1.49
10 A_2		980.07	−19.47	−3.38	−0.51	−3.38
11 A_2		395.62	−2.28	−1.67	−0.21	−1.67
12 B_1		1,027.51	0.78	−2.07	−1.85	−2.07
13 B_1		990.04	−5.65	−0.95	0.03	−0.95
14 B_1		825.78	−4.26	−4.67	−0.87	−4.67
15 B_1		729.46	−7.41	−2.82	−0.24	−2.82
16 B_1		346.32	−5.24	−1.73	−0.13	−1.73
17 B_2		2,955.03	−1.24	1.78	2.25	1.79
18 B_2		1,602.21	−1.88	−1.36	−0.29	−1.36
19 B_2		1,421.64	−3.02	−1.93	−0.37	−1.92
20 B_2		1,353.41	−10.92	−5.21	−1.05	−5.21
21 B_2		1,258.79	−2.40	−0.97	−0.33	−0.97
22 B_2		1,219.13	−6.49	−2.25	−0.98	−2.25
23 B_2		1,074.63	−5.06	−3.05	−0.35	−3.05
24 B_2		612.30	−1.95	−1.20	−0.62	−1.20
MAD			4.78	2.29	0.79	2.29
MaxAD			19.47	6.41	4.29	6.40

Mean absolute deviations (MAD) and the maximum absolute deviations (MaxAD) are reported. The construction of the 2D PES is based on M06/DZV single point calculations. Vibrational calculation: VMP2

[a] See footnote 4 for the ADGA thresholds

In order to investigate the effect of the basis set in the vibrational calculations during the ADGA, a fourth iterative PES, "D" was constructed matching the threshold of "B" but using HO functions in the VSCF calculations. The construction of "D" required 9,351 single point calculations and the analysis of the frequencies confirmed the equivalence with the quality achieved with the PES "B". For the pyrimidine molecule, it was therefore found that the basis set used in the vibrational calculation has little impact to the quality of the final results, confirming that the Gaussian basis sets may successfully replace the HO basis sets in ADGA vibrational calculations even when the latter are well-suited as in this specific case.

6 Summary and outlook

A new and highly black-box algorithm for constructing accurate semi-global analytical representations of potential energy and molecular property surfaces for use in vibrational structure calculations has been implemented and tested on a selection of model potentials as well as on water, difluoromethane and pyrimidine.

The algorithm has full flexibility with respect to mode coupling expansion of the potential and it has been fully integrated with the methods and techniques previously investigated by this group [2, 8, 30]. The main source of novelty of the ADGA resides in the use of the density as guidance for the definition of the configurational space where the potential is evaluated and in the definition of energy contributions as testing quantity to determine the convergence of the PES.

The use of localized Gaussians as primitive basis functions has been explored as well. The results demonstrate that the localized basis sets perform at least as good as the global sets in the well-conditioned problems whereas the drawbacks of the HO global basis set are avoided.

The robustness and reliability of the engine generating the PES has been tested in the generation of both mono and multidimensional surfaces. The tests demonstrate the high flexibility and robustness of the proposed approach and the ADGA is found to be accurate and cost-effective with respect to the number of evaluation points. In particular, the algorithm allows for inspection of higher mode couplings where all the many unimportant terms automatically will be treated with only a minimal number of points.

With this study, we hope to initiate the development of new ways for optimal PESs construction with algorithms requiring a small number of numerical thresholds as opposed to manual system-dependent optimization.

The basic ideas underlying the adaptive approach presented in this article can be combined with several other strategies for efficient construction of potential energy surfaces including multilevel approximations and Shepard interpolation technique. In particular, we are currently considering pre-screening of mode combination such that unimportant mode-combinations are not evaluated [5–7, 20] from the outset. The adaptive method may also be extended to use correlated densities and simultaneous convergence of the mono and multidimensional surfaces. Thus, the adaptive strategy outlined in the article is expected to be very useful in applications to both energies and properties of extended molecular systems. Work is underway for systems with up to hundred atoms.

Acknowledgments This work has been supported by the Lundbeck Foundation, the Danish national research foundation, the Danish

Center for Scientific Computing (DCSC), and EUROHORCs through a EURYI award.

Appendix

Integrals for distributed Gaussian basis sets

Gaussian integrals can be easily done analytically from a few standard integrals ($\int_{-\infty}^{\infty} e^{-\lambda x^2}\,\mathrm{d}x = \sqrt{\frac{\pi}{\lambda}}, \int_{-\infty}^{\infty} x^n e^{-\lambda x^2}\,\mathrm{d}x = (-\frac{\mathrm{d}}{\mathrm{d}\lambda})^{n/2}\sqrt{\frac{\pi}{\lambda}}$ with n an even positive integer number) and the fact that the product of two Gaussians is another Gaussian:

$$G_{ij}(x) = G_i(x)G_j(x) = \left(\frac{2\zeta_{ij}}{\pi}\right)^{\frac{1}{4}} \exp(-\zeta_{ij}(Z_i - Z_j)^2) \times \left(\frac{2(\zeta_i+\zeta_j)}{\pi}\right)^{\frac{1}{4}} \exp(-(\zeta_i+\zeta_j)(x - Z_{ij})^2) \quad (22)$$

with

$$\zeta_{ij} = \frac{\zeta_i\zeta_j}{\zeta_i+\zeta_j} \quad (23)$$

$$Z_{ij} = \frac{\zeta_i Z_i + \zeta_j Z_j}{\zeta_i+\zeta_j} \quad (24)$$

The overlap integrals are

$$O_{ij} = \langle G_i|G_j\rangle = \int_{-\infty}^{\infty} G_i^*(x)G_j(x)\,\mathrm{d}x = \left(\frac{4\zeta_{ij}}{(\zeta_i+\zeta_j)}\right)^{(\frac{1}{4})} \exp(-\zeta_{ij}(Z_i - Z_j)^2) \quad (25)$$

Clearly, the basis set of distributed Gaussians is a non-orthogonal basis with overlaps decaying fast with the distance between the centers of the Gaussians and with increasing exponents. Other integrals are

$$(x^n)_{ij} = \langle G_i|x^n|G_j\rangle = \sum_{k=0}^{int(n/2)} \binom{n}{2k} \frac{(2k-1)!!}{2^k(\zeta_i+\zeta_j)^k} Z_{ij}^{n-2k} O_{ij} \quad (26)$$

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)_{ij} = \langle G_i|\mathrm{d}/\mathrm{d}x|G_j\rangle = 2\zeta_j O_{ij}(Z_j - Z_{ij}) \quad (27)$$

$$T_{ij} = -\frac{1}{2}\langle G_i|\frac{d^2}{dx^2}|G_j\rangle = \zeta_j O_{ij}\left[1 - 2\zeta_j\left(\frac{1}{2(\zeta_i+\zeta_j)} + Z_{ij}^2 - 2Z_jZ_{ij} + Z_j^2\right)\right] \quad (28)$$

with $(2k-1)!! = (2k-1)\times(2k-3)\ldots\times 1$ and $0!! = 1$.

Other integrals have to be computed if the Hamiltonian contains Watson kinetic energy terms, which can be easily derived from the integrals listed above:

$$\langle G_i|x^n\frac{\mathrm{d}}{\mathrm{d}x}|G_j\rangle = -2\zeta_j\langle G_i|x^{n+1}|G_j\rangle + 2\zeta_j Z_j\langle G_i|x^n|G_j\rangle \quad (29)$$

$$\langle G_i|\frac{\mathrm{d}}{\mathrm{d}x}x^n|G_j\rangle = n\langle G_i|x^{n-1}|G_j\rangle + \langle G_i|x^n\frac{\mathrm{d}}{\mathrm{d}x}|G_j\rangle \quad (30)$$

$$\langle G_i|\frac{\mathrm{d}}{\mathrm{d}x}x^n\frac{\mathrm{d}}{\mathrm{d}x}|G_j\rangle = (4\zeta_j^2Z_j^2 - 2\zeta_j(n+1))\langle G_i|x^n|G_j\rangle + 2n\zeta_jZ_j\langle G_i|x^{n-1}|G_j\rangle + 4\zeta_j^2\langle G_i|x^{n+2}|G_j\rangle - 8\zeta_j^2Z_j\langle G_i|x^{n+1}|G_j\rangle \quad (31)$$

When fitting with scaled coordinates all these integrals are multiplied by a common factor of $(\sqrt{\omega})^n$ when the scaled coordinate x enters with a power of n.

References

1. Carter S, Culik SJ, Bowman JM (1997) J Chem Phys 107:10458
2. Kongsted J, Christiansen O (2006) J Chem Phys 125:124108
3. Bowman JM, Carter S, Huang XC (2003) Int Rev Phys Chem 22:533
4. Jung JO, Gerber RB (1996) J Chem Phys 105:10332
5. Rauhut G (2004) J Chem Phys 121:9313
6. Benoit DM (2004) J Chem Phys 120:562
7. Yagi K, Hirata S, Hirao K (2007) Theor Chem Acc 118:681
8. Toffoli D, Kongsted J, Christiansen O (2007) J Chem Phys 127:204106
9. Chaban GM, Jung JO, Gerber RB (1999) J Chem Phys 111:1823
10. Yagi K, Taketsugu T, Hirao K, Gordon M (2000) J Chem Phys 113:1005
11. Collins M (2002) Theor Chem Acc 108:313
12. Manzhos S, Carrington T (2006) J Chem Phys 125:084109
13. Manzhos S, Carrington T (2007) J Chem Phys 127:014103
14. Dawes R, Thompson DL, Guo Y, Wagner AF, Minkoff M (2007) J Chem Phys 126:184108
15. Dawes R, Thompson DL, Wagner AF, Minkoff M (2008) J Chem Phys 128:084107
16. Maeda S, Watanabe Y, Ohno K (2008) J Chem Phys 128:144111
17. Huang X, Braams BJ, Bowman JM (2005) J Chem Phys 122:044308
18. Tew D, Handy N, Carter S (2006) J Chem Phys 125:084313
19. Vendrell O, Gatti F, Lauvergnat D, Meyer HD (2007) J Chem Phys 127:184302
20. Pele L, Gerber RB (2008) J Chem Phys 128:165105
21. Oyanagi C, Yagi K, Taketsugu T, Hirao K (2006) J Chem Phys 124:064311
22. Midascpp (molecular interactions, dynamics and simulation chemistry program package in c++) (2007). http://www.chem.au.dk/~midas
23. Hamilton IP, Light JC (1986) J Chem Phys 84:306
24. Poirier B, Light J (2000) J Chem Phys 113:211
25. Partridge H, Schwenke DW (1997) J Chem Phys 106:4618
26. Dalton, A molecular electronic structure program, release 2.0 (2005). http://www.kjemi.uio.no/software/dalton/dalton.html
27. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su SJ, Windus TL, Dupuis M, Montgomery JA (1993) J Comput Chem 14:1347
28. Watson JKG (1968) Mol Phys 15:479
29. Beck MH, Jackle A, Worth GA, Meyer HD (2000) Phys Rep 324:1

30. Christiansen O (2007) Phys Chem Chem Phys 9:2942
31. Bowman JM (1978) J Chem Phys 68:608
32. Bowman JM (1986) Acc Chem Res 19:202
33. Gerber RB, Ratner MA (1988) Adv Chem Phys 70:97
34. Norris LS, Ratner MA, Roitberg AE, Gerber RB (1996) J Chem Phys 105:11261
35. Christiansen O (2003) J Chem Phys 119:5773
36. Matsunaga N, Chaban GM, Gerber RB (2002) J Chem Phys 117:3541
37. Bowman JM, Christoffel K, Tobin F (1979) J Phys Chem 83:905
38. Christoffel KM, Bowman JM (1982) Chem Phys Lett 85:220
39. Carter S, Bowman JM, Handy NC (1998) Theor Chem Acc 100:191
40. Christiansen O (2004) J Chem Phys 120:2149
41. Christiansen O (2004) J Chem Phys 120:2140
42. Christiansen O (2005) J Chem Phys 122:194105
43. Seidler P, Christiansen O (2007) J Chem Phys 126:204101
44. Seidler P, Hansen MB, Christiansen O (2008) J Chem Phys 128:154113
45. Coon JB, Naugle NW, Mckenzie RD (1966) J Mol Spectr 20:107
46. Lin CK, Chang HC, Lin SH (2007) J Phys Chem A 111:9347
47. Xie T, Bowman JM (2002) J Chem Phys 117:10487
48. Yanai T, Tew DP, Handy NC (2004) Chem Phys Lett 393:51
49. Zhao Y, Truhlar D (2008) Theor Chem Acc 120:215
50. Dunning Jr TH, Hay PJ (1977) In: Schaefer III HF (ed) Modern theoretical chemistry. Plenum Press, New York, p 1