

# Applying Conventional Ab Initio and Density Functional Theory Approaches to Electric Property Calculations. Quantitative Aspects and Perspectives

George Maroulis

**Abstract** We have examined the predictive capability of density functional theory methods in calculations of electric polarizability and hyperpolarizability. We have focused on test cases belonging to three high-priority classes of molecular systems: “soft” metal clusters, novel types of compounds, and weakly bonded molecules. The performance of theoretical methods over arbitrary collections of molecular properties can be analyzed and classified by the introduction of a new methodology based on graph theoretic considerations and pattern recognition techniques.

**Keywords** Ab initio · Density functional theory · Electric (hyper)polarizability

## Contents

1	Introduction .....	96
2	Theoretical Considerations .....	97
2.1	Electric Properties of Atoms, Molecules, and Clusters. Basic Theory and Computational Aspects .....	97
2.2	Interaction-Induced Electric Properties .....	99
2.3	Proximity, Similarity, and Order in Spaces of Theoretical Descriptions .....	100
3	Results .....	102
3.1	Sodium Tetramer, a Very Soft Molecule .....	102
3.2	New Classes of Molecules, the Case of HXeI .....	111
3.3	Interaction-Induced Polarizability and Hyperpolarizability of Two Water Molecules .....	119
4	Final Remarks and Conclusions .....	125
	References .....	126

---

G. Maroulis (✉)

Department of Chemistry, University of Patras, Patras 26500, Greece  
e-mail: [maroulis@upatras.gr](mailto:maroulis@upatras.gr)

## 1 Introduction

The theory of electric polarizability is of fundamental importance to the rational approach and interpretation of large classes of phenomena [1]. In particular, these properties are of fundamental importance to intermolecular interaction studies [2], nonlinear optics [3], collision-induced spectroscopy [4], and the simulation of fluids [5, 6]. They are also routinely associated with general molecular characteristics as hardness [7], softness [8], hypersoftness [9], stiffness [10], and compressibility [11]. Understandably, polarizability is also linked to reactivity [12]. Another important field is QSAR, QSPR studies, and the understanding of pharmacological activity [13, 14].

In view of the important applications, the theoretical determination of electric properties of atoms, molecules, clusters, and even larger molecular architectures is rapidly expanding. The predictive capability of theoretical methods and convergence to the available experimental data has been closely examined in comprehensive reviews [15].

Two wide classes of theoretical methods are preferentially applied to the determination of electric polarizabilities: *ab initio* methods and density functional theory (DFT)-based approaches. *Ab initio* methods have been known to be converging reliably, displaying a high level of agreement to experimentally determined quantities. Very accurate *ab initio* treatments of electric polarizabilities are available for atoms and relatively small molecules. The distinct advantage of DFT methods lies in the possibility of economical, in a computational sense, treatments of relatively large molecular architectures. There, their advantage ends. It is usually very hard to determine the predictive capability of DFT-based methods or just to reasonably compare their performance to the presumably more accurate *ab initio* methodologies.

In this paper, we investigate the possibilities offered by widely used DFT methods. We have chosen test cases in three different, difficult classes of problems: (1) the linear and nonlinear polarizabilities of metal clusters, (2) the polarizabilities of novel compounds, and (3) the interaction-induced polarizability in weakly bonded systems. In particular, the three test cases are the sodium tetramer, a particularly “soft” molecule, the new compound HXeI, and the interaction polarizability of two water molecules in the dimer  $(\text{H}_2\text{O})_2$ .

In previous work we employed an information theoretic approach to classify and systematically improve theoretical descriptions of molecules, introduced as arbitrary collections of data/properties [16, 17]. This approach relies on the availability of reference or accurate theoretical descriptions and is now easily accessible [18]. More recently, we have developed a more subtle, general approach based on graph theoretic arguments and pattern recognition techniques [19]. We rely on generalized metrics to introduce distance/proximity, order, and classification in spaces of theoretical descriptions. In addition, we introduce clustering in such spaces by the construction of a unique mathematical object, the minimum spanning tree (MST), and the performance of single-linkage cluster analysis (SLCA).

## 2 Theoretical Considerations

### 2.1 *Electric Properties of Atoms, Molecules, and Clusters. Basic Theory and Computational Aspects*

Our use of the theory of electric polarizability follows in all aspects, including the basic theoretical philosophy and terminology, the classic papers of Buckingham [20] and McLean and Yoshimine [21].

The energy ( $E^P$ ) and perturbed electric moments ( $\mu_\alpha^P$ ,  $\Theta_{\alpha\beta}^P$ ,  $\Omega_{\alpha\beta\gamma}^P$ ) of an uncharged molecule in a weak, static electric field can be expanded as

$$\begin{aligned}
 E^P &\equiv E^P(F_\alpha, F_{\alpha\beta}, F_{\alpha\beta\gamma}, F_{\alpha\beta\gamma\delta}, \dots) \\
 &= \mathbf{E}^0 - \boldsymbol{\mu}_\alpha F_\alpha - (1/3)\boldsymbol{\Theta}_{\alpha\beta} F_{\alpha\beta} - (1/15)\boldsymbol{\Omega}_{\alpha\beta\gamma} F_{\alpha\beta\gamma} - (1/105)\boldsymbol{\Phi}_{\alpha\beta\gamma\delta} F_{\alpha\beta\gamma\delta} + \dots \\
 &\quad - (1/2)\boldsymbol{\alpha}_{\alpha\beta} F_\alpha F_\beta - (1/3)\mathbf{A}_{\alpha,\beta\gamma} F_\alpha F_{\beta\gamma} - (1/6)\mathbf{C}_{\alpha\beta,\gamma\delta} F_\alpha F_\beta F_\gamma \\
 &\quad - (1/15)\mathbf{E}_{\alpha,\beta\gamma\delta} F_\alpha F_{\beta\gamma\delta} + \dots \\
 &\quad - (1/6)\boldsymbol{\beta}_{\alpha\beta\gamma} F_\alpha F_\beta F_\gamma - (1/6)\mathbf{B}_{\alpha\beta,\gamma\delta} F_\alpha F_\beta F_\gamma F_\delta + \dots \\
 &\quad - (1/24)\boldsymbol{\gamma}_{\alpha\beta\gamma\delta} F_\alpha F_\beta F_\gamma F_\delta + \dots, \tag{1}
 \end{aligned}$$

$$\begin{aligned}
 \mu_\alpha^P &= \boldsymbol{\mu}_\alpha + \boldsymbol{\alpha}_{\alpha\beta} F_\beta + (1/3)\mathbf{A}_{\alpha,\beta\gamma} F_{\beta\gamma} + (1/2)\boldsymbol{\beta}_{\alpha\beta\gamma} F_\beta F_\gamma + (1/3)\mathbf{B}_{\alpha\beta,\gamma\delta} F_\beta F_\gamma F_\delta \\
 &\quad + (1/6)\boldsymbol{\gamma}_{\alpha\beta\gamma\delta} F_\beta F_\gamma F_\delta + \dots, \tag{2}
 \end{aligned}$$

$$\Theta_{\alpha\beta}^P = \Theta_{\alpha\beta} + \mathbf{A}_{\gamma,\alpha\beta} E_\gamma + \mathbf{C}_{\alpha\beta,\gamma\delta} F_\gamma F_\delta + (1/2)\mathbf{B}_{\gamma\delta,\alpha\beta} F_\gamma F_\delta + \dots, \tag{3}$$

$$\Omega_{\alpha\beta\gamma}^P = \boldsymbol{\Omega}_{\alpha\beta\gamma} + \mathbf{E}_{\delta,\alpha\beta\gamma} F_\delta + \dots, \tag{4}$$

where the variables  $F_\alpha$ ,  $F_{\alpha\beta}$ ,  $F_{\alpha\beta\gamma}$ , etc., are the field, field gradient, etc. at the origin of the molecule. The terms **in bold** are the permanent properties of the system: energy ( $\mathbf{E}^0$ ), dipole ( $\boldsymbol{\mu}_\alpha$ ), quadrupole ( $\boldsymbol{\Theta}_{\alpha\beta}$ ), octopole ( $\boldsymbol{\Omega}_{\alpha\beta\gamma}$ ), and hexadecapole ( $\boldsymbol{\Phi}_{\alpha\beta\gamma\delta}$ ) moment. The second-, third-, and fourth-order properties are the dipole polarizability ( $\boldsymbol{\alpha}_{\alpha\beta}$ ), dipole–quadrupole polarizability ( $\mathbf{A}_{\alpha,\beta\gamma}$ ), quadrupole polarizability ( $\mathbf{C}_{\alpha\beta,\gamma\delta}$ ), dipole–octopole polarizability ( $\mathbf{E}_{\alpha,\beta\gamma\delta}$ ), first dipole hyperpolarizability ( $\boldsymbol{\beta}_{\alpha\beta\gamma}$ ), dipole–dipole–quadrupole hyperpolarizability ( $\mathbf{B}_{\alpha\beta,\gamma\delta}$ ), and second dipole hyperpolarizability ( $\boldsymbol{\gamma}_{\alpha\beta\gamma\delta}$ ). The subscripts denote Cartesian components and a repeated subscript implies summation over  $x$ ,  $y$ , and  $z$ . The number of independent components needed to specify the above tensors is strictly regulated by symmetry. In addition to the Cartesian components, of interest are the various invariants of some tensors. For the dipole polarizability and second dipole hyperpolarizability the mean is defined as

$$\bar{\alpha} = (1/3)\alpha_{\alpha\alpha} \text{ and } \bar{\gamma} = (1/5)\gamma_{\alpha\alpha\beta\beta}. \tag{5}$$

For some systems of simple symmetry, as linear molecules, the invariants of dipole (hyper)polarizability tensors, mean values, and anisotropies are easily defined and commonly used. Most of them are measurable quantities or can be deduced from experimental observations.

$$\begin{aligned}
 \bar{\alpha} &= (\alpha_{zz} + 2\alpha_{xx})/3 \\
 \Delta\alpha &= \alpha_{zz} - \alpha_{xx} \\
 \bar{\beta} &= (3/5)(\beta_{zzz} + 2\beta_{zxx}) \\
 \Delta\beta &= \beta_{zzz} - 3\beta_{zxx} \\
 \bar{\gamma} &= (3\gamma_{zzzz} + 8\gamma_{xxxx} + 12\gamma_{xxzz})/15 \\
 \Delta_1\gamma &= 3\gamma_{zzzz} - 4\gamma_{xxxx} + 3\gamma_{xxzz} \\
 \Delta_2\gamma &= \gamma_{zzzz} + \gamma_{xxxx} - 6\gamma_{xxzz}.
 \end{aligned} \tag{6}$$

When sufficiently weak electric fields are applied it is possible to extract the electric properties of the molecule from the above expansions. In previous work, we have applied various computational schemes based on the finite-field [22] approach to the calculation of electric properties from perturbed atomic/molecular energies and induced multipole moments [23–28].

Various computational aspects of the theoretical determination of electric properties are available in books [29, 30] or comprehensive collections [31–33].

All calculations reported in the following sections have been performed with quantum chemical methods easily accessible via the widely used GAUSSIAN suite of programs. See GAUSSIAN 98 [34] and GAUSSIAN 03 [35]. This arsenal includes conventional ab initio methods and density functional theory (DFT) approaches. We will not give further details about the structure and predictive capability of these methods here. For the interested reader, extensive presentations of the above classes of methods are clearly presented in standard references [36–39].

The ab initio methods used are:

- *SCF*, self-consistent-field
- *MP2*, *MP3*, and *MP4*: second-, third-, and fourth-order Møller–Plesset perturbation theory
- *DQ-MP4* and *SDQ-MP4*, partial fourth-order Møller–Plesset
- *CCSD*, singles and doubles coupled cluster
- *CCSD(T)*, which includes an estimate of connected triples via a perturbational treatment. This is the method with the, presumably, highest predictive potential

The DFT methods include the widely used B3LYP, B3PW91, and mPW1PW91 and many more that have been occasionally employed in electric property calculations.

Basis sets in modern quantum chemistry is too broad a subject to be examined in detail here [40]. The search for Gaussian basis sets suitable for molecular property calculations is vital to computational quantum chemistry [41]. Many computational

schools have reported significant work in the direction of basis set construction [42–44] and testing of their performance. The construction of purpose-oriented basis sets has attracted particular attention [45–49]. Early work by Dykstra and coworkers [50, 51] or Spackman [52] has elucidated many important aspects of the performance of basis sets in electric property calculations. In our computational efforts we favor the use of molecule-specific, purpose-oriented basis sets. The construction of molecule-specific, purpose-oriented basis sets for large molecular architectures or low-symmetry polyatomics is largely impractical. Nevertheless, previous work shows that for systems of reasonable size as atoms [23, 53], atomic anions or cations [54, 55], diatomics [56, 57] and triatomics [58–61], symmetric polyatomics [62–64], and clusters [65–68], one can easily control the construction of the basis set in order to obtain suitable basis sets for electric property calculations.

## 2.2 Interaction-Induced Electric Properties

Interaction electric properties, as dipole moment, polarizability, and hyperpolarizability are of fundamental importance to the analysis and interpretation of measurements and observations in collision- and interaction-induced spectroscopies [69, 70]. Considerable progress has been recorded in recent years either on theoretical issues or computational advances. We single out a few significant papers related to the above field. Głaz et al. [71] reported a study of the collision-induced hyper-Rayleigh light-scattering spectra of He–Ne atomic pairs. Chrysos et al. [72] reported a study of the CO<sub>2</sub>–Ar collision-induced  $\nu_3$  CO<sub>2</sub> band. A study of fundamental importance for collision-induced spectroscopy of gaseous CO<sub>2</sub>, the determination of the exact low-order classical moments in collision-induced bands of linear rotors, was reported by Chrysos et al. [73]. A new treatment of the collision-induced Raman scattering by Ne–Ne was reported by Chrysos et al. [74]. Baranowska et al. [75] reported a theoretical study of the interaction-induced dipole moment and polarizability of CO–Ne. El-Kader et al. [76] determined the contributions of multipolar polarizabilities to the isotropic and anisotropic light scattering induced by intermolecular interactions in gaseous CH<sub>4</sub>. Zvereva-Loëte et al. [77] and Buldakov et al. [78] reported an extensive study on the dipole moment surface and dipole polarizability surface for the CH<sub>4</sub>–N<sub>2</sub> complex, a system of importance for the atmospheric physics of Titan. The calculation of the interaction-induced dipole moment, polarizability, and first and second hyperpolarizability of the H<sub>2</sub>O–Rg (Rg = He, Ne, Ar, Kr and Xe) complexes was reported by Haskopoulos and Maroulis [79]. Hartmann et al. [80] reported an investigation of the far-infrared collision-induced absorption band in gaseous CO<sub>2</sub>.

Our work on the interaction-induced polarizability and hyperpolarizability of the water dimer relies on the conventional supermolecule approach. Details on this approach have been given in previous work on interaction-induced electric properties. In this approach, the interaction-induced properties of the A...B supermolecule are obtained as

$$P_{\text{int}}(A \cdots B) = P(A \cdots B) - P(A) - P(B). \quad (7)$$

In practice, the above equation is nearly exact when very large, nearly saturated basis sets are used. When truncated basis sets are used the basis set superposition error is removed by the counterpoise-correction method of Boys and Bernardi [81]. Thus, the above equation is replaced by

$$P_{\text{int}}(A \cdots B) = P(A \cdots B) - P(A \cdots X) - P(X \cdots B), \quad (8)$$

where  $P(A \cdots X)$  denote calculation of the property for subsystem A in the presence of the ghost orbitals of subsystem B and  $P(X \cdots B)$  for subsystem B in the presence of the ghost orbitals of subsystem A. It is easily shown that as the flexibility of the basis set increases one approaches a nearly ideal situation where  $P(A \cdots X) \approx P(A)$  and  $P(X \cdots B) \approx P(B)$ .

For all computational aspects of the interaction-induced (hyper)polarizability of the water dimer, we lean heavily on previous experience of systems as  $\text{CO}_2\text{-Rg}$  [82],  $\text{Ne-Ar}$  [83],  $\text{Xe-Xe}$  [84],  $\text{Kr-Xe}$  [85],  $\text{He-Ar}$  [86],  $\text{H}_2\text{-Ar}$  [87],  $\text{H}_2\text{-Ne}$  [88],  $\text{Kr-He}$  [89],  $\text{Kr-Ne}$  [90], and  $\text{H}_2\text{-Ne}$  [91].

### 2.3 Proximity, Similarity, and Order in Spaces of Theoretical Descriptions

The evaluation of the performance of theoretical methods in atomic/molecular property calculations has been recognized as a formidable problem early enough. The quantification of the relative merit of theoretical methods is an essential part of modern computational quantum chemistry. Sometime ago we presented a methodology that relies on graph theoretic arguments and pattern recognition techniques to introduce order, classification, and clustering in spaces of arbitrary theoretical descriptions of atomic/molecular systems [92]. Our theory uses metric considerations to define distance/proximity and similarity in such spaces. The utility of such a methodology has been brought forth in various applications. Pattern recognition has long found application in Chemistry [93]. Our use of such techniques extends the application of pattern recognition (PR) to a hitherto unreached field: computational quantum chemistry (CQC). The logical strength of the analogy is made obvious by the following diagram:

Pattern recognition (PR)		Computational quantum chemistry (CQC)
Object	$\leftrightarrow$	Method
Features	$\leftrightarrow$	Molecular property values
Pattern	$\leftrightarrow$	Theoretical description (TD)
Pattern space	$\leftrightarrow$	Space of all TD

Thus in PR one has *objects* and in CQC *methods*. Objects are characterized by *features* and methods by *molecular property values*. A collection of features is a

*pattern*; a collection of molecular property values a *theoretical description* (TD). A collection of patterns forms the pattern space, a collection of theoretical descriptions the *space of all theoretical descriptions*.

The above-mentioned methodology has found application in various quantum chemistry problems [94–96]. We give here only a few essential points and definitions.

Let  $TD_i$  be a collection of properties  $Q_{m\alpha}$  where the index  $m$  denotes methods and the index  $\alpha$  denotes properties. The two indices take values in the index sets  $I_\alpha$  and  $I_m$  (respectively).

$$TD_i = \{Q_{m\alpha}, m \in I_m, \alpha \in I_\alpha\}. \quad (9)$$

We denote by TD the space of all theoretical descriptions  $TD_i$ :

$$TD = \{TD_1, TD_2, \dots, TD_N\} \quad \text{where } i = 1, 2, \dots, N. \quad (10)$$

We define a generalized distance in the space of all TD by using the Minkowski metric. The distance between two theoretical descriptions  $TD_i$  and  $TD_j$  is defined as

$$D_{ij} \equiv D(TD_i, TD_j) = \left( \sum_{\alpha} \frac{|Q_{i\alpha} - Q_{j\alpha}|^p}{\left( \max_{ij} |Q_{i\alpha} - Q_{j\alpha}| \right)^p} \right)^{1/p}, \quad p \geq 1, 1 \leq i, j \leq N. \quad (11)$$

In most applications we have used the Euclidean metric, that is  $p = 2$ .

*Similarity* between two theoretical descriptions  $TD_i$  and  $TD_j$  is then defined on the basis of distance/proximity as

$$S_{ij} = 1 - \frac{D_{ij}}{\max_{ij} D_{ij}}, \quad 1 \leq i, j \leq N. \quad (12)$$

By definition,  $0 \leq S_{ij} \leq 1$ .

To make a connection with graph theory we need a few standard definitions and interpretations. A reliable source of basic graph theory is the work of Chartrand and Lesniak [97]. A graph  $G$  is a finite nonempty set of objects called vertices together with a set of unordered pairs of vertices called edges. The vertex set of  $G$  is denoted  $\mathbf{V}(G)$  and the edge set  $\mathbf{E}(G)$ . The cardinality  $\mathbf{p}$  of  $V(G)$  is called the order of  $G$  and the cardinality  $\mathbf{q}$  of  $E(G)$  the size of  $G$ . Let us consider the Cartesian product  $TD \times TD$ . The *graph of theoretical descriptions*  $G_{TD}$  has as vertex set  $V(G_{TD})$  the set of theoretical descriptions  $TD_i$ . The edge set  $E(G_{TD})$  is a subset of above defined Cartesian product,  $E(G_{TD}) \subset TD \times TD$ .

We assign to each edge of  $G_{TD}$  a real number, a *weight*. Consider the edge defined by  $TD_i$  and  $TD_j$ . We assign to the edge  $\{TD_i, TD_j\}$  the real number  $D_{ij} \equiv D(TD_i, TD_j)$ .  $G_{TD}$  is now a *weighted graph*.

The diameter  $\text{Diam } G_{TD}$  of the graph of TD is defined as

$$\text{Diam } G_{\text{TD}} = \max_{i,j \in V(G_{\text{TD}})} D_{ij} \equiv \max_{\text{TD}_i, \text{TD}_j \in V(G_{\text{TD}})} D(\text{TD}_i, \text{TD}_j). \quad (13)$$

In view of the above, the definition of the similarity can now be written as

$$S_{ij} \equiv S(\text{TD}_i, \text{TD}_j) = 1 - \frac{D(\text{TD}_i, \text{TD}_j)}{\text{Diam } G_{\text{TD}}}, \text{TD}_i, \text{TD}_j \in V(G_{\text{TD}}). \quad (14)$$

Two more definitions are particular useful at this stage. The distance of fixed vertex  $u$  from a subset of the vertex set  $S \subseteq V(G_{\text{TD}})$  is defined as

$$d(u, S) = \min_{x \in S} \{d(u, x)\}. \quad (15)$$

A point of major importance to our methodology is the definition of a *minimum spanning tree* (MST). A *spanning subgraph*  $H$  of a graph  $G$  has vertex and edge sets  $V(H) \subseteq V(G)$  and  $E(H) \subseteq E(G)$  and is of the same order as  $G$ . A graph  $G$  of order  $p$  and size  $q$  is a tree if and only if it is acyclic and  $p = q + 1$ . The weight of a spanning tree in a connected graph is the sum of the weights of its edges. Thus a *minimum spanning tree* of  $G$  is a *spanning tree* of  $G$  of minimum weight.

Suitable algorithms exist for the construction of a MST [98].

Here is now a synoptic form of our methodology step by step:

1. Define the space of theoretical descriptions TD.
2. Calculate all distances between vertices or  $\text{TD}_i$  of the graph  $G_{\text{TD}}$ .
3. Calculate the distance and similarity matrix.
4. Construct a weighted minimum spanning tree *MST* using known algorithms.
5. Perform a *single linkage cluster analysis* (SLCA) by removing all edges from the *MST* characterized by weights above a given threshold  $D_T$ .

This analysis creates a partition of the *MST* in distinct clusters  $C_1, C_2, \dots, C_K$ . The union of all clusters is the vertex set of the  $G_{\text{TD}}$  graph:  $V(G_{\text{TD}}) = \bigcup_{1 \leq i \leq K} C_i$ .

Last, it is important to introduce a distance between clusters for our analysis. The nearest neighbor distance between clusters  $C_m$  and  $C_n$  is defined as

$$D_{nn}(C_m, C_n) = \min_{\text{TD}_i \in C_m, \text{TD}_j \in C_n} D(\text{TD}_i, \text{TD}_j). \quad (16)$$

## 3 Results

### 3.1 Sodium Tetramer, a Very Soft Molecule

The electric dipole polarizability of sodium clusters has been extensively studied, both experimentally [99–101] and theoretically [102–104]. In addition, the polarizability of the sodium atom is accurately known. The latest experimental value is



$\alpha(\text{Na}) = 162.7 \pm 0.8$ , reported by Ekstroem et al. [105]. Reliable theoretical values have been reported both for the Hartree-Fock limit and higher levels of theory. The most accurate theoretical value  $\alpha(\text{Na}) = 162.88 \pm 0.6$  has been reported by Thakkar and Lupinetti [106]. Consequently, reference values are readily available for the differential polarizability (DP) and the differential-per-atom-polarizability (DPA), defined as [66, 107]:

$$\text{DP} \equiv \bar{\alpha}_{\text{diff}} = (\bar{\alpha}(\text{Na}_n) - n\alpha(\text{Na})),$$

$$\text{DPA} \equiv \bar{\alpha}_{\text{diff}}/n = (\bar{\alpha}(\text{Na}_n) - n\alpha(\text{Na}))/n = (\bar{\alpha}(\text{Na}_n))/n - \alpha(\text{Na}).$$

The respective hyperpolarizability quantities, differential hyperpolarizability (DH) and the differential-per-atom-hyperpolarizability (DHPA), are defined as

$$\text{DH} \equiv \bar{\gamma}_{\text{diff}} = (\bar{\gamma}(\text{Na}_n) - n\gamma(\text{Na})),$$

$$\text{DHPA} \equiv \bar{\gamma}_{\text{diff}}/n = (\bar{\gamma}(\text{Na}_n) - n\gamma(\text{Na}))/n = (\bar{\gamma}(\text{Na}_n))/n - \gamma(\text{Na}).$$

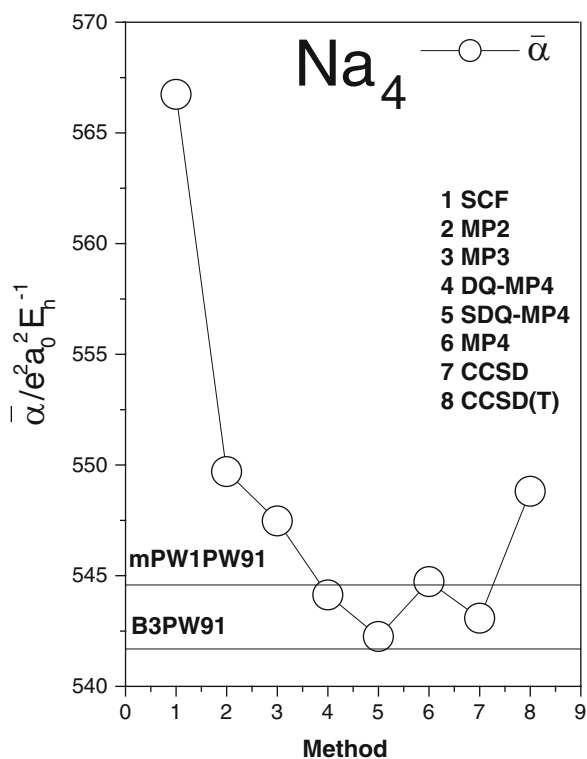
DH and DHPA are even more basis set and method sensitive than DP and DPA. As it has been shown elsewhere the above quantities are very useful as rigorous criteria for the analysis of the performance of quantum chemical methods [108, 109].

We have chosen as testing ground for our theoretical treatment the sodium tetramer ( $\text{Na}_4$ ). Most of the data pertaining to  $\text{Na}_4$  and used in this section have been published elsewhere [110]. The molecular geometry of the tetramer is a rhombus with side 3.64 Å and a short diagonal of 3.30 Å. The molecule is placed on the  $xz$  plane with the four Na atoms at  $(0,0,\pm 1.65)$  and  $(\pm 3.244549,0,0)$ . All electric property calculations have been performed with a purpose-oriented molecule-specific basis set of [7s5p2d] type. The theoretical methods used are conventional ab initio and a selection of DFT approaches. The ab initio methods are SCF, MP2, MP3, DQ-MP4, SDQ-MP4, MP4, CCSD, and CCSD(T). The choice of DFT methods leans heavily on previous experience [111–115]: B3LYP, B3PW91, PBEPBE, PBEPW91, mPW1PW91, and mPW1PBE as implemented in the GAUSSIAN suite of codes.

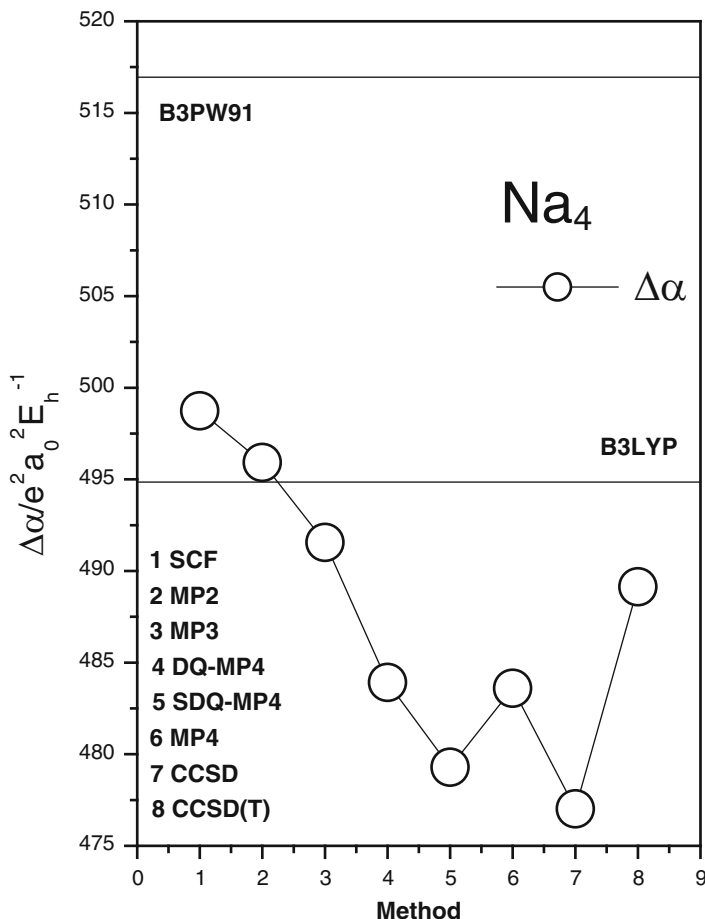
The independent Cartesian components of the polarizability ( $\alpha_{\alpha\beta}$ ) and hyperpolarizability ( $\gamma_{\alpha\beta\gamma\delta}$ ) are given in Table 1. The method-dependence of the mean and the anisotropy of the dipole polarizability and the mean second hyperpolarizability are shown in Figs. 1, 2, and 3, respectively. Figure 1 shows clearly that electron correlation reduces the size of the mean dipole polarizability. The sequence  $\bar{\alpha}(\text{SCF}) < \bar{\alpha}(\text{MP2}) < \bar{\alpha}(\text{MP3}) < \bar{\alpha}(\text{MP4})$  shows a monotonic decrease of the values calculated with Møller–Plesset perturbation theory. Our best theoretical value  $\bar{\alpha}(\text{CCSD(T)})$  is quite close to  $\bar{\alpha}(\text{MP2})$  and  $\bar{\alpha}(\text{MP3})$ . For comparison, we show on the same figure the DFT values for the mean calculated with mPW1PW91 and B3PW91. The former is close to the MP4 value and the latter close to SDQ-MP4. Figure 2 shows that the method-dependence of the dipole polarizability

**Table 1** Static (hyper)polarizability of  $\text{Na}_4$  calculated with ab initio and DFT methods ( $10^{-3} \times \gamma_{\alpha\beta\gamma\delta}$ ). All quantities in atomic units

Method	$\alpha_{xx}$	$\alpha_{yy}$	$\alpha_{zz}$	$\gamma_{xxxx}$	$\gamma_{yyyy}$	$\gamma_{zzzz}$	$\gamma_{xxyy}$	$\gamma_{yyzz}$	$\gamma_{xxzz}$
SCF	894.87	356.17	449.18	16448	2563	3236	3369	881	3465
MP2	875.48	338.02	435.61	12941	2377	3153	2740	840	2721
MP3	870.31	337.28	434.84	11030	2222	2976	2368	796	2271
DQ-MP4	861.83	336.70	433.86	10443	2090	2789	2176	753	2037
SDQ-MP4	856.81	336.39	433.53	10036	1950	2597	2022	704	1861
MP4	862.06	336.75	435.39	10340	2016	2703	2109	728	1950
CCSD	855.83	336.83	436.57	11337	2035	2732	2247	729	2275
CCSD(T)	869.53	337.39	439.50	11700	2082	2833	2357	749	2414
B3LYP	835.22	299.40	395.43	12250	1990	2655	2722	733	2579
B3PW91	881.89	323.89	419.28	9946	1889	2614	2457	689	2466
PBEPBE	859.05	311.76	406.94	11143	1883	2640	2642	712	2511
PBEPW91	854.05	309.38	404.83	11309	1874	2614	2634	709	2500
mPW1PW91	884.63	327.59	421.52	9143	1868	2524	2366	694	2340
mPW1PBE	889.76	329.96	423.63	8754	1878	2548	2361	699	2334



**Fig. 1** Method dependence of the mean dipole polarizability of  $\text{Na}_4$



**Fig. 2** Method dependence of the anisotropy the dipole polarizability of Na<sub>4</sub>

anisotropy values bear close resemblance to that of the mean. We observe again the monotonic  $\Delta\alpha(\text{SCF}) < \Delta\alpha(\text{MP2}) < \Delta\alpha(\text{MP3}) < \Delta\alpha(\text{MP4})$ . Agreement with the DFT methods is less obvious in this case.  $\Delta\alpha(\text{B3LYP})$  is quite close to  $\Delta\alpha(\text{MP2})$  but  $\Delta\alpha(\text{B3PW91})$  is significantly higher than the ab initio values. Inspecting Fig. 3, we are surprised to find that the various orders of Møller–Plesset perturbation theory display once more a monotonic behavior for the mean second dipole hyperpolarizability:  $\bar{\gamma}(\text{SCF}) < \bar{\gamma}(\text{MP2}) < \bar{\gamma}(\text{MP3}) < \bar{\gamma}(\text{MP4})$ .  $\bar{\gamma}(\text{MP3})$  is quite close to our, presumably, most accurate  $\bar{\gamma}$  (CCSD) and  $\bar{\gamma}$  (CCSD(T)). What is more three conventional DFT methods as B3LYP, B3PW91 and mPW1PW91 yield mean second hyperpolarizability values close enough to the high-level ab initio ones.

We have used the data in Table 1 to calculate distance/proximity and similarities for the ab initio and DFT theoretical descriptions (TD) of the electric (hyper) polarizability of the sodium tetramer. We have a total of nine independent Cartesian components so we have in our hands a 9D problem. The similarities between the

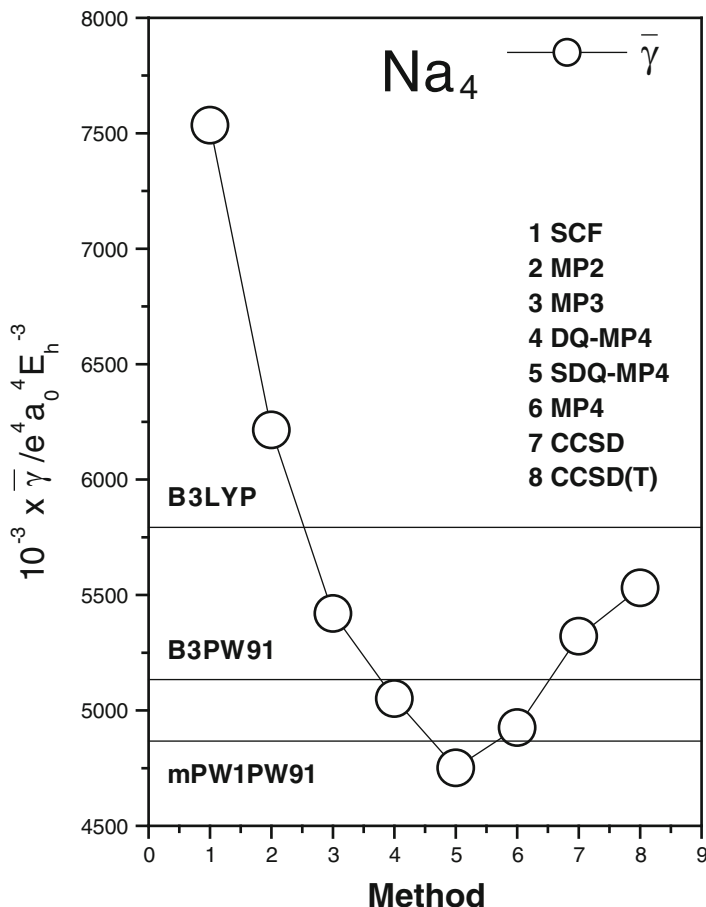


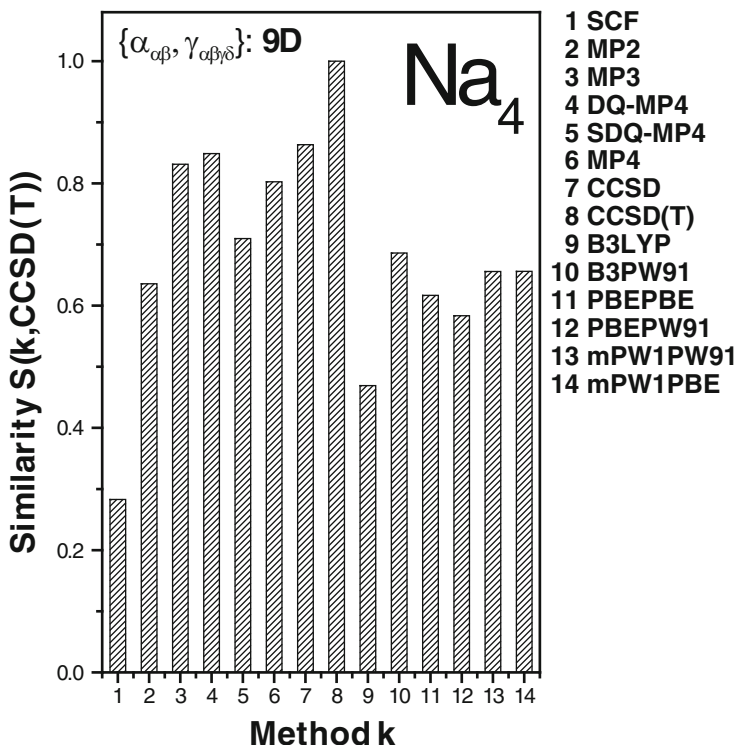
Fig. 3 Method dependence of the mean second hyperpolarizability of  $\text{Na}_4$

theoretical descriptions (TD) of Table 1 are given in Table 2. The quantities shown in Table 2 allow us a direct quantitative evaluation of the similarity of the performance of two arbitrary methods over the calculation of the (hyper)polarizability of the sodium tetramer. In Fig. 4 we show a histogram with the similarity of any method to the presumably most accurate one CCSD(T). As this is our reference value,  $S(\text{CCSD(T)}, \text{CCSD(T)}) = 1$ . The method closest to CCSD(T) is CCSD:  $S(\text{CCSD}, \text{CCSD(T)}) = 0.86337$ . Among the DFT methods, the closest to the reference is B3PW91, with  $S(\text{B3PW91}, \text{CCSD(T)}) = 0.68609$ . We are also able to glean more specific information about the relative performance of the various methods. Of all DFT methods, B3PW91 is closest to mPW1PW91,  $S(\text{B3PW91}, \text{mPW1PW91}) = 0.90949$ . PBEPBE is closest to PBEPW91,  $S(\text{PBEPBE}, \text{PBEPW91}) = 0.95325$  and mPW1PW91 closest to mPW1PBE,  $S(\text{mPW1PBE}, \text{mPW1PW91}) = 0.94854$ .

The minimum spanning tree (MST) of the theoretical descriptions of Table 1 is shown in Fig. 5. The tree is a Graph  $G(p,q)$  of order 14 and magnitude 13

**Table 2** Similarity of the performance of theoretical methods on the electric (hyper)polarizability of  $\text{Na}_4$

	SCF	MP2	MP3	DQ-MP4	SDQ-MP4	MP4	CCSD	CCSD(T)	B3LYP	B3PW91	PBEPBE	PBEPW91	mPWIPW91	mPWIPBE
SCF	1	0.57488	0.33884	0.16830	0.00280	0.09796	0.16868	0.28289	0	0.07459	0.05613	0.02549	0.02893	0.04121
MP2	0.57488	1	0.74056	0.55296	0.36835	0.46939	0.52730	0.63584	0.32852	0.40102	0.3940	0.36280	0.35548	0.36370
MP3	0.33884	0.74056	1	0.80334	0.61266	0.71629	0.74301	0.83139	0.42900	0.58075	0.53971	0.50707	0.55275	0.56012
DQ-MP4	0.16830	0.55296	0.80334	1	0.80770	0.90759	0.88283	0.84868	0.48311	0.67442	0.61870	0.59192	0.66610	0.66250
SDQ-MP4	0.00280	0.36835	0.61266	0.80770	1	0.89422	0.81645	0.70972	0.46637	0.68950	0.61724	0.60095	0.70972	0.69245
MP4	0.09796	0.46939	0.71629	0.90759	0.89422	1	0.88070	0.80267	0.47075	0.70139	0.62366	0.59946	0.70706	0.69923
CCSD	0.16868	0.52730	0.74301	0.88283	0.81645	0.88070	1	0.86337	0.51992	0.69631	0.65619	0.63300	0.67626	0.66199
CCSD(T)	0.28289	0.63584	0.83139	0.84868	0.70972	0.80267	0.86337	1	0.46914	0.68609	0.61677	0.58336	0.65589	0.65614
B3LYP	0	0.32852	0.42900	0.48311	0.46637	0.47075	0.51992	0.46914	1	0.54306	0.76642	0.80058	0.48531	0.44434
B3PW91	0.07459	0.40102	0.58075	0.67442	0.68950	0.70139	0.69631	0.68609	0.54306	1	0.77031	0.73162	0.90949	0.88024
PBEPBE	0.05613	0.39400	0.53971	0.61870	0.61724	0.62366	0.65619	0.61677	0.76642	0.77031	1	0.95325	0.70809	0.66765
PBEPW91	0.02549	0.36280	0.50707	0.59192	0.60095	0.59946	0.63300	0.58336	0.80058	0.73162	0.95325	1	0.67413	0.63113
mPWIPW91	0.02893	0.35548	0.55275	0.66610	0.70972	0.70706	0.67626	0.65589	0.48531	0.90949	0.70809	0.67413	1	0.94854
mPWIPBE	0.04121	0.36370	0.56012	0.66250	0.69245	0.69923	0.66199	0.65614	0.44434	0.88024	0.66765	0.63113	0.94854	1



**Fig. 4** Histogram of the similarities  $S(k,CCSD(T))$  where  $k =$  ab initio and DFT for the 9D descriptions  $\{\alpha_{\alpha\alpha}, \alpha_{\gamma\gamma}, \alpha_{\alpha\alpha}, \gamma_{\alpha\alpha\alpha\alpha}, \gamma_{\gamma\gamma\gamma\gamma}, \gamma_{\alpha\alpha\alpha\alpha}, \gamma_{\alpha\alpha\gamma\gamma}, \gamma_{\gamma\gamma\alpha\alpha}, \gamma_{\alpha\alpha\alpha\alpha}\}$

( $p = q + 1$ ). Removing all edges  $D(i,j)$  above a threshold of  $D_T = 0.4$ ,  $D(i,j) > D_T$  results in the clustering shown also in Fig. 5. The obtained clusters are as follows:

C1 = {SCF}

C2 = {MP2}

C3 = {MP3}

C4 = {DQ-MP4, SDQ-MP4, MP4, CCSD, CCSD(T)}

C5 = {B3LYP}

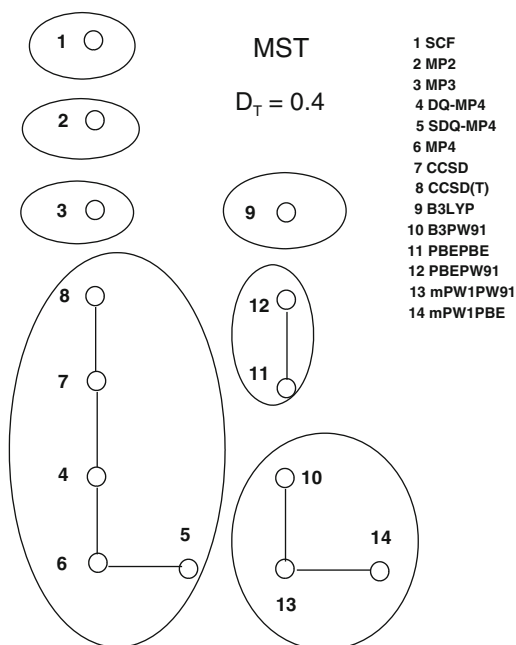
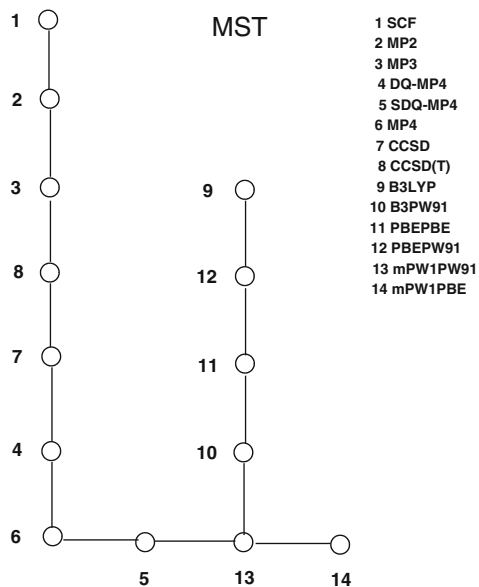
C6 = {PBEPBE, PBEPW91}

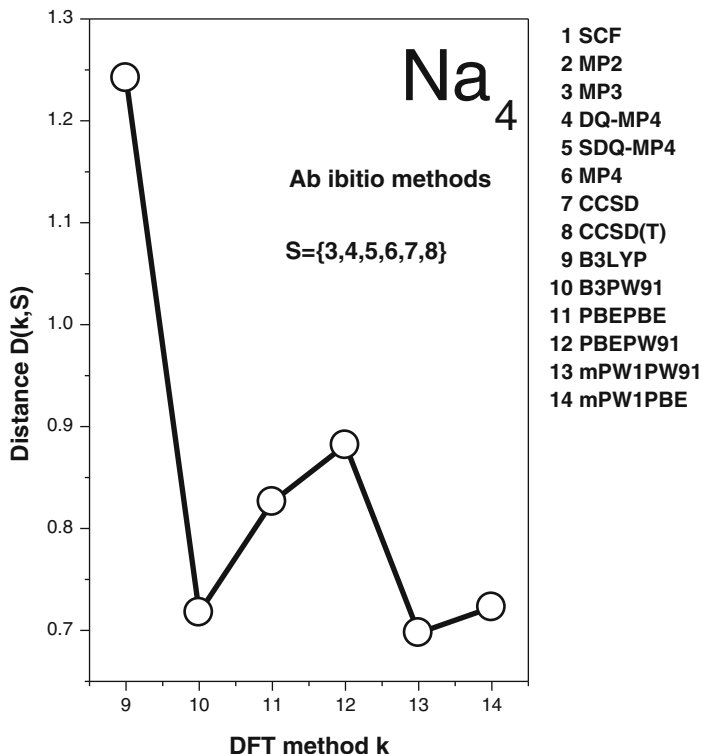
C7 = {B3PW91, mPW1PW91, mPW1PBE}

$$TD = C1 \cup C2 \cup C3 \cup C4 \cup C5 \cup C6 \cup C7$$

High-level ab initio results form a large cluster  $C4 = \{DQ-MP4, SDQ-MP4, MP4, CCSD, CCSD(T)\}$  which contains the most accurate theoretical descriptions CCSD and CCSD(T). The DFT methods form three distinct clusters C5, C6, and

**Fig. 5** Minimum spanning tree (*MST*) for the theoretical descriptions of the sodium tetramer and subsequent clustering with a distance threshold  $D_T = 0.4$ .





**Fig. 6** Distance of the DFT theoretical descriptions from the  $C4 \equiv S$  cluster that contains the high-level ab initio methods

C7. The theoretical description of all DFT closest to CCSD(T) is B3PW91 which belongs to the C7 cluster,  $B3PW91 \in C7$ .

In Fig. 6 we show the variation of the distance of all DFT methods from the  $C4 \equiv S$  cluster that contains all high-level ab initio methods. Instead of comparing DFT methods individually to one ab initio theoretical description we define directly an element-set distance as follows:

$$D(x, S) = \min_{s \in S} (x, s).$$

Here,  $D(x, S)$  defines the distance of a theoretical description  $x$  from the  $S$  cluster. In this case  $S$  is the cluster that contains all high-level ab initio theoretical descriptions and  $x$  a DFT method. In Fig. 6 we see clearly the distance of all DFT methods from the  $C4 \equiv S$  cluster. This is a very realistic view of the highly complex matter of the performance of DFT methods. Figure 6 reveals that the DFT method most distant from the  $S$  cluster is the B3LYP one. The respective distance is  $D(B3LYP, S) = 1.242$ . The methods most close to the  $S$  cluster are



B3PW91 and mPW1PW91. The respective distances are  $D(\text{B3PW91},S) = 0.7175$  and  $D(\text{mPW1PW91},S) = 0.6975$ . Ordering the DFT methods with respect to their distance from  $S$  results in the sequence  $D(\text{B3LYP},S) > D(\text{PBEPW91},S) > D(\text{PBEPBE},S) > D(\text{mPW1PBE},S) > D(\text{B3PW91},S) > D(\text{mPW1PW91},S)$ .

### 3.2 New Classes of Molecules, the Case of HXeI

HXeI is a typical representative of a fascinating new class of molecules. HRgX compounds are produced by UV radiation of hydrogen halides (HX) in rare gas (RG) matrices [116]. They have attracted considerable experimental and theoretical attention [117]. The importance of the HXeI molecule in particular was brought forth by the work of Buck and Farnik [118]. In their experiments this linear molecule is detected by orientation in strong laser and weak electric fields [119]. Computational experience on HXeI is relatively limited. An empirical estimate of the anisotropy of the dipole polarizability has been proposed by Nahler et al. [120] in their work on the photodissociation of oriented HXeI molecules generated from HI-Xe<sub>n</sub> clusters.

In this section we turn our attention on the electric dipole moment, polarizability, and hyperpolarizability of this important species. We lean heavily for molecular data and insights on our recent paper on the electric properties of HXeI [121].

We take into account two classes of molecular properties. Ab initio results calculated with basis sets (given in I/Xe/H order) B5  $\equiv$  [10s9p8d1f/9s8p7d1f/6s3p1d] (179 CGTF) and B9  $\equiv$  [11s10p10d3f/9s8p7d1f/7s5p1d] (214 CGTF). The B9 basis set was used in the DFT calculations. With basis B5 we calculated SCF, MP2, SDQ-MP4, MP4, CCSD, and CCSD(T) and with the larger B9 basis SCF and MP2 values. All ab initio values were taken from the above-cited paper. The DFT results were calculated with basis B9 and the methods B1LYP, B3LYP, B3PW91, mPW1PW91, HCTH, BHandH, BHandHLYP, PBEPBE, and PBEPW91. The molecular geometry of this linear molecule is defined by the bond lengths  $R(\text{I-Xe}) = 3.0577$  and  $R(\text{Xe-I}) = 1.7077$  Å. It was obtained at the MP2(Full)/B9 level of theory. At the same MP2/B9 level, a natural bond orbital analysis (NBO) yields charges of  $-0.55803$  (I),  $0.57691$  (Xe), and  $-0.01888$  (H). The calculated independent Cartesian components ( $z$  is the molecular axis) of the electric property tensors are given in Table 3.

The molecular values listed in Table 3 show clearly the difficulty in predicting reliable (hyper)polarizabilities for HXeI. This is particularly evident in the observed variations of the longitudinal components of the first and second hyperpolarizability. We base most of the presentation and analysis in this part on the invariants of (hyper)polarizability: mean values  $\bar{\alpha}$ ,  $\Delta\alpha$ ,  $\bar{\beta}$ , and  $\bar{\gamma}$ . Figure 7 shows the method dependence of the dipole moment. We also give a few characteristic DFT values. The high-level ab initio data are well grouped together. The DFT method closest to the most accurate CCSD(T)/B5 value is the B3PW91/B9 one which yields  $\mu_z = 2.4181$  ea<sub>0</sub>. The BHandHLYP method predicts a dipole moment

**Table 3** Electric properties of HXeI at the theoretical molecular geometry

Basis	Method	$\mu$	$\alpha_{zz}$	$\alpha_{xx}$	$\beta_{zzz}$	$\beta_{zxx}$	$\gamma_{zzzz}$	$\gamma_{xxxx}$	$\gamma_{xxzz}$
B5	SCF	3.1642	158.1475	58.7304	-2788.7	179.5	159534	25857	7392
	MP2	2.6450	178.4193	61.3754	-2502.8	198.1	45910	33854	13524
	SDQ-MP4	2.6243	177.7069	61.1082	-2489.1	208.9	36556	32441	11775
	MP4	2.4836	181.7763	61.9752	-2111.3	205.5	-19509	35827	13888
	CCSD	2.6229	176.8633	60.9417	-2338.8	214.2	21232	31478	11800
	CCSD(T)	2.4584	181.2297	61.5717	-1858.2	220.2	-51560	33852	13652
B9	SCF	3.1621	158.1235	58.8893	-2776.3	182.4	157783	24797	6106
	MP2	2.6102	178.3887	61.5569	-2453.5	196.3	19795	8707	-2040
B9	B1LYP	2.4300	164.8947	60.9186	-1568.5	194.6	124931	36455	15993
	B3LYP	2.3857	164.4301	60.9369	-1487.3	193.0	124643	36478	16142
	B3PW91	2.4181	161.6319	59.8737	-1438.6	182.0	111561	32474	14202
	mPW1PW91	2.4670	161.3968	59.6325	-1492.7	179.9	108993	31914	13952
	HCTH	2.2235	163.5979	61.9986	-1344.2	206.4	164649	47502	20850
	BHandH	2.7623	162.1835	59.4488	-1970.6	175.6	111730	29621	11917
	BHandHLYP	2.6863	163.1892	59.2536	-1947.4	182.3	108529	29737	11884
	PBEPBE	2.2242	162.5679	61.9794	-1239.9	192.8	137383	40680	18205
	PBEPW91	2.2194	162.6710	62.0161	-1236.2	193.5	137304	40812	18287

All quantities in atomic units

close enough to the MP2 one. In Fig. 8 the evolution of the method dependence of the mean dipole polarizability shows a clear gap between SCF and the post-Hartree-Fock methods, a sign of a very strong electron correlation effect. We observe a clear discrepancy between DFT and ab initio methods. Among the most reliable DFT methods the HCTH and B3PW91 yield mean dipole polarizabilities  $\bar{\alpha} = 95.87$  and  $93.79 e^2 a_0^2 E_h^{-1}$ , respectively, both clearly below the most accurate CCSD and CCSD(T) results. The anisotropy of the dipole polarizability is normally a more severe test for DFT methods than that of the mean. In Fig. 9, the method dependence of the anisotropy resembles closely that of the mean. Among all other DFT methods, we note the performance of BHandHLYP, B3PW91, and PBE which give  $\Delta\alpha = 103.94$ ,  $101.76$ , and  $100.59 e^2 a_0^2 E_h^{-1}$ , respectively, all three close enough to the SCF value. The ab initio values for the mean first hyperpolarizability are shown in Fig. 10. Of all DFT methods, BHandHLYP performs best, yielding a value  $\bar{\beta} = -949.6 e^3 a_0^3 E_h^{-2}$ . This result is close to both MP4 and CCSD(T). The performance of other DFT methods is characteristically poor. B3PW91 yields  $\bar{\beta} = -644.8 e^3 a_0^3 E_h^{-2}$ , a value significantly lower in magnitude than the BHandHLYP one. Last, ab initio and DFT values for the mean second hyperpolarizability are given in Fig. 11. Electron correlation lowers significantly the magnitude of this important property. With the notable exception of BH and HLYP and B3PW91, which predict values close to the SCF one,  $\bar{\gamma} = -50993$  and  $47073 e^4 a_0^4 E_h^{-3}$ , respectively, all other DFT methods yield considerably higher values.

We have calculated distance/proximity values for all methods/basis sets used in this section. The respective theoretical descriptions are 8D strings of molecular properties  $\{\mu_z, \alpha_{zz}, \alpha_{xx}, \beta_{zzz}, \beta_{zxx}, \gamma_{zzzz}, \gamma_{xxxx}, \gamma_{xxzz}\}$ . The calculated values are given

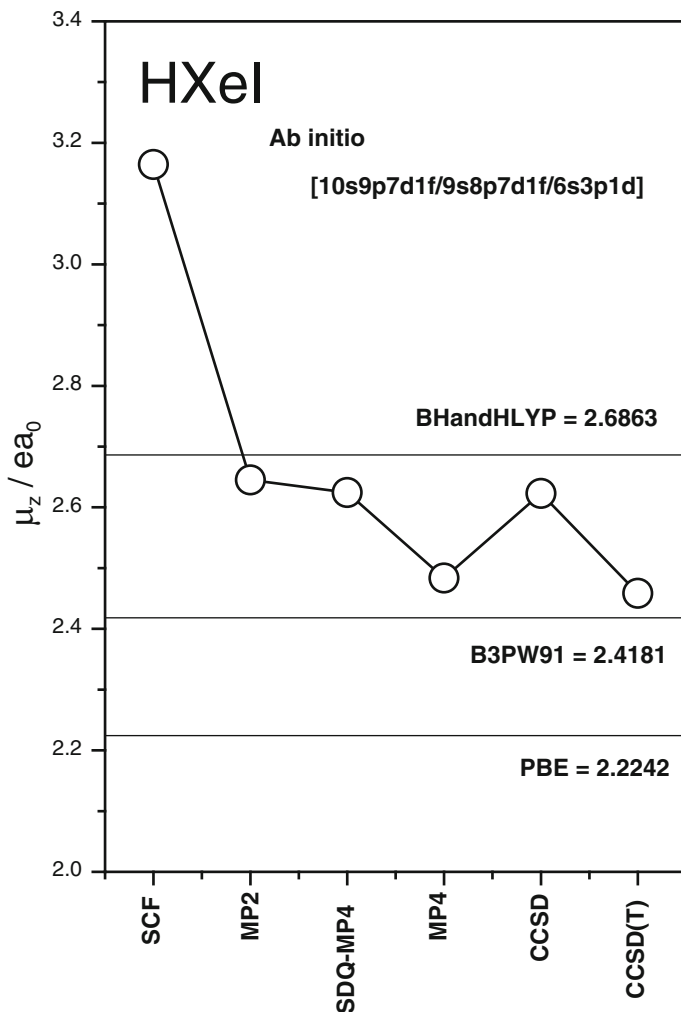
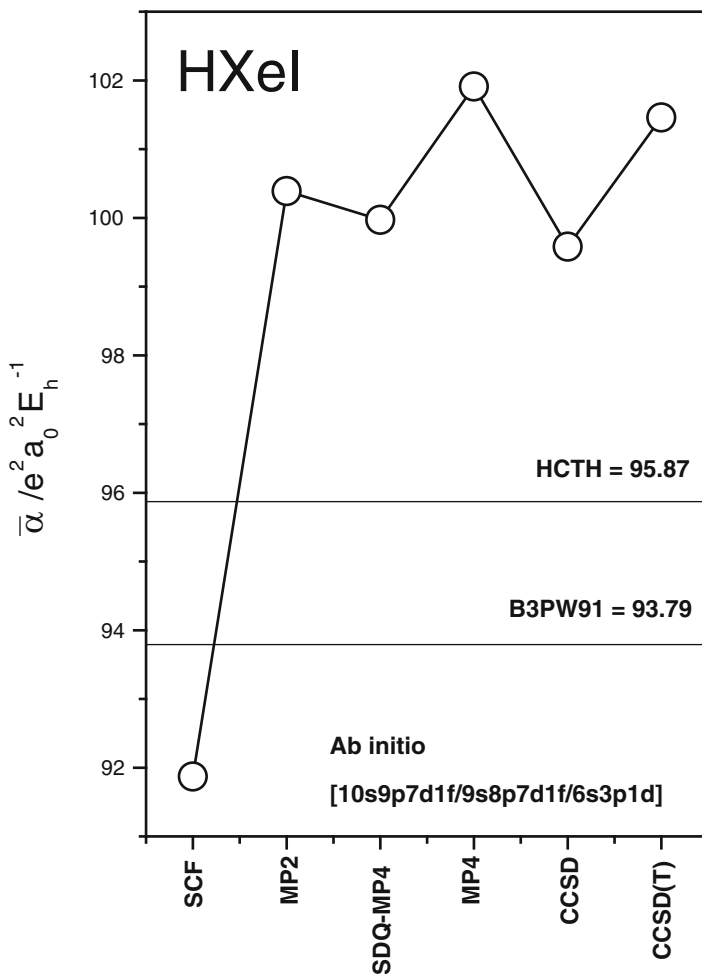


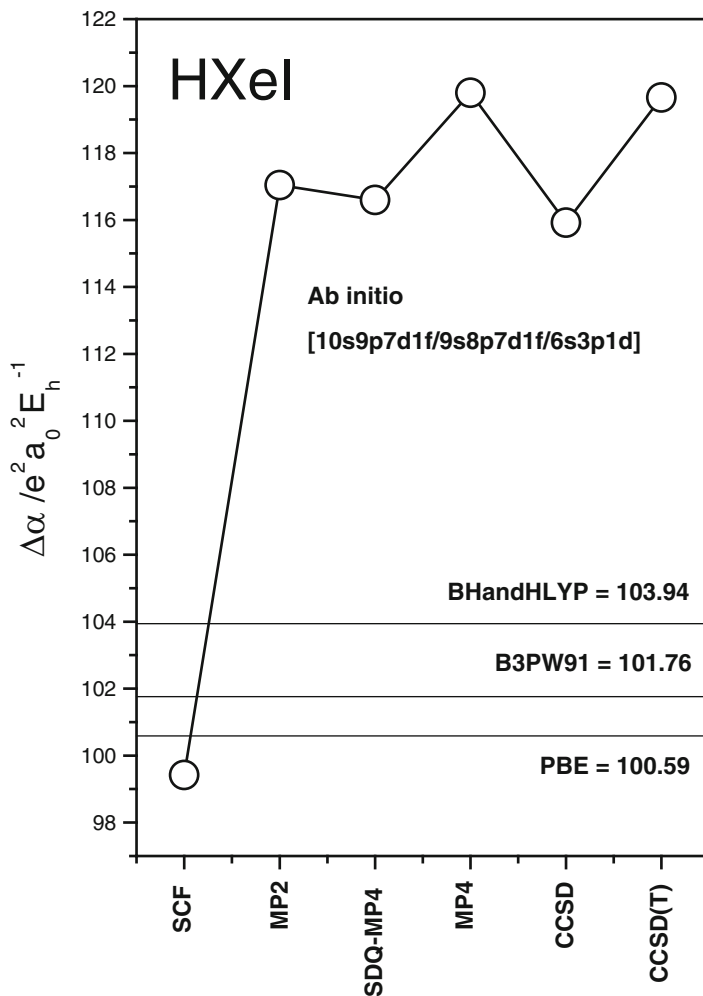
Fig. 7 Method dependence of the dipole moment of HXeI

in Table 4. To simplify matters, we have assigned numbers to all methods: 1 = SCF, 2 = MP2, 3 = SDQ-MP4, 4 = MP4, 5 = CCSD, 6 = CCSD(T), 7 = SCF, 8 = MP2, 9 = B1LYP, 10 = B3LYP, 11 = B3PW91, 12 = mPW1PW91, 13 = HCTH, 14 = BHandH, 15 = BHandHLYP, 16 = PBEPBE, and 17 = PBEPW91. To avoid confusion we sometimes denote methods 7 and 8 by 7 = SCF-B9 and 8 = MP2-B9, respectively, to avoid confusion with 1 = SCF and 2 = MP2 results that have been calculated with basis set B5. The most distant or dissimilar theoretical descriptions are 1 = SCF and 6 = CCSD(T):  $S(1,6) = 0$ . The two SCF/B5 and SCF/B9 descriptions are very similar,  $S(1,7) = 0.95157$ . This is not the case for the MP2/B5 and MP2/B9 descriptions as  $S$



**Fig. 8** Method dependence of the mean polarizability of HXeI

(2,8) = 0.55279. In Fig. 12 we show an histogram of the evolution of the similarity  $S(k,6) \equiv S(k, \text{CCSD(T)})$ . As 6 = CCSD(T) is the reference theoretical description here the maximum similarity for the histogram in Fig. 12 is  $S(6,6) = 1$ . The ab initio theoretical description closest to 6 = CCSD(T) is 4 = MP4:  $S(4,6) = 0.80257$ . From the DFT descriptions the closest to 6 = CCSD(T) is 9 = B11YP:  $S(6,9) = 0.41201$ . Thus, as one easily gathers from the histogram, the similarity between ab initio (calculated with basis set B5) and DFT descriptions is rather poor. This is also the case for the comparison of ab initio results calculated with basis sets B5 and B9. To view similarity from another perspective, we show in Fig. 13 the evolution of the similarity  $S(k,11) \equiv S(k, \text{B3PW91})$ . The 11 = B3PW91 method is similar enough to MP2, SDQ-MP4, and CCSD. It is most close to the 12 = mPW1PW91 method:  $S(11,12) = 0.94834$ .



**Fig. 9** Method dependence of the polarizability anisotropy of HXeI

The minimum spanning tree (MST) corresponding to the graph representing the theoretical descriptions (Table 3) is shown in Fig. 14. This tree is a graph  $G(p,q)$  of order  $p = 17$  and size  $q = p - 1$ . One expects this graph to be of sufficient complex structure. Clustering by removing all edges greater than a threshold value  $D_T = 0.4$  results in the following partitioning of the spaces of theoretical descriptions:

- C1  $\equiv \{1,7\}$
- C2  $\equiv \{2,3,5\}$
- C3  $\equiv \{4\}$
- C4  $\equiv \{6\}$

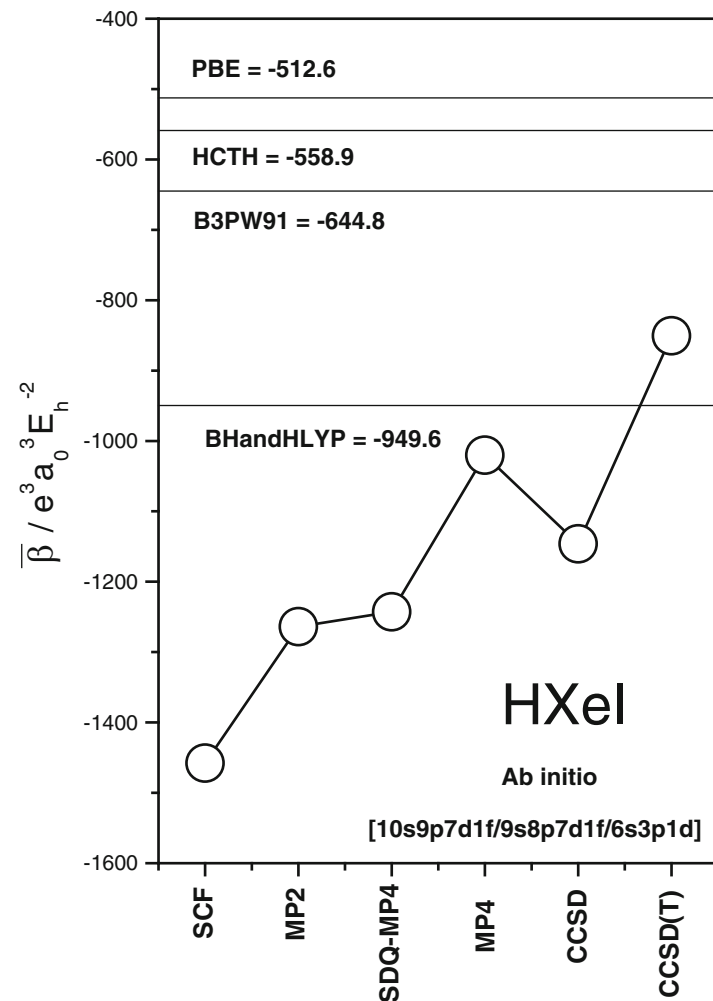


Fig. 10 Method dependence of the mean first hyperpolarizability of HXeI

C5  $\equiv$  {8}

C6  $\equiv$  {9,10}

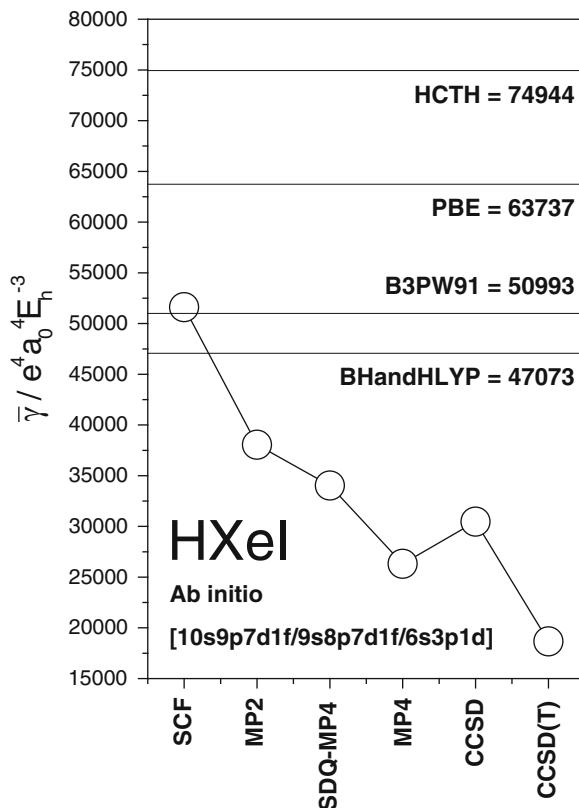
C7  $\equiv$  {11,12}

C8  $\equiv$  {13,16,17}

C9  $\equiv$  {14,15}

$$TD = C1 \cup C2 \cup C3 \cup C4 \cup C5 \cup C6 \cup C7 \cup C8 \cup C9 \equiv \bigcup_i C_i, i = 1, 2, \dots, 9$$

**Fig. 11** Method dependence of the mean second hyperpolarizability of HXeI



Some characteristics of this clustering are easily discernible. The two SCF and SCF-B9 descriptions form cluster C1. Three high-level ab initio theoretical descriptions form cluster C2  $\equiv \{2,3,5\} \equiv \{\text{MP2}, \text{SDQ-MP4}, \text{CCSD}\}$ . The composition of the C6, C7, and C8 clusters grouping DFT methods is clearly understood: B3PW91 is very close to mPW1PW91, BH and H is close to BH and HLYP, and PBEPBE is close to PBEPW91. HCTH belongs to the same cluster as PBEPBE and PBEPW91 and for good reason: it is most similar to these two DFT methods with  $S(\text{HCTH}, \text{PBEPBE}) \approx S(\text{HCTH}, \text{PBEPW91}) \approx 0.8$ .

Last, in Fig. 15 we show the evolution of the distance of the ab initio descriptions calculated with basis set B5 from the reference group of widely used B3LYP, B3PW91, and mPW1PW91. The distance  $d(x, S)$ , where  $S = \{\text{B3LYP}, \text{B3PW91}, \text{mPW1PW91}\}$  and  $x \in \Omega \equiv \{\text{SCF}, \text{MP2}, \text{SDQ-MP4}, \text{MP4}, \text{CCSD}, \text{CCSD(T)}\}$ , is shortest for the most accurate CCSD(T) method:

$$d(\text{CCSD(T)}, S) \leq d(x, S) \quad x \in \Omega$$

**Table 4** Similarity of the performance of theoretical methods on the electric (hyper)polarizability of HXeI

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	1	0.28938	0.27216	0.07659	0.24691	0	0.95157	0.21474	0.31802	0.29578	0.38810	0.42457	0.06871	0.62357	0.59859	0.12434	0.11749
2	0.28938	1	0.87055	0.75645	0.79704	0.61276	0.30389	0.55279	0.54984	0.52172	0.43077	0.42155	0.37835	0.46129	0.48752	0.40343	0.40237
3	0.27216	0.87055	1	0.75570	0.91459	0.67538	0.29179	0.56777	0.52845	0.49821	0.40435	0.39383	0.36635	0.42352	0.46524	0.37277	0.37264
4	0.07659	0.75645	0.75570	1	0.75124	0.80257	0.09307	0.49525	0.47353	0.45494	0.33823	0.31800	0.36556	0.29868	0.33247	0.38519	0.38670
5	0.24691	0.79704	0.91459	0.75124	1	0.72627	0.26807	0.54984	0.52566	0.49619	0.40304	0.39116	0.36818	0.40579	0.45498	0.36943	0.36977
6	0	0.61276	0.67538	0.80257	0.72627	1	0.01971	0.42909	0.41201	0.39452	0.28160	0.25971	0.32456	0.21827	0.26705	0.32338	0.32609
7	0.95157	0.30389	0.29179	0.09307	0.26807	0.01971	1	0.24443	0.32761	0.30446	0.38960	0.42390	0.07897	0.61925	0.59852	0.13342	0.12682
8	0.21474	0.55279	0.56777	0.49525	0.54984	0.42909	0.24443	1	0.3193	0.30055	0.27217	0.26888	0.10919	0.32457	0.34357	0.18465	0.18269
9	0.31802	0.54984	0.52845	0.47353	0.52566	0.41201	0.32761	0.3193	1	0.96155	0.77561	0.73615	0.69901	0.62023	0.65465	0.77331	0.76842
10	0.29578	0.52172	0.49821	0.45494	0.49619	0.39452	0.30446	0.30055	0.96155	1	0.78799	0.74413	0.70584	0.60903	0.64114	0.79780	0.79211
11	0.38810	0.43077	0.40435	0.33823	0.40304	0.28160	0.38960	0.27217	0.77561	0.78799	1	0.94834	0.51457	0.74095	0.76795	0.62996	0.62168
12	0.42457	0.42155	0.39383	0.31800	0.39116	0.25971	0.42390	0.26888	0.73615	0.74413	0.94834	1	0.46906	0.78064	0.80443	0.58152	0.57313
13	0.06871	0.37835	0.36635	0.36556	0.36818	0.32456	0.07897	0.10919	0.69901	0.70584	0.51457	0.46906	1	0.33563	0.36696	0.81210	0.81885
14	0.62357	0.46129	0.42352	0.29868	0.40579	0.21827	0.61925	0.32457	0.62023	0.60903	0.74095	0.78064	0.33563	1	0.91223	0.43192	0.42385
15	0.59859	0.48752	0.46524	0.33247	0.45498	0.26705	0.59852	0.34357	0.65465	0.64114	0.76795	0.80443	0.36696	0.91223	1	0.45053	0.44312
16	0.12434	0.40343	0.37277	0.38519	0.36943	0.32338	0.13342	0.18465	0.77331	0.79780	0.62996	0.58152	0.81210	0.43192	0.45053	1	0.99007
17	0.11749	0.40237	0.37264	0.38670	0.36977	0.32609	0.12682	0.18269	0.76842	0.79211	0.62168	0.57313	0.81885	0.42385	0.44312	0.99007	1

1 = SCF

2 = MP2

3 = SDQ-MP4

4 = MP4

5 = CCSD

6 = CCSD(T)

7 = SCF-B9

8 = MP2-B9

9 = B1LYP

10 = B3LYP

11 = B3PW91

12 = mPW/PPW91

13 = HCTH

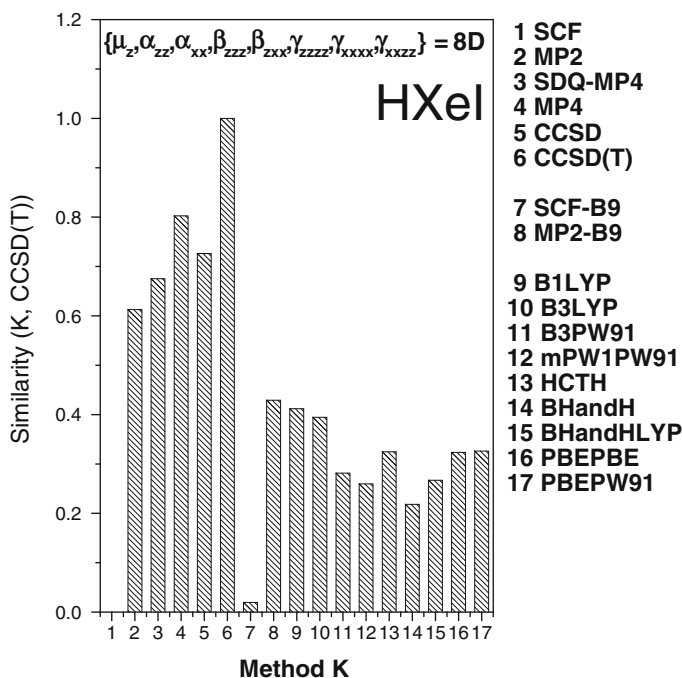
14 = BHandH

15 = BHandHLYP

16 = PBEPBE

17 = PBEPW91





**Fig. 12** Histogram of the similarities  $S(k, \text{CCSD}(T))$  where  $k = \text{ab initio}$  and DFT for the 8D descriptions  $\{\mu_z, \alpha_{zz}, \alpha_{xx}, \beta_{zzz}, \beta_{zxx}, \gamma_{zzzz}, \gamma_{xxxx}, \gamma_{xxzz}\}$

The distance  $d(x, S)$  increases as follows:

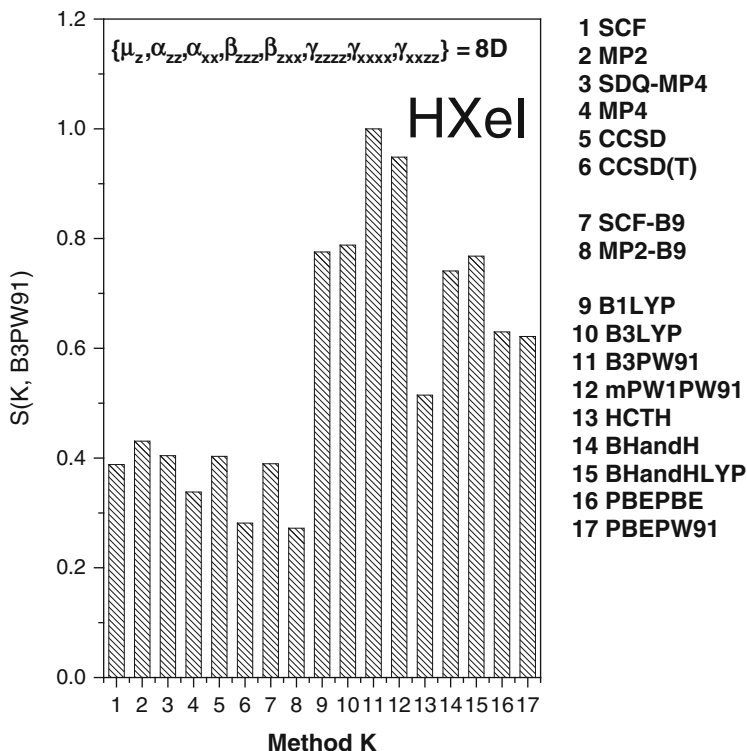
$$d(\text{CCSD}(T), S) < d(\text{SCF}, S) < d(\text{MP4}, S) < d(\text{CCSD}, S) < d(\text{SDQ} - \text{MP4}, S) < d(\text{MP2}, S).$$

### 3.3 Interaction-Induced Polarizability and Hyperpolarizability of Two Water Molecules

In previous work [122] we presented an extended computational study of the interaction-induced electric properties of the water dimer  $(\text{H}_2\text{O})_2$ . Our findings strongly suggest that the interaction-induced mean dipole polarizability and hyperpolarizability are nearly additive, as

$$\bar{\alpha}(\text{H}_2\text{O})_2 \approx 2\bar{\alpha}(\text{H}_2\text{O}) \text{ and } \bar{\gamma}(\text{H}_2\text{O})_2 \approx 2\bar{\gamma}(\text{H}_2\text{O}).$$

This surprising result appears to hold for the dipole polarizability of certain classes of water clusters, as brought forth in the work of Rodriguez et al. [123] or Ghanty and Ghosh [124]. The above defined additivity results clearly suggests that



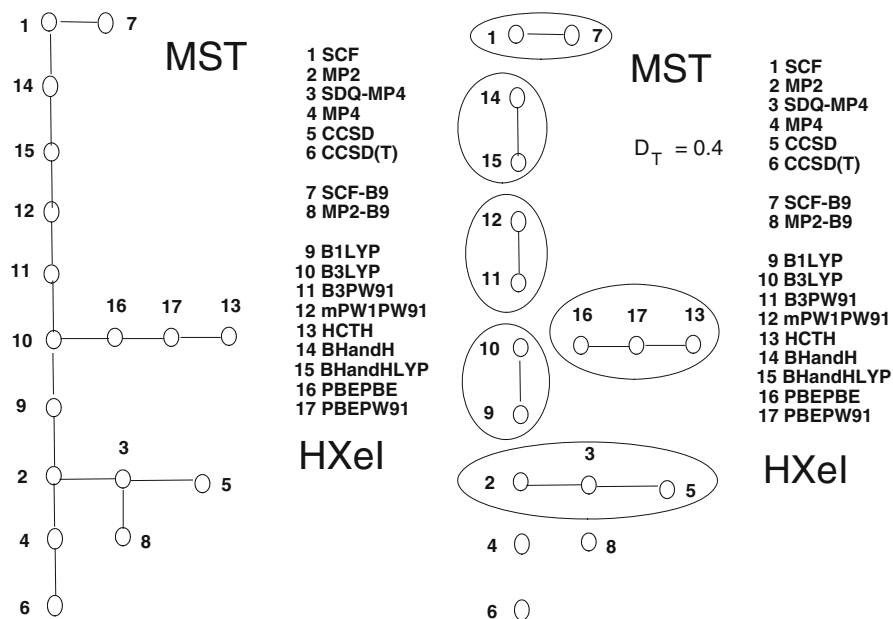
**Fig. 13** Histogram of the similarities  $S(k, B3PW91)$  where  $k = ab\ initio$  and DFT for the 8D descriptions  $\{\mu_z, \alpha_{zz}, \alpha_{xx}, \beta_{zzz}, \beta_{zxx}, \gamma_{zzzz}, \gamma_{xxxx}, \gamma_{xxzz}\}$

the respective interaction-induced mean (hyper)polarizability is very small. Identifying the proton acceptor (A) and proton donor (D) in the two moieties of  $(H_2O)_2 \equiv AD$ , we define the interaction properties in the counterpoise corrected (CP) scheme as

$$\begin{aligned}\bar{\alpha}_{\text{int}}(H_2O)_2 &= \bar{\alpha}(H_2O)_2 - \bar{\alpha}(H_2O-X) - \bar{\alpha}(X-H_2O), \\ \bar{\gamma}_{\text{int}}(H_2O)_2 &= \bar{\gamma}(H_2O)_2 - \bar{\gamma}(H_2O-X) - \bar{\gamma}(X-H_2O).\end{aligned}$$

In a very recent paper, we demonstrated that the calculation of the interaction-induced (hyper)polarizability of a molecular system as the water dimer constitutes a sever test of the validity of DFT methods [125]. In this section we add more calculations in order to extend our observations to the performance of a larger class of DFT-based approximations.

Full computational details about the work presented in this section are given in our papers on the water dimer. See Fig. 16 for the relative orientation and respective role of the two water molecules. We adopt a dimer geometry that keeps the

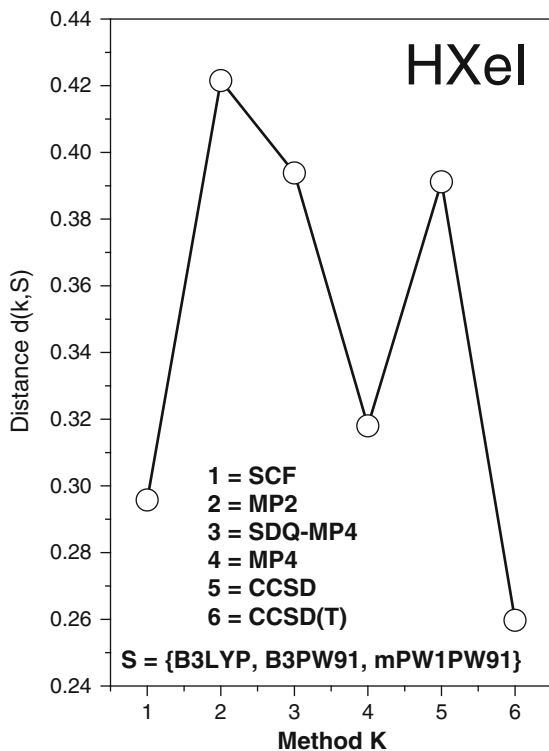


**Fig. 14** Minimum spanning tree (MST) for the theoretical descriptions of HXeI and subsequent clustering with a distance threshold  $D_T = 0.4$

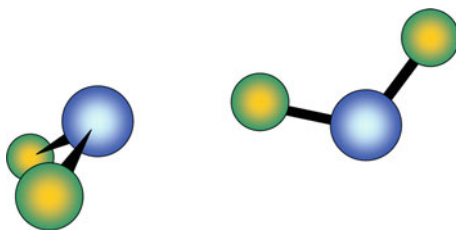
monomer  $\text{H}_2\text{O}$  geometry frozen [126]. Thus, comparison of supermolecule to monomer properties makes eminent sense. All calculations presented here have been performed with the water-molecule-specific [6s4p3d1f/4s3p1d] basis set. The ab initio methods employed are SCF, MP2, SDQ-MP4, MP4, CCSD, and CCSD(T). The DFT methods are B1LYP, B3LYP, B3PW91, mPW1PW91, mPW1PBE, PBEPBE, PBEPW91, HCTH, BHandH, and BHandHLYP.

In Table 5 we give all monomer and dimer value used in the analysis of the results. The  $\text{H}_2\text{O}$  monomer (M) values pertain to the molecular geometry of the moiety in the dimer. The per-monomer (PM) values are simply defined as  $\text{PM} = \text{AD}/2$  that is  $\bar{\alpha}(\text{H}_2\text{O})_2/2$  for the polarizability and  $\bar{\gamma}(\text{H}_2\text{O})_2/2$  for the hyperpolarizability. The properties of interest are the interaction mean (hyper) polarizabilities obtained by the formula  $\text{AD} - \text{AX} - \text{XD}$  (see above). A first and valuable observation here is the strong disagreement of ab initio and DFT methods on the (hyper)polarizability of the water dimer. Such behavior for the DFT methods has been noted and analyzed early enough [127]. We must emphasize at this point the essential difference of the BHandH and BHandHLYP DFT methods. Both seem to be very close to the ab initio ones and quite distinct from the other DFT approaches.

In Fig. 17 we show the histogram of interaction-induced mean polarizability for all methods. With the notable exception of the BHandHLYP method the interaction property  $\bar{\alpha}_{\text{int}}(\text{H}_2\text{O})_2$  is negative for all ab initio methods and positive for the DFT



**Fig. 15** Distance of the ab initio theoretical descriptions calculated with basis B5 from the reference DFT group  $S = \{B3LYP, B3PW91, mPW1PW91\}$



**A** = Proton Acceptor

**D** = Proton Donor

**Fig. 16** Relative orientation of the two interacting water monomers in  $(H_2O)_2$  and definition of the interaction-induced properties

$$P_{\text{int}} = P(AD) - P(AX) - P(XD)$$

**Table 5** Analysis of the interaction-induced electric properties for the theoretical equilibrium molecular geometry of the water dimer (AD)

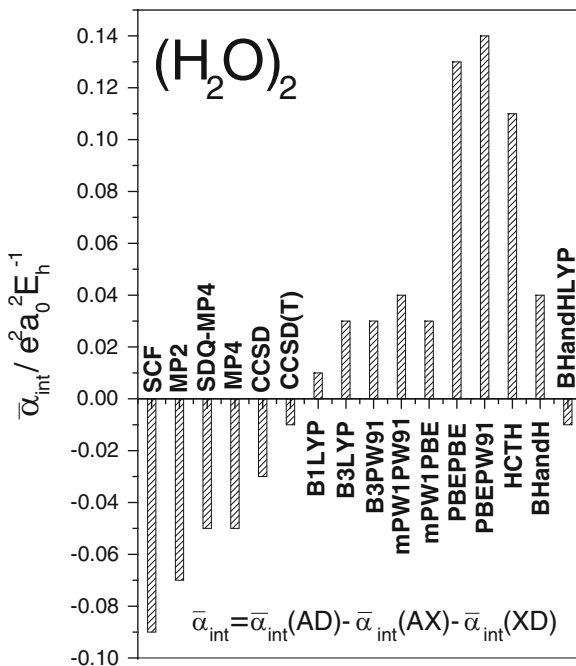
Property	Method	M	PM	AD	AX	XD	AD – AX – XD
$\bar{\alpha}$	SCF	8.48	8.44	16.89	8.49	8.49	<b>–0.09</b>
	MP2	9.85	9.81	19.62	9.84	9.85	<b>–0.07</b>
	SDQ-MP4	9.66	9.63	19.26	9.65	9.66	<b>–0.05</b>
	MP4	9.95	9.92	19.84	9.94	9.95	<b>–0.05</b>
	CCSD	9.58	9.56	19.13	9.58	9.58	<b>–0.03</b>
	<b>CCSD(T)</b>	<b>9.78</b>	<b>9.77</b>	<b>19.54</b>	<b>9.77</b>	<b>9.78</b>	<b>–0.01</b>
	B1LYP	9.85	9.87	19.74	9.86	9.86	<b>0.01</b>
	B3LYP	9.94	9.97	19.95	9.96	9.96	<b>0.03</b>
	B3PW91	9.78	9.81	19.62	9.79	9.79	<b>0.03</b>
	mPW1PW91	9.67	9.71	19.41	9.68	9.69	<b>0.04</b>
	mPW1PBE	9.67	9.71	19.41	9.69	9.69	<b>0.03</b>
	PBEPBE	10.54	10.62	21.25	10.56	10.56	<b>0.13</b>
	PBEPW91	10.53	10.62	21.24	10.55	10.55	<b>0.14</b>
	HCTH	10.28	10.35	20.71	10.30	10.30	<b>0.11</b>
	BHandH	9.27	9.30	18.60	9.28	9.28	<b>0.04</b>
	BHandHLYP	9.19	9.20	18.38	9.20	9.20	<b>–0.01</b>
$\bar{\gamma}$	SCF	975	957	1914	997	995	<b>–78</b>
	MP2	1742	1743	3486	1795	1790	<b>–99</b>
	SDQ-MP4	1698	1715	3429	1750	1746	<b>–67</b>
	MP4	1929	1952	3905	1992	1987	<b>–74</b>
	CCSD	1645	1671	3342	1690	1688	<b>–36</b>
	<b>CCSD(T)</b>	<b>1796</b>	<b>1834</b>	<b>3669</b>	<b>1847</b>	<b>1842</b>	<b>–20</b>
	B1LYP	2113	2309	4618	2249	2231	<b>137</b>
	B3LYP	2219	2449	4898	2366	2347	<b>185</b>
	B3PW91	2083	2260	4520	2215	2193	<b>113</b>
	mPW1PW91	1975	2145	4291	2102	2074	<b>115</b>
	mPW1PBE	1976	2116	4231	2079	2053	<b>99</b>
	PBEPBE	3090	3575	7150	3308	3266	<b>576</b>
	PBEPW91	3073	3557	7114	3288	3247	<b>578</b>
	HCTH	3115	3653	7306	3294	3307	<b>704</b>
	BHandH	1505	1571	3141	1569	1554	<b>19</b>
	BHandHLYP	1496	1529	3058	1537	1519	<b>1</b>

M  $\equiv$  Monomer, PM  $\equiv$  Per Molecule, AD  $\equiv$  (H<sub>2</sub>O)<sub>2</sub>  
Reference values in *bold*. All quantities in atomic units

ones. An almost extreme behavior is obvious for the PBEPBE, PBEPW91, and HCTH methods.

The interaction-induced mean hyperpolarizabilities  $\bar{\gamma}_{\text{int}}(\text{H}_2\text{O})_2$  are shown in Fig. 18. The very same pattern as in the case of  $\bar{\alpha}_{\text{int}}(\text{H}_2\text{O})_2$  is visible here as well. The PBEPBE, PBEPW91, and HCTH methods yield very large values for the

**Fig. 17** Histogram of the performance of theoretical methods on the calculation of the interaction-induced mean dipole polarizability of the water dimer

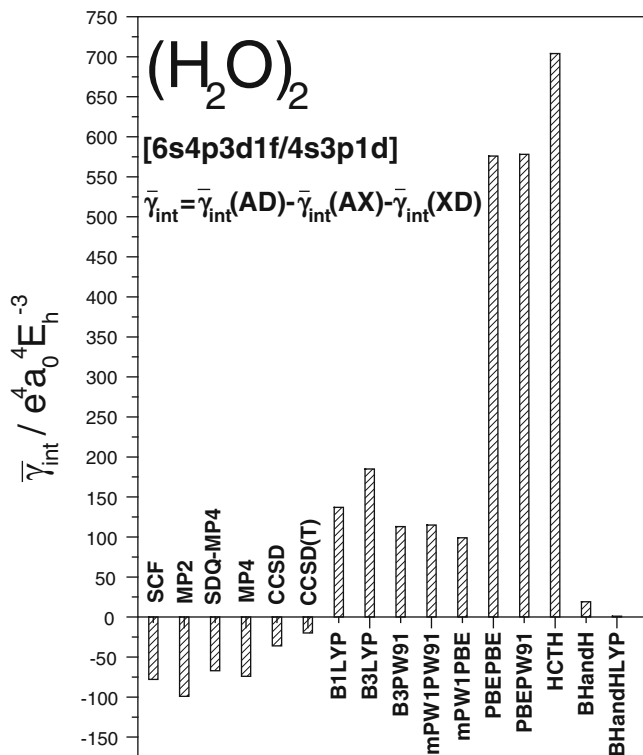


interaction hyperpolarizability. The BHandH and BHandHLYP methods again give surprisingly small  $\bar{\gamma}_{\text{int}}(\text{H}_2\text{O})_2$  values.

Last, in Fig. 19 we show the evolution of the DFT values for the interaction-induced hyperpolarizability for the monomer ( $M$ ) and the per-monomer ( $PM$ ) value. In fact the  $PM = \bar{\gamma}(\text{H}_2\text{O})_2/2$  and  $M = \bar{\gamma}(\text{H}_2\text{O})$  values determine the differential-per-monomer hyperpolarizability, defined as

$$\text{DHPM} = [\bar{\gamma}(\text{H}_2\text{O})_2 - 2\bar{\gamma}(\text{H}_2\text{O})]/2 = \bar{\gamma}(\text{H}_2\text{O})_2/2 - \bar{\gamma}(\text{H}_2\text{O}) \equiv PM - M.$$

In Fig. 19 we also show the reference  $M$  and  $P$  values calculated at the CCSD(T)/[6s4p3d1f/4s3p1d] level of theory. We observe that the sequence B3LYP, B3PW91, mPW1PW91, mPW1PBE displays a smooth convergence towards the reference CCSD(T) values. The PBEPBE, PBEPW91, and HCTH methods are characterized by very large  $M$  and  $PM$  values. Obviously, the BHandH and BHandHLYP methods appear as something of an anomaly here. Both the  $PM$  and  $M$  values for these two methods are lower than the respective reference CCSD(T) values.

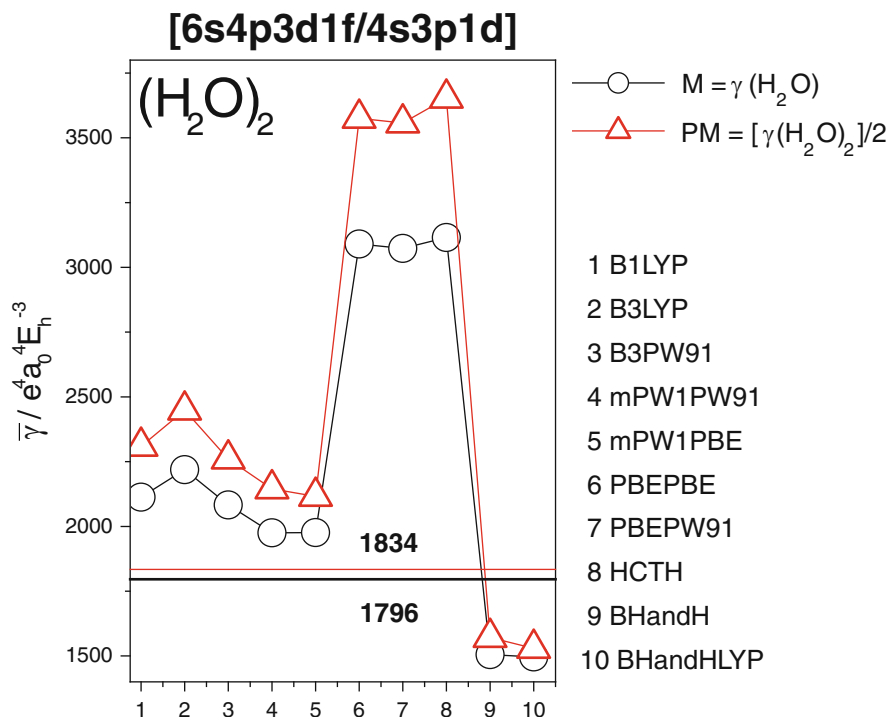


**Fig. 18** Histogram of the performance of theoretical methods on the calculation of the interaction-induced mean second hyperpolarizability of the water dimer

## 4 Final Remarks and Conclusions

We have investigated and closely analyzed the predictive capability of DFT methods in the calculation of electric polarizabilities and hyperpolarizabilities for three difficult cases: the sodium tetramer, the new compound HXeI, and the interaction (hyper)polarizability of the two moieties in the weakly bonded water dimer. In the first two cases we have shown that the ab initio and DFT methods offer a quite different picture of the electric polarizability and hyperpolarizability. In the case of  $(H_2O)_2$  we are led to two drastically divergent pictures as the two classes of methods clearly differ even in the sign of the interaction-induced mean polarizability and hyperpolarizability.

We show that it is probably a distinct advantage to talk about the quality or predictive capability of a set of DFT methods instead of trying to obtain a picture of the potential of a single one. This strongly corroborates our recent strategy to rely on a given set of DFT methods in order to form a clear idea about the DFT perspective in electric property calculations [115, 128, 129].



**Fig. 19** Mean second hyperpolarizability for the water monomer (M) and dimer (per monomer PM) for all DFT methods. Reference CCSD(T) values are also included

## References

1. Buckingham AD in Pullman B (ed) (1978) Intermolecular interactions: from diatomics to biopolymers, Wiley, New York, p. 1
2. Kaplan IG (2006) Intermolecular interactions. Wiley, Chichester
3. Hanna DC, Yuratich MA, Cotter D (1979) Nonlinear optics of free atoms and molecules. Springer, Berlin
4. Hartmann JM, Boulet C, Robert D (2008) Collisional effects on molecular spectra: laboratory experiments and models, consequences for applications. Elsevier, Amsterdam
5. Gray CG, Gubbins KE (1984) Theory of molecular fluids. Clarendon, Oxford
6. Gray CG, Gubbins KE, Joslin CG (2012) Theory of molecular fluids. volume 2: applications. Clarendon, Oxford
7. Berkowitz M, Parr RG (1988) J Chem Phys 88:2554
8. Vela A, Gázquez JL (1990) J Am Chem Soc 112:1490
9. Liu PH, Hunt KLC (1995) J Chem Phys 103:10597
10. Torrent-Sucarrat M, De Profit F, Geerlings P (2005) J Phys Chem A 109:6071
11. Donald KJ (2006) J Phys Chem A 110:2283
12. Matito E, Putz MV (2011) J Phys Chem A 115:12459–12462
13. Karelson K, Lobanov VS (1996) Phys Rev 96:1027–1043



14. Hansch C, Steinmetz WE, Leo AJ, Mekapati SB, Kurup A, Hoekman D (2003) *J Chem Inf Comput Sci* 43:120–125
15. Shelton DP, Rice JE (1994) *Chem Rev* 94:3–29
16. Maroulis G, Sana M, Leroy G (1981) *Int J Quant Chem* 19:43–60
17. Maroulis G (1988) *Int J Quant Chem* 24:185–190
18. Sordo JA (1988) *Comp Phys Comm* 113:85–104
19. Maroulis G, In: Sen KD (2002) *Reviews of modern quantum chemistry. A celebration of Robert G. Parr*, World Scientific, Singapore, pp 320–339
20. Buckingham AD (1967) *Adv Chem Phys* 12:107–142, and references therein
21. McLean AD, Yoshimine M (1967) *J Chem Phys* 47:1927–1935
22. Cohen HD, Roothaan CCJ (1965) *J Chem Phys* 43:S34–S39
23. Maroulis G, Bishop DM (1985) *Chem Phys Lett* 114:182–186
24. Bishop DM, Maroulis G (1985) *J Chem Phys* 82:2380–2391
25. Maroulis G, Thakkar AJ (1988) *J Chem Phys* 88:7623–7632
26. Maroulis G, Thakkar AJ (1988) *J Chem Phys* 89:7320–7323
27. Maroulis G (1991) *J Chem Phys* 94:1182–1190
28. Maroulis G (1998) *J Chem Phys* 108:5432–5448
29. Dykstra CE (1988) *Ab initio calculation of the structures and properties of molecules*. Elsevier, Amsterdam
30. Sauer SPA (2011) *Molecular electromagnetism: a computational chemistry approach*. Oxford University Press, Oxford
31. Maroulis G (ed) (2006) *Atoms, molecules and clusters in electric fields: theoretical approaches to the calculation of electric polarizability*. Imperial College Press, London
32. Maroulis G (ed) (2006) *Computational aspects of electric polarizability calculations: atoms, molecules and clusters*. Ios Press, Amsterdam
33. Papadopoulos MG, Sadlej AJ, Leszczynski J (eds) (2006) *Non-linear optical properties of matter. from molecules to condensed phases*. Springer, Berlin
34. Frisch MJ, Trucks GW, Schlegel HB et al (1998) *GAUSSIAN 98, Revision A.7*. Gaussian, Pittsburgh PA
35. Frisch MJ, Trucks GW, Schlegel HB et al (2004) *GAUSSIAN 03, Revision D.01*. Gaussian, Wallingford, CT
36. Szabo A, Ostlund NS (1982) *Modern quantum chemistry*. McMillan, New York
37. Wilson S (1984) *Electron correlation in molecules*. Clarendon, Oxford
38. Urban U, Cernusak I, Kellö V, Noga J (1987) *Methods Comput Chem* 1:117
39. Helgaker T, Jørgensen P, Olsen J (2000) *Molecular electronic-structure theory*. Wiley, Chichester
40. Wilson S (1987) *Adv Chem Phys* 67:439–500
41. Davidson ER, Feller D (1986) *Chem Rev* 86:681
42. Thakkar AJ, Koga T, Saito M, Hoffmeyer RE (1993) *S27*, 343–354
43. Koga T, Saito M, Hoffmeyer RE, Thakkar AJ (1994) *JMolStruct (THEOCHEM)* 306:249–260
44. Prascher BP, Woon DE, Peterson KA, Dunning TH Jr, Wilson AK (2011) *Theor Chem Acc* 128:69–82
45. Kellö V, Sadlej AJ (1995) *Theor Chim Acta* 91:353–371
46. Zuev MB, Nefediev SE (2004) *J Comput Methods Sci Eng* 4:481–491
47. Arruda PM, Canal Neto A, Jorge FE (2009) *Int J Quant Chem* 109:1189–1199
48. Baranowska AJ, Sadlej AJ (2010) *J Comput Chem* 31:552–560
49. de Berrêdo RC, Jorge FE, Jorge SS, Centoducatte R (2011) *Comput Theor Chem* 965:236–239
50. Liu SY, Dykstra CE (1987) *J Phys Chem* 91:1749–1754
51. Dykstra CE, Liu SY, Malik DJ (1989) *Adv Chem Phys* 75:37–111
52. Spackman MA (1989) *J Phys Chem* 93:7594–7603
53. Maroulis G, Bishop DM (1985) *J Phys B* 24:4675–4682

54. Maroulis G, Bishop DM (1986) *Mol Phys* 57:359–367
55. Maroulis G, Bishop DM (1986) *J Phys B* 19:369–377
56. Maroulis G, Bishop DM (1986) *Mol Phys* 58:273–283
57. Maroulis G, Makris C, Hohm U, Goebel D (1997) *J Phys Chem A* 101:953–956
58. Maroulis G (1994) *J Chem Phys* 101:4949–4955
59. Maroulis G (1992) *Chem Phys Lett* 199:250–256
60. Maroulis G, Pouchan C (1998) *Phys Rev* 57:2440–2447
61. Maroulis G (2003) *Chem Phys* 291:81–95
62. Maroulis G, Thakkar AJ (1991) *J Chem Phys* 95:9060–9064
63. Maroulis G (1996) *Chem Phys Lett* 259:654–660
64. Maroulis G (1996) *J Chem Phys* 105:8467–8468
65. Maroulis G, Xenides D (1999) *J Phys Chem A* 103:4590–4593
66. Maroulis G, Pouchan C (2003) *Phys Chem Chem Phys* 5:1992–1995
67. Karamanis P, Maroulis G, Pouchan C (2006) *J Chem Phys* 124:071101
68. Maroulis G, Karamanis P, Pouchan C (2007) *J Chem Phys* 126:154316
69. Birnbaum G (1985) *Phenomena induced by intermolecular interactions*. Plenum, New York
70. Tabisz GC, Neuman MN (eds) (1995) *Collision- and interaction-induced spectroscopy*. Kluwer, Dordrecht
71. Głaz W, Bancewicz T, Godet JL, Maroulis G, Haskopoulos A (2006) *Phys Rev A* 73:042708
72. Chrysos M, Ratchet F, Egorova NI, Kouzov AP (2007) *Phys Rev A* 75:012707
73. Chrysos M, Kouzov AP, Egorova NI, Ratchet F (2008) *Phys Rev Lett* 100:133007
74. Chrysos M, Dixneuf S, Ratchet F (2009) *Phys Rev A* 80:054701
75. Baranowska A, Fernández B, Rizzo A, Jansik B (2009) *Phys Chem Chem Phys* 11:9871–9883
76. El-Kader, El-Sheikh, Bancewicz T, Hellmenn R (2009) *J Chem Phys* 131:044314
77. Zvereva-Loëte N, Kalugina YN, Boudon V, Buldakov MA, Cherepanov VN (2010) *J Chem Phys* 133:184302
78. Buldakov MA, Cherepanov VN, Kalugina YN, Zvereva-Loëte N, Boudon V (2010) *J Chem Phys* 132:164304
79. Haskopoulos A, Maroulis G (2010) *J Phys Chem A* 114:8730–8741
80. Hartmann JM, Boulet C, Jacquemart D (2011) *J Chem Phys* 134:094316
81. Boys Sf, Bernardi F (1970) 19:553–556
82. Maroulis G, Haskopoulos A (2001) *Chem Phys Lett* 349:335–341
83. Maroulis G, Haskopoulos A (2002) *Chem Phys Lett* 358:64–70
84. Maroulis G, Haskopoulos A, Xenides D (2004) *Chem Phys Lett* 396:59–65
85. Haskopoulos A, Xenides D, Maroulis G (2005) *Chem Phys* 309:271–275
86. Maroulis G, Haskopoulos A, Głaz W, Bancewicz T, Godet JL (2006) *Chem Phys Lett* 428:28–33
87. Bancewicz T, Głaz W, Godet JL, Maroulis G (2008) *J Chem Phys* 129:124306
88. Haskopoulos A, Maroulis G (2010) *Chem Phys Lett* 367:127–135
89. Xenides D, Hantzis A, Maroulis G (2011) *Chem Phys* 382:80–87
90. Chantzis A, Maroulis G (2011) *Chem Phys Lett* 507:42–47
91. Głaz W, Godet JL, Haskopoulos A, Bancewicz T, Maroulis G (2011) *Phys Rev A* 84:012503
92. Maroulis G (1995) *Int J Quant Chem* 55:173–180
93. Varmuza K (1980) *Pattern recognition in chemistry*. Springer, Heidelberg
94. Maroulis G (1999) *J Chem Phys* 111:583–591
95. Xenides D (2007) *J Mol Struct (THEOCHEM)* 804:41–46
96. Christodouleas C, Xenides D, Simos TE (2010) *J Comput Chem* 31:412–420
97. Chartrand G, Lesniak L (1986) *Graphs and digraphs*. Wadsworth, Belmont
98. Spath H (1980) *Cluster analysis algorithms*. Ellis Horwood, Chichester
99. Antoine R, Rayane D, Allouche AR, Aubert-Frecon M, Benichou E, Dalby FW, Dugourd PH, Broyer M, Guet C (1999) *J Chem Phys* 110:5568–5577
100. Tikhonov G, Kasperovich K, Wong K, Kresin VV (2001) *Phys Rev A* 64:063202

101. Bowlan J, Liang A, de Heer WA (2011) *Phys Rev Lett* 106:043401
102. Calaminici P, Jug K, Köster A (1999) *J Chem Phys* 111:4613–4620
103. Jiemchooraj A, Norman P, Semelius B (2006) *J Chem Phys* 125:124306
104. Aguado A, Vega A, Balbás LC (2011) *Phys Rev B* 84:165450
105. Ekstrom CR, Schmiedmayer J, Chapman MS, Hammond TD, Pritchard DE (1995) *Phys Rev A* 51:3883–3888
106. Thakkar AJ, Lupinetti C (2005) *Chem Phys Lett* 402:270–273
107. Maroulis G, Begué D, Pouchan C (2003) *J Chem Phys* 119:794–797
108. Maroulis G (2004) *J Chem Phys* 121:10519–10524
109. Papadopoulos MG, Reis H, Avramopoulos A, Erkoc S, Amirouche L (2005) *J Phys Chem B* 109:18822–18830
110. Xenides D, Maroulis G (2007) *J Comput Methods Sci Eng* 7:431–442
111. Karamanis P, Maroulis G (2005) *Match Commun Math Comput Chem* 53:269–282
112. Maroulis G, Haskopoulos (2009) *J Comput Theor Nanosci* 6:418–427
113. Maroulis G (2009) *J Comput Theor Nanosci* 6:886–893
114. Maroulis G (2011) *Theor Chem Acc* 129:437–445
115. Karamanis P, Maroulis G (2011) *J Phys Org Chem* 24:588–599
116. Petterson M, Lundell J, Räsänen M (1995) *J Chem Phys* 102:6423–6431
117. Petterson M, Lundell J, Räsänen M (1999) *Eur J Inorg Chem* 729–737 and references therein
118. Buck U, Farnik M (2006) *Int Rev Phys Chem* 25:583–612, and references therein
119. Buck U (2002) *J Phys Chem A* 106:10049–10062
120. Nahler NH, Baumfalk R, Buck U, Bihary Z, Benny Gerber R, Friedrich B (2003) *J Chem Phys* 119:224–231
121. Maroulis G (2008) *J Chem Phys* 129:044314
122. Maroulis G (2000) *J Chem Phys* 113:1813–1820
123. Rodriguez J, Laria D, Marceca EJ, Estrin DA (1999) *J Chem Phys* 110:9039–9047
124. Ghanty TK, Ghosh SK (2003) *J Chem Phys* 118:8547–8550
125. Maroulis G (2012) *Int J Quant Chem* 112:2231–2241
126. Frisch MJ, Pople JA, Del Bene JE (1985) *J Phys Chem* 89:3664–3669
127. Champagne B, Perpète EA, Jacquemin D, van Gisbergen SJA, Baerends EJ, Soubra-Ghaoui C, Robins KA, Kirtman B (2000) *J Phys Chem A* 104:4755–4763
128. Maroulis G (2012) *Chem Phys Lett* 525–526:49–53
129. Maroulis G, Haskopoulos A (2012) *Comp Theor Chem* 988:34–41