MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. 53 (2005) 269-282

ISSN 0340 - 6253

Comparison of conventional ab initio and DFT quantum chemical methods. The electric moments and (hyper)polarizability of H-C≡C-C≡C-H as a test case

Panaghiotis Karamanis and George Maroulis *
Department of Chemistry, University of Patras,
GR-26500 Patras, Greece

(Received 30 October 2003. Revised 10 April 2004. Accepted 12 April 2004)

Abstract

We report an analysis of the performance of widely used density functional theory methods in theoretical predictions of electric multipole moments, polarizabilities and hyperpolarizabilities. Our method relies on graph theoretic arguments and pattern recognition techniques. Theoretical descriptions of the molecular properties of diacetylene (H-C≡C-C≡C-H), obtained via Møller-Plesset perturbation theory, Coupled-cluster methods and a collection of density functional theory methods, are analyzed and compared. Our analysis brings forth essential characteristics of the relative merit of the various quantum chemical methods in molecular property calculations.

1 Introduction

Density functional theory (DFT) has emerged as a source of powerful computational methods in modern quantum chemistry [1]. Conventional ab initio quantum chemical methods are known to be of high predictive capability [2,3]. High-level approaches as Møller-Plesset Perturbation Theory (MP) and Coupled Cluster (CC) techniques yield very reliable atomic and molecular property values but are computationally expensive and their applicability is limited to relatively small systems. On the other hand DFT based

*Corresponding author. Electronic address: maroulis@upatras.gr

methods are less demanding. As the possibility of applying rigorous quantum chemical treatments to large and very large molecules has never been lost sight of, this advantage of DFT methods has prompted systematic investigations on their predictive potential [4-6]. The main objective of such investigations is to provide extensive comparisons of the performance of conventional ab initio and DFT on classes of important problems. These include the determination of molecular geometries, spectroscopic constants, electric properties et al. The wealth of the collected results has constituted a firm basis for the creation of new DFT methods [7].

The rigorous evaluation of the performance of theoretical methods on atomic/molecular property calculations has been the object of many studies [8,9]. In previous we proposed an information theoretic approach to this important problem [10,11]. A rather more general method, based on graph theoretic arguments and pattern recognition techniques has since been proposed [12] and applied to the analysis of the performance of well-defined theoretical methods [13-15].

In this paper we propose an analysis of the relative merit of conventional ab initio and widely used DFT methods. Our analysis focuses on the performance of a class of MP, CC and DFT methods on the calculation of electric properties of a linear molecule with two triple carbon-carbon bonds, diacetylene (H-C≡C-C≡C-H). This is central to a major project on the linear and nonlinear polarizability of substituted diacetylenes, in progress in our laboratory [16-18]. The assessment of the predictive capability of commonly available and widely used DFT methods is highly desirable, given the current interest in applying quantum chemical approaches to the design of NLO materials for advanced technologies [19,20].

2 Theoretical considerations

We follow Buckingham's conventions and notation in our definition of the molecular properties of interest in this work [21]. The energy of an uncharged molecule in a general, static

weak electric field can be expanded as

$$E^p \equiv E^p(F_\alpha F_{\alpha\beta}, F_{\alpha\beta\gamma}F_{\alpha\beta\gamma\delta}, \ldots) = E^0 - \mu_\alpha F_\alpha - (1/3)\Theta_{\alpha\beta}F_{\alpha\beta}$$
$$-(1/15)\Omega_{\alpha\beta\gamma}F_{\alpha\beta\gamma} - (1/105)\Phi_{\alpha\beta\gamma\delta}F_{\alpha\beta\gamma\delta} + \ldots - (1/2)\alpha_{\alpha\beta}F_\alpha F_\beta$$
$$-(1/3)A_{\alpha,\beta\gamma}F_\alpha F_{\beta\gamma} - (1/6)C_{\alpha\beta,\gamma\delta}F_{\alpha\beta}F_{\gamma\delta} - (1/15)E_{\alpha,\beta\gamma\delta}F_\alpha F_{\beta\gamma\delta} + \ldots$$
$$-(1/6)\beta_{\alpha\beta\gamma}F_\alpha F_\beta F_\gamma - (1/6)B_{\alpha\beta,\gamma\delta}F_\alpha F_\beta F_{\gamma\delta} + \ldots$$
$$-(1/24)\gamma_{\alpha\beta\gamma\delta}F_\alpha F_\beta F_\gamma F_\delta + \ldots \qquad (1)$$

In this expansion the subscripts denote Cartesian components and a repeated subscript implies summation over x, y and z. F_α F_α, etc. are the field, field gradient, etc. at the origin. E^0, μ, Θ_β Ω_β and $\Phi_{\beta\delta}$re the energy and the dipole, quadrupole, octopole and hexadecapole moment of the free molecule and the second, third and fourth-order properties are the dipole and quadrupole polarizabilities and hyperpolarizabilities α_β β_β, $\gamma_{\beta\delta}$ $A_{\alpha\beta}$, $C_{\alpha,\gamma}$, $E_{\alpha\beta\delta}$nd $B_{\alpha,\gamma}$. For a linear centrosymmetric molecule as diacetylene $\mu = \Omega_\beta = \beta_\beta = A_{\alpha\beta}$ [21]. The independent components of the $\Theta_\beta$quadrupole moment), $\Phi_{\beta\delta}$hexadecapole), $\alpha_\beta$dipole polarizability) and $\gamma_{\beta\delta}$dipole hyperpolarizability) tensors adopted for this molecule are the same as in previous work [17]. The invariants of interest in this paper are defined with z as the reference molecular axis as (observable in principle) mean values and anisotropies:

$$\begin{aligned}
\overline{\alpha} &= (\alpha_{zz} + 2\,\alpha_{xx})/3 \\
\Delta &= \alpha_{zz} - \alpha_{xx} \\
\overline{\gamma} &= (3\gamma_{zzzz} + 8\gamma_{xxxx} + 12\gamma_{xxzz})/15 \\
\Delta_1\gamma &= 3\gamma_{zzzz} - 4\gamma_{xxxx} + 3\gamma_{xxzz} \\
\Delta_2\gamma &= \gamma_{zzzz} + \gamma_{xxxx} - 6\gamma_{xxzz}
\end{aligned} \qquad (2)$$

Detailed presentations of our approach to the calculation of the above properties, including

the computational philosophy underlying the design of basis sets, may be found in previous work [13,17,22].

The conventional methods used in this work are self-consistent field (SCF), second (MP2) and fourth-order (MP4) Møller-Plesset Perturbation Theory and CCSD(T) (single and double excitation coupled cluster theory including an estimate of connected triple excitations by a perturbational treatment) [2,3]. The DFT methods are B3LYP, B3PW91, B3PW86 and MPW1PW91, as implemented in the GAUSSIAN 98 set of programmes [23].

Our analysis of the performance of methods is based upon comparisons of string of numbers that is calculated molecular properties [10,11]. Let $Q_{m\alpha}$e the value of the molecular property αtaking values in the index set I_A) calculated with quantum chemical method m (taking values in I_M)

$$Q_{m\alpha}, m \in I_M(methods),\ \alpha \in I_A(properties) \tag{3}$$

The theoretical description of the molecule associated with method i, TD_i, is a collection of values

$$TD_i = \{Q_{i\alpha}, \alpha \in I_A\} \tag{4}$$

We define a distance function in the space of all theoretical descriptions (TD) using the Minkowski metric. The distance between TD_i and TD_j is defined as

$$D_{ij} = \left(\sum_{\alpha} \frac{|Q_{i\alpha} - Q_{j\alpha}|^p}{\left(\max_{ij} |Q_{i\alpha} - Q_{j\alpha}| \right)^p} \right)^{1/p}, p \geq 1 \tag{5}$$

Equation (4) then defines distance or proximity on TD. Similarity between two theoretical descriptions TD_i and TD_j is subsequently defined as

$$S_{ij} = 1 - \frac{D_{ij}}{\max_{ij} D_{ij}}, \qquad 0 \leq S_{ij} \leq 1 \tag{6}$$

Let G be the graph with vertex set the whole of TD, $V(G) = \{TD_i, i \in I_M\}$. and as the edge set E(G), a subset of the Cartesian product TD X TD . A generalized distance can be defined on G by assigning to each edge e = $\{TD_i, TD_j$ a weight equal to the distance D_{ij}. The **minimum spanning tree** (MST) of G is a spanning tree for which the sum of the weights of the edges is minimal [24]. In a final step clustering is achieved by removing from MST all edges with weights greater than a given threshold D_T. This method of clustering is called **single linkage cluster analysis** (SLCA) [25].

3 Results and discussion

All molecular property values in Tables 1 and 3 have been calculated with three large basis sets Q0 ≡[9s6p2d/6s2p], Q1 ≡[9s6p3d1f/6s3p1d] and Q2 ≡ [9s6p4d2f/6s4p2d]. A full description of Q2 has been given elsewhere [26]. Basis Q0 contains only the optimised d-GTF (on C) and p-GTF (on H). Basis Q1 is a truncated version of Q2: The α_βnd $\gamma_{\beta\delta}$alues obtained with Q2 have been taken from previous work [17], where all computational details of interest to the calculations reported here are presented in sufficient detail. Atomic units are used throughout this paper [27].

Table 1 shows theoretical descriptions pertaining to the molecular properties E^0, Θ Φ $\bar{\alpha}$, Δ, $\bar{\gamma}$, $\Delta\gamma$nd $\Delta\gamma$ We use the symbol B3LYP/Q0 to denote the theoretical description obtained with method B3LYP and basis set Q0. It is fairly obvious that the calculated electric property values are strongly method and basis set dependent. This is most apparent for the hyperpolarizability where even the sign of $\Delta\gamma$s not certain. Table 2 contains similarity values for the theoretical descriptions in Table 1. The most dissimilar TD are SCF/Q0 (5) and B3LYP/Q2 (13), as S(5,13) ≡(SCF/Q0,B3LYP/Q2) = 0. Table 2 contains a wealth of information on the relative merit of the employed methods. The similarity of SCF/Q0, SCF/Q1 and SCF/Q2 values is evidenced from S(5,11)=0.8368, S(5,17)=0.7989 and S(11,17)=0.9082. At the MP2 level S(6,12)=0.7901, S(6,18)=0.7611

Table 1. Energies, electric multipole moments, polarizabilities and hyperpolarizabilities of diacetylene calculated with conventional ab initio and density functional theory methods. All quantities in atomic units.

No	Method	Basis set	$-E^0/E_h$	Θ	Φ	$\bar{\alpha}$	$\Delta\alpha$	$\bar{\gamma}$	$\Delta_1\gamma$	$\Delta_2\gamma$
1	B3LYP	Q0 [a]	153.546640034	9.7497	502.83	51.96	61.98	14947	81656	1902
2	B3PW91	Q0	153.435730578	9.9207	500.27	51.15	61.95	13514	73204	4015
3	B3PW86	Q0	153.958061587	9.9678	501.71	51.33	61.91	12989	71737	4569
4	MPW1PW91	Q0	153.454825517	10.0613	506.70	50.89	61.54	13717	75564	4538
5	SCF	Q0	152.543853149	11.1155	558.51	50.23	56.93	11141	56389	3478
6	MP2	Q0	153.099544508	9.4333	490.73	48.60	54.17	14135	86952	9020
7	B3LYP	Q1 [b]	153.546840830	9.9207	501.88	52.18	61.81	17827	76750	-3574
8	B3PW91	Q1	153.444925092	9.9678	505.76	51.56	61.45	14031	64714	3572
9	B3PW86	Q1	153.967531833	10.0613	508.41	51.68	61.41	15698	67542	-1734
10	MPW1PW91	Q1	153.463542034	9.9207	510.82	51.32	61.08	17873	73387	-6977
11	SCF	Q1	152.544095580	11.1042	556.75	50.41	56.89	13227	55394	1709
12	MP2	Q1	153.107163076	9.4152	491.37	49.06	54.21	16449	83575	5785
13	B3LYP	Q2 [c]	153.547715532	9.7252	500.41	52.34	61.69	18119	74141	-2990
14	B3PW91	Q2	153.472831971	9.9622	507.75	51.62	61.26	16273	67476	-1693
15	B3PW86	Q2	153.995674475	10.0131	509.80	51.64	61.26	15745	66002	-1146
16	MPW1PW91	Q2	153.490306510	10.0858	513.03	51.47	60.90	16126	66490	-1471
17	SCF	Q2	152.545208950	11.0988	555.54	50.61	56.85	13136	49375	657
18	MP2	Q2	153.120704889	9.5132	493.21	49.17	54.45	16227	78541	5178

[a] Q0 ≡ [9s6p2d/6s2p]. [b] Q1 ≡ [9s6p3d1f/6s3p1d]. [c] Q2 ≡ [9s6p4d2f/6s4p2d].

Table 2. Similarity matrix obtained for the theoretical descriptions shown in Table 1

Method		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
B3LYP	1	1.0000	0.7900	0.7239	0.7694	0.1024	0.2545	0.7162	0.7385	0.7071	0.6099	0.1455	0.3149	0.7012	0.7392	0.6947	0.7089	0.1174	0.3458
B3PW91	2	0.7900	1.0000	0.8110	0.9137	0.2213	0.3159	0.5799	0.8571	0.6703	0.5228	0.2440	0.3414	0.5649	0.7052	0.6668	0.6979	0.2184	0.3855
B3PW86	3	0.7239	0.8110	1.0000	0.7952	0.1527	0.2447	0.5257	0.7770	0.7051	0.4589	0.1664	0.2606	0.5121	0.6375	0.7085	0.6380	0.1453	0.3043
MPW1MPW91	4	0.7694	0.9137	0.7952	1.0000	0.2642	0.3463	0.5673	0.8226	0.6628	0.5222	0.2873	0.3686	0.5439	0.6920	0.6596	0.6989	0.2541	0.4118
SCF	5	0.1024	0.2213	0.1527	0.2642	1.0000	0.0724	0.0283	0.2666	0.1305	0.0841	0.8368	0.0632	0.0000	0.1882	0.1336	0.2396	0.7989	0.1228
MP2	6	0.2545	0.3159	0.2447	0.3463	0.0724	1.0000	0.1220	0.2602	0.1487	0.1349	0.0803	0.7901	0.1181	0.2032	0.1545	0.2104	0.0305	0.7611
B3LYP	7	0.7162	0.5799	0.5257	0.5673	0.0283	0.1220	1.0000	0.5989	0.7307	0.8149	0.1144	0.2441	0.9227	0.7974	0.7104	0.7543	0.0987	0.2803
B3PW91	8	0.7385	0.8571	0.7770	0.8226	0.2666	0.2602	0.5989	1.0000	0.7211	0.5493	0.3032	0.3067	0.5946	0.7647	0.7257	0.7672	0.2903	0.3618
B3PW86	9	0.7071	0.6703	0.7051	0.6628	0.1305	0.1487	0.7307	0.7211	1.0000	0.6949	0.1938	0.2354	0.7128	0.8209	0.9639	0.8222	0.1850	0.2861
MPW1MPW91	10	0.6099	0.5228	0.4589	0.5222	0.0841	0.1349	0.8149	0.5493	0.6949	1.0000	0.1803	0.2679	0.7841	0.7774	0.6782	0.7601	0.1677	0.3101
SCF	11	0.1455	0.2440	0.1664	0.2873	0.8368	0.0803	0.1144	0.3032	0.1938	0.1803	1.0000	0.1070	0.0873	0.2647	0.1967	0.3176	0.9082	0.1693
MP2	12	0.3149	0.3414	0.2606	0.3686	0.0632	0.7901	0.2441	0.3067	0.2354	0.2679	0.1070	1.0000	0.2459	0.3051	0.2411	0.3088	0.0638	0.9184
B3LYP	13	0.7012	0.5649	0.5121	0.5439	0.0000	0.1181	0.9227	0.5946	0.7128	0.7841	0.0873	0.2459	1.0000	0.7856	0.6996	0.7352	0.0752	0.2822
B3PW91	14	0.7392	0.7052	0.6375	0.6920	0.1882	0.2032	0.7974	0.7647	0.8209	0.7774	0.2647	0.3051	0.7856	1.0000	0.8119	0.9353	0.2545	0.3572
B3PW86	15	0.6947	0.6668	0.7085	0.6596	0.1336	0.1545	0.7104	0.7257	0.9639	0.6782	0.1967	0.2411	0.6996	0.8119	1.0000	0.8167	0.1888	0.2930
MPW1MPW91	16	0.7089	0.6979	0.6380	0.6989	0.2396	0.2104	0.7543	0.7672	0.8222	0.7601	0.3176	0.3088	0.7352	0.9353	0.8167	1.0000	0.3067	0.3642
SCF	17	0.1174	0.2184	0.1453	0.2541	0.7989	0.0305	0.0987	0.2903	0.1850	0.1677	0.9082	0.0638	0.0752	0.2545	0.1888	0.3067	1.0000	0.1303
MP2	18	0.3458	0.3855	0.3043	0.4118	0.1228	0.7611	0.2803	0.3618	0.2861	0.3101	0.1693	0.9184	0.2822	0.3572	0.2930	0.3642	0.1303	1.0000

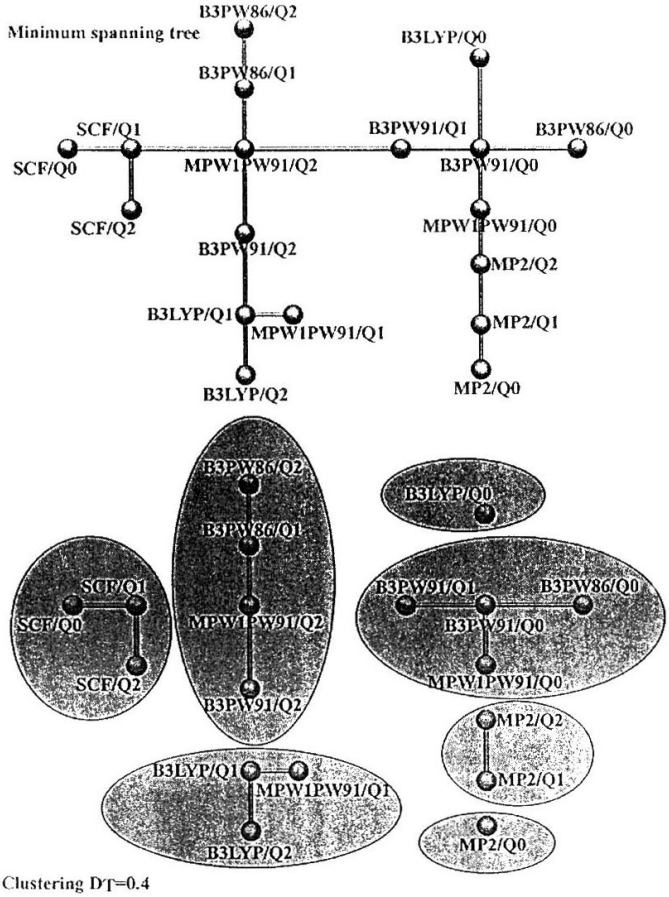

Figure 1: Minimum spanning tree (MST) and clustering in the space of theoretical descriptions shown in Table 1

Table 3. Energies, polarizabilities and hyperpolarizabilities of diacetylene calculated with conventional **ab initio** and density functional theory methods. All quantities in atomic units.

N°	Method	Basis set	Energy/E_h	$\bar{\alpha}$	$\Delta\alpha$	$\bar{\gamma}$	$\Delta_1\gamma$	$\Delta_2\gamma$
1	B3LYP	Q0	153.546640034	51.96	61.98	14947	81656	1902
2	B3PW91	Q0	153.435730578	51.15	61.95	13514	73204	4015
3	B3PW86	Q0	153.958061587	51.33	61.91	12989	71737	4569
4	MPW1PW91	Q0	153.454825517	50.89	61.54	13717	75564	4538
5	SCF	Q0	152.543853149	50.23	56.93	11141	56389	3478
6	MP2	Q0	153.099544508	48.60	54.17	14135	86952	9020
7	MP4	Q0	152.543853149	50.23	56.93	11141	56389	3478
8	CCSD(T)	Q0	153.099544508	48.60	54.17	14135	86952	9020
9	B3LYP	Q1	153.546840830	52.18	61.81	17827	76750	-3574
10	B3PW91	Q1	153.444925092	51.56	61.45	14031	64714	3572
11	B3PW86	Q1	153.967531833	51.68	61.41	15698	67542	-1734
12	MPW1PW91	Q1	153.463542034	51.32	61.08	17873	73387	-6977
13	SCF	Q1	152.544095580	50.41	56.89	13227	55394	1709
14	MP2	Q1	153.107163076	49.06	54.21	16449	83575	5785
15	B3LYP	Q2	153.547715532	52.34	61.69	18119	74141	-2990
16	B3PW91	Q2	153.472831971	51.62	61.26	16273	67476	-1693
17	B3PW86	Q2	153.995674475	51.64	61.26	15745	66002	-1146
18	MPW1PW91	Q2	153.490306510	51.47	60.90	16126	66490	-1471
19	SCF	Q2	152.545208950	50.61	56.85	13136	49375	657
20	MP2	Q2	153.120704889	49.17	54.45	16227	78541	5178

Table 4. Similarity matrix obtained for the theoretical descriptions shown in Table 3.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1.0000	0.7678	0.6914	0.7615	0.2101	0.1617	0.3389	0.1609	0.6782	0.7092	0.6831	0.5599	0.2637	0.2312	0.6565	0.7114	0.6650	0.6973	0.2133	0.2609
2	0.7678	1.0000	0.7831	0.9330	0.3744	0.2384	0.4387	0.2760	0.5171	0.8436	0.6299	0.4572	0.3973	0.2680	0.5045	0.6669	0.6262	0.6741	0.3464	0.3113
3	0.6914	0.7831	1.0000	0.7705	0.2490	0.1590	0.3446	0.2044	0.4541	0.7453	0.6666	0.3811	0.2575	0.1771	0.4448	0.5853	0.6716	0.5959	0.2189	0.2200
4	0.7615	0.9330	0.7705	1.0000	0.3723	0.2949	0.4916	0.3242	0.5071	0.7995	0.6124	0.4525	0.3955	0.3216	0.4921	0.6475	0.6096	0.6577	0.3366	0.3611
5	0.2101	0.3744	0.2490	0.3723	1.0000	0.3100	0.4565	0.4419	0.0778	0.4031	0.1694	0.1159	0.8126	0.2931	0.0766	0.2714	0.1749	0.2982	0.7699	0.3585
6	0.1617	0.2384	0.1590	0.2949	0.3100	1.0000	0.7903	0.8071	0.0139	0.1832	0.0602	0.0363	0.3091	0.7587	0.0000	0.1172	0.0655	0.1439	0.2157	0.7274
7	0.3389	0.4387	0.3446	0.4916	0.4565	0.7903	1.0000	0.7859	0.1653	0.3838	0.2270	0.1742	0.4537	0.7085	0.1526	0.2899	0.2321	0.3162	0.3583	0.7243
8	0.1609	0.2760	0.2044	0.3242	0.4419	0.8071	0.7859	1.0000	0.0175	0.2430	0.1055	0.0561	0.4278	0.6892	0.0074	0.1577	0.1145	0.1901	0.3444	0.7144
9	0.6782	0.5171	0.4541	0.5071	0.0778	0.0139	0.1653	0.0175	1.0000	0.5395	0.6989	0.8030	0.1900	0.1575	0.9424	0.7732	0.6751	0.7405	0.1603	0.1921
10	0.7092	0.8436	0.7453	0.7995	0.4031	0.1832	0.3838	0.2430	0.5395	1.0000	0.6808	0.4824	0.4489	0.2380	0.5430	0.7293	0.6860	0.7419	0.4180	0.2937
11	0.6831	0.6299	0.6666	0.6124	0.1694	0.0602	0.2270	0.1055	0.6989	0.6808	1.0000	0.6523	0.2484	0.1639	0.6978	0.7963	0.9637	0.7991	0.2291	0.2139
12	0.5599	0.4572	0.3811	0.4525	0.1159	0.0363	0.1742	0.0561	0.8030	0.4824	0.6523	1.0000	0.2415	0.1944	0.7802	0.7453	0.6306	0.7299	0.2171	0.2351
13	0.2637	0.3973	0.2575	0.3955	0.8126	0.3091	0.4537	0.4278	0.1900	0.4489	0.2484	0.2415	1.0000	0.3566	0.1938	0.3756	0.2539	0.4022	0.8949	0.4289
14	0.2312	0.2680	0.1771	0.3216	0.2931	0.7587	0.7085	0.6892	0.1575	0.2380	0.1639	0.1944	0.3566	1.0000	0.1481	0.2387	0.1687	0.2642	0.2767	0.9135
15	0.6565	0.5045	0.4448	0.4921	0.0766	0.0000	0.1526	0.0074	0.9424	0.5430	0.6978	0.7802	0.1938	0.1481	1.0000	0.7764	0.6786	0.7439	0.1681	0.1854
16	0.7114	0.6669	0.5853	0.6475	0.2714	0.1172	0.2899	0.1577	0.7732	0.7293	0.7963	0.7453	0.3756	0.2387	0.7764	1.0000	0.7846	0.9585	0.3515	0.2908
17	0.6650	0.6262	0.6716	0.6096	0.1749	0.0655	0.2321	0.1145	0.6751	0.6860	0.9637	0.6306	0.2539	0.1687	0.6786	0.7846	1.0000	0.7920	0.2361	0.2203
18	0.6973	0.6741	0.5959	0.6577	0.2982	0.1439	0.3162	0.1901	0.7405	0.7419	0.7991	0.7299	0.4022	0.2642	0.7439	0.9585	0.7920	1.0000	0.3780	0.3188
19	0.2133	0.3464	0.2189	0.3366	0.7699	0.2157	0.3583	0.3444	0.1603	0.4180	0.2291	0.2171	0.8949	0.2707	0.1681	0.3515	0.2361	0.3780	1.0000	0.3467
20	0.2609	0.3113	0.2200	0.3611	0.3585	0.7274	0.7243	0.7144	0.1921	0.2937	0.2139	0.2351	0.4289	0.9135	0.1854	0.2908	0.2203	0.3188	0.3467	1.0000

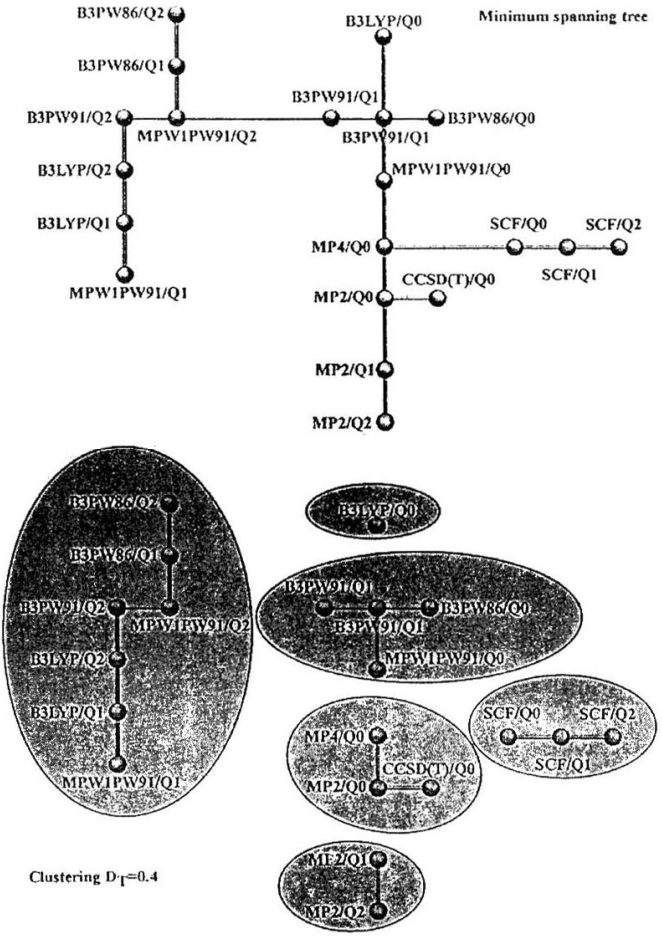

Figure 2: Minimum spanning tree (MST) and clustering in the space of theoretical descriptions shown in Table 3

and S(12,18)=0.9184. Similar observations are possible for the relative merits of the DFT methods. In figure 1 we show the minimum spanning tree (MST) for the space of TD in Table 1. We show also a clustering with threshold distance $D_T = 0.4$. The complexity of the MST is significantly reduced by the clustering. We observe that DFT and conventional ab initio belong to different clusters. There is also a clear separation of SCF and MP2 descriptions.

Table 3 shows theoretical descriptions pertaining over a subset of the available molecular properties: E^0, $\overline{\alpha}$, Δ, $\overline{\gamma}$, $\Delta\gamma$nd $\Delta\gamma$ For these properties we have also included higher order ab initio methods as MP4 and CCSD(T). In this sense, the table includes a large variety of theoretical descriptions obtained with the Q0 basis set. Table 4 reveals that the most distant or dissimilar descriptions are MP2/Q0 and B3LYP/Q2, S(6,15)=0. The DFT/Q0 descriptions bear little resemblance to the presumably most accurate CCSD(T)/Q0 one. For the ab initio TD obtained with Q0 we have S(5,8)=0.4419, S(6,8)=0.8071 and S(7,8)=0.7859. Thus MP2/Q0 is closer to CCSD(T)/Q0 than the higher MP4/Q0. The MST in figure 2 is quite analogous to the previous one. Clustering with D_T=0.4 produces clusters which reveal the fundamental dissimilarity of DFT and ab initio methods. Three clusters contain all DFT related TD. All SCF descriptions are included in one cluster while the CCSD(T)/Q0 clusters with the closely related MP2/Q0 and MP4/Q0. Another cluster contains MP2/Q1 and MP2/Q2. It is worth noticing that all MP2 descriptions appear in the part of the MST.

Conclusions

We have shown that a rigorous approach based on graph theoretic principles and pattern recognition techniques can be used to present and analyze in some depth theoretical descriptions obtained via quantum chemical methods. Thus the overall performance of methods and/or basis sets is more easily brought to light. Similarity or dissimilarity between theoretical descriptions is defined on metric considerations and leads in an unbiased

way to the classification of the performance of the methods.

Acknowledgment

The authors gratefully acknowledge support of this work by the University of Patras through a grant from the CARATHEODORY project (Grant 02449).

References

[1] W. Koch and M.C. Holthausen, *A Chemist's Guide to Density Functional Theory*, 2nd Edition (Wiley, Chichester, 2001).

[2] A. Szabo and N.S. Ostlund, *Modern Quantum Chemistry* (MacMillan, New York, 1982).

[3] T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic-Structure Theory* (Wiley, Chichester, 2000).

[4] S.J.A. Van Gisbergen, V.P. Osinga, O.V. Gritsenko, R.Vanleeuwen, J.G.Snijders and E.J.Baerends, J. Chem. Phys. 105, 3142 (1996).

[5] V.P. Osinga, S.J.A. Van Gisbergen, J.G. Snijders and E.J. Baerends, J. Chem. Phys. 106, 5091 (1997).

[6] B. Champagne, E. A. Perpète, D. Jacquemin, S. J. A. van Gisbergen, E. J. Baerends, C. Soubra-Ghaoui, K. A. Robins and B. Kirtman, J.Phys.Chem. A 104, 4755 (2000).

[7] S.J.A. van Gisbergen, C. Fonseca Guerra and E.J. Baerends, J.Comput.Chem. 16, 1511 (2000) and references on the Amsterdam Density Functional (ADF) programme therein.

[8] B.J. Ransil, Int. J. Quant. Chem. 17, 467 (1980)

[9] J.A. Sordo, Comp. Phys. Comm. 113, 85 (1998) and references therein.

[10] G. Maroulis, M. Sana and G. Leroy, Int. J. Quant. Chem. 19, 43 (1981).

[11] G. Maroulis, Int. J. Quant. Chem. 24, 185 (1988).

[12] G. Maroulis, Int. J. Quant. Chem. 55, 173 (1995).

[13] G. Maroulis, J.Chem.Phys. 108, 5432 (1998).

[14] G. Maroulis, J.Chem.Phys. 111, 583 (1999).

[15] G. Maroulis, J.Chem.Phys. 113, 1813 (2000).

[16] P. Karamanis and G. Maroulis, J. Mol. Struct. (THEOCHEM) 621, 157 (2003).

[17] P. Karamanis and G. Maroulis, Chem. Phys. Lett. 376, 403 (2003).

[18] P. Karamanis and G. Maroulis, in T.E. Simos (Ed.), *Computational Methods in Sciences and Engineering 2003* (World Scientific, Singapore, 2003), pp 285-288.

[19] J.L. Brédas, C. Adant, P. Tackx and A. Persoons, Chem. Rev. 94, 243 (1994).

[20] H.S. Nalwa and S. Miyata (Eds), *Nonlinear Optics of Organic Molecules and Polymers* (CRC Press, Boca Raton, 1997).

[21] A.D. Buckingham, Adv. Chem. Phys. 12, 107 (1967).

[22] G. Maroulis, J. Chem. Phys. 118, 2673 (2003).

[23] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M.W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople, GAUSSIAN 98, Revision A.7 (Gaussian, Inc., Pittsburgh PA, 1998) and references therein.

[24] G. Chatrand and L. Lesniak, *Graphs and Digraphs* (Wadsworth, Belmont, 1986).

[25] H.Spath, *Cluster Analysis Algorithms* (Ellis Horwood, Chichester, 1980).

[26] See reference 17. All three basis sets are available from the corresponding author.

[27] Conversion factors to SI units are: energy, 1 $E_h = 4.3597482 \times 10^{-18}$ J, Θ 1 $ea_0^2 = 4.486554 \times 10^{-40} Cm^2$, Φ 1 $ea_0^4 = 1.256363 \times 10^{-60}$ Cm^4, α 1 $e^2a_0^2E_h^{-1} = 1.648778 \times 10^{-41}$ $C^2m^2J^{-1}$ and γ 1 $e^4a_0^4E_h^{-3} = 6.235378 \times 10^{-65}$ $C^4m^4J^{-3}$.